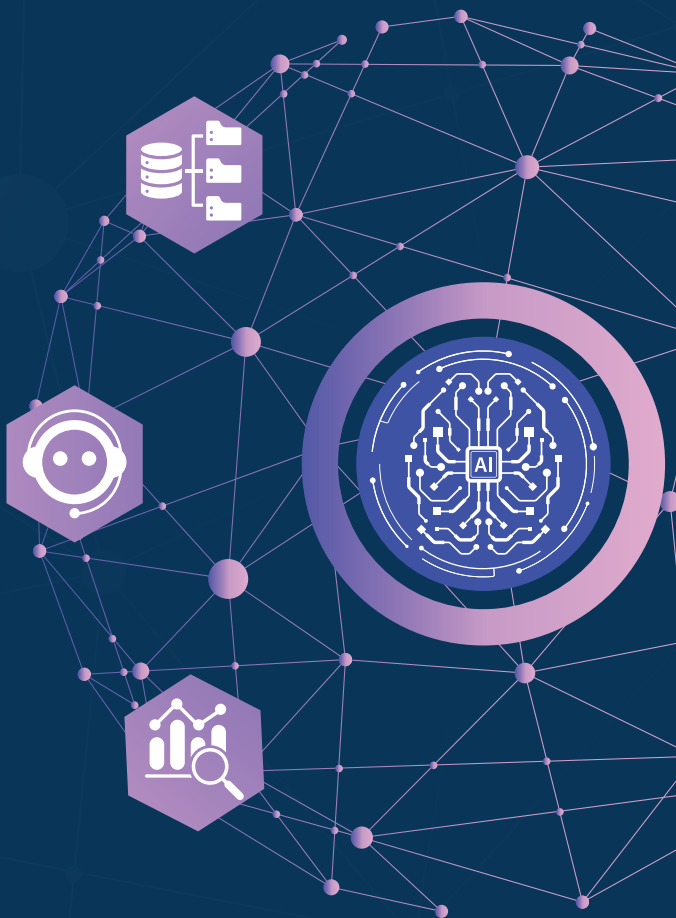
TS. NGUYỄN THÀNH TRUNG

# ĐỊA PHƯƠNG
## TRONG NỀN KINH TẾ
# TÁC NHÂN

Từ thương hiệu
đến chủ quyền thuật toán



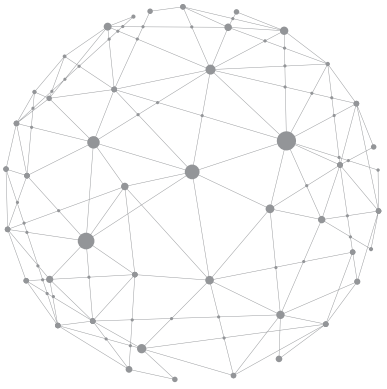

NHÀ XUẤT BẢN
CÔNG THƯƠNG

TS. NGUYỄN THÀNH TRUNG

# ĐỊA PHƯƠNG TRONG NỀN KINH TẾ TÁC NHÂN

## Từ thương hiệu đến chủ quyền thuật toán

NHÀ XUẤT BẢN CÔNG THƯƠNG



# LỜI NÓI ĐẦU

Tháng 8 năm 1998, hai nghiên cứu sinh của Đại học Stanford thành lập một công ty nhỏ trong gara của một ngôi nhà thuê tại Menlo Park. Ý tưởng của họ nghe có vẻ quá đơn giản, nếu internet ngày càng lớn, con người sẽ cần một cách tốt hơn để tìm thấy điều mình muốn. Larry Page và Sergey Brin không phát minh ra internet, cũng không tạo ra thêm thông tin. Họ chỉ làm một việc là giảm chi phí tìm kiếm thông tin. Thành công của Google không đến từ việc tạo ra một nhu cầu mới. Con người từ lâu đã muốn biết điều mình chưa biết. Google chỉ khiến việc đó trở nên nhanh hơn rất nhiều.

Câu chuyện ấy lặp lại nhiều lần trong lịch sử internet. Amazon không phát minh ra mua bán, họ giảm chi phí tìm và mua một món hàng trong số hàng triệu món hàng. Airbnb không phát minh ra khách sạn, họ giảm chi phí xây dựng niềm tin giữa hai người hoàn toàn xa lạ. Nhìn từ góc độ ấy, lịch sử của các nền tảng số không phải là lịch sử của công nghệ. Nó gần hơn với lịch sử của việc liên tục cắt giảm những loại chi phí vô hình mà con người phải trả để đạt được một điều mình vốn luôn mong muốn.

Nhưng internet, sau gần ba mươi năm phát triển, vẫn để lại một loại chi phí rất lớn. Đó là chi phí ra quyết định. Một người chuẩn bị đến Quảng Ninh vẫn phải đọc hàng trăm đánh giá khách sạn. Một gia đình muốn cho con đi du học vẫn phải mở hàng chục website để so sánh các trường. Khi mọi thứ đều có thể được tìm thấy, điều khó nhất không còn là tìm kiếm mà là lựa chọn.

Nhà kinh tế học Herbert Simon từng nhận xét rằng, một thế giới giàu thông tin sẽ tiêu thụ một thứ ngày càng khan hiếm, là sự chú ý của con người. Chúng ta không thiếu thông tin. Chúng ta thiếu thời gian để đọc, năng lực để đánh giá và sự chắc chắn để đưa ra quyết định. Đó chính là khoảng trống mà các tác nhân trí tuệ nhân tạo xuất hiện.

Nếu trước đây con người sử dụng internet để thu thập thông tin rồi tự mình cân nhắc, thì ngày nay họ ngày càng có xu hướng giao cả quá trình ấy cho một hệ thống có khả năng đọc, so sánh, suy luận và hành động thay mình. Đó là sự khác biệt căn bản giữa công cụ tìm kiếm và tác nhân trí tuệ nhân tạo. Một công cụ tìm kiếm trả lời câu hỏi: "Thông tin ở đâu?". Một tác nhân AI trả lời câu hỏi khác: "Theo tất cả những gì tôi biết về anh, tôi nghĩ anh nên chọn phương án này". Sự khác biệt ấy nghe có vẻ nhỏ. Nhưng về kinh tế học, đó là sự thay đổi của chủ thể thực hiện việc lựa chọn.

Con người trực tiếp tìm kiếm, trực tiếp so sánh và trực tiếp quyết định. Internet giúp họ làm những việc ấy hiệu quả hơn nhưng không thay họ thực hiện. Việc ủy nhiệm một phần quyết định cho người khác thì không mới, con người từ lâu đã ủy nhiệm cho chuyên gia, đại lý, cố vấn và bộ máy hành chính. Cái mới là lần đầu tiên việc ủy nhiệm một phần quá trình lựa chọn cho hệ thống máy có khả năng diễn ra ở quy mô đại chúng, với chi

phí biên gần bằng không. Con người vẫn chịu trách nhiệm cuối cùng, nhưng không còn nhất thiết phải đọc hàng trăm đánh giá, so sánh hàng chục phương án hay tự mình xây dựng mọi lịch trình. Quyền lựa chọn không biến mất, nó được ủy nhiệm.

Chính từ sự chuyển dịch ấy mà khái niệm nền kinh tế tác nhân được hình thành. Đây là một nền kinh tế trong đó ngày càng nhiều quyết định được thực hiện thông qua các tác nhân trí tuệ nhân tạo thay vì trực tiếp bởi con người. Những tác nhân ấy không chỉ trả lời câu hỏi mà còn có thể tìm kiếm, thương lượng, đặt chỗ, thanh toán và phối hợp với các hệ thống khác để hoàn thành một mục tiêu. Không phải ngẫu nhiên mà cuộc cạnh tranh lớn nhất của giới công nghệ trong thập kỷ này không còn là cuộc đua tạo ra mô hình mạnh nhất, mà là cuộc đua giành quyền trở thành người đại diện của con người trong các quyết định hằng ngày. Sự chuyển dịch này đặt ra một câu hỏi hoàn toàn mới cho các địa phương.

Trong hơn nửa thế kỷ qua, gần như toàn bộ lý thuyết về marketing địa phương, thương hiệu điểm đến và phát triển du lịch đều dựa trên cùng một giả định là người đưa ra quyết định cuối cùng là con người. Vì thế, các địa phương học cách kể những câu chuyện hấp dẫn hơn, thiết kế những trải nghiệm đáng nhớ hơn, xây dựng những biểu tượng dễ nhận biết hơn và truyền thông hiệu quả hơn. Tất cả những nỗ lực ấy đều nhằm trả lời một câu hỏi: "Làm thế nào để một con người yêu một nơi chốn?".

Nhưng nếu người đọc đầu tiên của một địa phương không còn là con người thì sao? Một tác nhân AI không đi dọc phố cổ Hội An để ngửi mùi gỗ sau cơn mưa. Nó không cảm nhận được sự tĩnh lặng của một ngôi chùa ở Ninh Bình, không rung động trước tiếng chuông nhà thờ ở Hà Nội và cũng không lưu giữ ký ức

về một bữa cơm gia đình ở miền Tây Việt Nam. Máy không trải nghiệm địa phương như con người. Máy đọc địa phương thông qua dữ liệu. Nó đọc tên gọi, địa chỉ, mối quan hệ giữa các thực thể, lịch mở cửa, khả năng đặt chỗ, mức độ tin cậy của nguồn thông tin, khả năng giao dịch và hàng triệu tín hiệu khác mà phần lớn con người không bao giờ nhìn thấy.

Ở đây xuất hiện một nghịch lý mới. Trong nền kinh tế trải nghiệm, địa phương phải đủ khác biệt để con người ghi nhớ. Trong nền kinh tế tác nhân, địa phương lại phải đủ khả đọc để máy hiểu. Càng khác biệt, đôi khi càng khó được máy biểu diễn. Càng chuẩn hóa để máy dễ đọc, đôi khi lại càng đánh mất những sắc thái tạo nên bản sắc. Cuộc cạnh tranh của các địa phương vì thế không còn chỉ diễn ra trong tâm trí con người mà còn diễn ra bên trong các hệ thống suy luận.

Trong nền kinh tế tác nhân, bên cạnh sự chú ý, một điểm nghẽn mới đang nổi lên, khả năng được hệ thống suy luận đưa vào tập cân nhắc. Một địa phương có thể có hàng nghìn bài báo, hàng triệu lượt xem trên mạng xã hội và hàng chục giải thưởng quốc tế, nhưng nếu những thông tin ấy phân tán, thiếu cấu trúc, mâu thuẫn hoặc không thể liên kết với nhau, chúng vẫn rất khó trở thành tri thức đối với máy. Muốn suy luận ổn định, máy phải có dữ liệu đủ khả truy hồi, đối chiếu và liên kết, trong đó tầng định danh, trạng thái và giao dịch cần có cấu trúc. Lợi thế cạnh tranh của một địa phương vì thế không còn chỉ nằm ở những gì nó sở hữu mà còn nằm ở cách những tài sản ấy được biểu diễn trong không gian số.

Đây là lý do cuốn sách sử dụng khái niệm tính khả đọc thuật toán thay vì những thuật ngữ quen thuộc như chuyển đổi số hay hiện diện trực tuyến. Một địa phương có thể đã chuyển đổi số

rất nhiều nhưng vẫn không khả đọc, hàng trăm trang web, hàng nghìn tài liệu PDF và vô số bài viết trên mạng xã hội tồn tại như những hòn đảo rời rạc. Đối với con người, những mảnh ghép ấy có thể vẫn đủ để hiểu một nơi. Đối với các hệ thống suy luận, chúng giống như hàng triệu mảnh giấy chưa bao giờ được sắp xếp thành một cuốn sách. Khoảng cách giữa có dữ liệu và được máy hiểu có thể trở thành khoảng cách quyết định năng lực cạnh tranh của một địa phương.

Chính từ nghịch lý ấy mà cuốn sách "**Địa phương trong nền kinh tế tác nhân**" được viết. Cuốn sách này không bàn về một mô hình trí tuệ nhân tạo cụ thể. ChatGPT rồi sẽ thay đổi. Gemini, Claude hay những hệ thống khác cũng sẽ thay đổi. Điều cuốn sách quan tâm là một sự chuyển dịch có khả năng kéo dài nhiều thập kỷ, sự thay đổi của chủ thể lựa chọn và những hệ quả của nó đối với thương hiệu, du lịch, đầu tư, phát triển địa phương và năng lực cạnh tranh quốc gia.

Nếu "**Địa phương trong nền kinh tế trải nghiệm**" nghiên cứu cách một địa phương tích lũy giá trị trong ký ức của con người thì cuốn sách này nghiên cứu cách một địa phương hiện diện trong ký ức của máy. Nếu cuốn "Địa phương trong nền kinh tế trải nghiệm" sử dụng hình ảnh của phù sa để giải thích sự tích lũy chậm nhưng bền vững của thương hiệu thì cuốn sách này sử dụng hình ảnh của bộ lọc, dòng chảy và van điều tiết để giải thích cách các hệ thống trí tuệ nhân tạo nhìn thấy thế giới. Hai cuốn không đối lập nhau. Chúng mô tả hai lực kéo đang đồng thời định hình tương lai của mọi địa phương, một lực kéo hướng về chiều sâu của trải nghiệm và một lực kéo hướng về tính khả đọc của dữ liệu.

Ba chương đầu của cuốn sách sẽ trả lời câu hỏi điều gì đã thực sự thay đổi, từ sự xuất hiện của nền kinh tế tác nhân đến cách

máy hình thành hiểu biết về một địa phương. Ba chương tiếp theo phân tích những hệ quả kinh tế của sự thay đổi ấy đối với niềm tin, dòng khách, giá trị và chuỗi lợi nhuận. Ba chương cuối cùng chuyển sang câu hỏi thực hành, địa phương cần được thiết kế như thế nào để vừa giữ được bản sắc trước con người, vừa có thể hiện diện trước các tác nhân trí tuệ nhân tạo, đồng thời bảo vệ chủ quyền của mình trong một thế giới mà hình ảnh của mỗi nơi ngày càng được kiến tạo bởi các hệ thống suy luận.

Cuốn sách không xem trí tuệ nhân tạo là chương cuối của chuyển đổi số. Ngược lại, AI chỉ là biểu hiện rõ nhất của một sự thay đổi sâu hơn, sự xuất hiện của một lớp hạ tầng mới, nơi ngày càng nhiều giao dịch kinh tế bắt đầu bằng quá trình suy luận của máy. Nếu đường sắt định hình địa lý kinh tế của thế kỷ XIX, điện định hình công nghiệp của thế kỷ XX và internet định hình nền kinh tế nền tảng đầu thế kỷ XXI, thì các tác nhân trí tuệ nhân tạo có thể sẽ định hình cách các địa phương được nhìn thấy, được hiểu và được lựa chọn trong nhiều thập kỷ tới. Câu hỏi không còn là địa phương có ứng dụng AI hay không. Câu hỏi là địa phương có đủ năng lực để tồn tại trong quá trình suy luận của các tác nhân AI hay không. Đó chính là câu hỏi mà chín chương tiếp theo sẽ cùng nhau tìm cách trả lời.

**TS. NGUYỄN THÀNH TRUNG**

# MỤC LỤC

*Lời nói đầu* *3*
*Giới thiệu nội dung các chương* *22*
*Danh mục từ viết tắt* *24*

**CHƯƠNG 1: TỪ TRẢI NGHIỆM ĐẾN ỦY NHIỆM** **31**
1.1. NƠI CUỐN SÁCH TRƯỚC DỪNG LẠI 34
1.1.1. Từ phù sa sang bộ lọc 34
1.1.2. Khi trải nghiệm trở thành lợi thế cạnh tranh 35
1.1.3. Điều chưa từng được đặt ra 36
1.2. KHI TÌM KIẾM BIẾN THÀNH ỦY NHIỆM 37
1.2.1. Bốn phút, ba phương án 37
1.2.2. Hai hợp đồng khác nhau 38
1.2.3. Không chỉ là chuyện du lịch 39
1.2.4. Tác nhân không trả lời, tác nhân hành động 40
1.3. NỀN KINH TẾ TÁC NHÂN LÀ GÌ 42
1.3.1. Từ internet đến nền kinh tế tác nhân 42
1.3.2. Tác nhân khác trợ lý hội thoại ở đâu 43
1.3.3. Chuỗi ra quyết định mới 45
1.3.4. Từ kinh tế chú ý sang kinh tế ủy nhiệm 47
1.4. KHI MÁY TRỞ THÀNH NGƯỜI ĐỌC ĐẦU TIÊN 48
1.4.1. Ba lần thay đổi của chủ thể đọc 48
1.4.2. Máy đọc thế giới như thế nào 49
1.4.3. Điều máy đọc được 50
1.4.4. Điều máy không đọc được 50
1.4.5. Nghịch lý Wikipedia 52
1.5. TRẢI NGHIỆM VÀ KHẢ ĐỌC: HAI LOGIC NGƯỢC CHIỀU 55
1.5.1. Cuộc tranh luận về lược đồ 55
1.5.2. Trải nghiệm là thứ khó số hóa nhất 57
1.5.3. Bản sắc càng sâu càng khó khả đọc 58
1.5.4. Nghịch lý của địa phương 59

1.6. CHÚNG TA ĐỨNG Ở ĐÂU TRONG BƯỚC CHUYỂN NÀY 61
1.6.1. Những con số tăng trưởng dựa trên hệ thống cũ 61
1.6.2. Nơi quyết định thực sự được đưa ra 61
1.6.3. Bài học từ thị trường Nga: Tăng trưởng đến từ đường bay 62
1.6.4. Vì sao thành công hiện tại làm chậm phản ứng 63
1.7. MỘT CUỘC CẠNH TRANH HOÀN TOÀN MỚI 65
1.7.1. Từ cạnh tranh điểm đến sang cạnh tranh khả đọc 65
1.7.2. Địa phương sẽ phải quản trị điều gì 66
1.7.3. Khả đọc trở thành hạ tầng mới 67
1.7.4. Khi khả đọc là công cụ phòng vệ 68
1.7.5. Dẫn nhập cho các chương tiếp theo 70

**CHƯƠNG 2: TÍNH KHẢ ĐỌC THUẬT TOÁN 71**
2.1. MỘT SỰ VẮNG MẶT KHÔNG AI BÁO CÁO 73
2.1.1. Khi quyết định bắt đầu trước khi con người tìm kiếm 73
2.1.2. Sự biến mất khỏi tập lựa chọn 74
2.1.3. Ba đặc tính của sự vô hình 75
2.1.4. Đặt tên cho vấn đề 77
2.2. PHẢ HỆ CỦA TÍNH KHẢ ĐỌC: SCOTT VÀ BỘ MÁY BIẾT ĐỌC 78
2.2.1. Cái nhìn của nhà nước 78
2.2.2. Vì sao mọi bộ máy đều cần một thế giới khả đọc 79
2.2.3. Khi máy trở thành bộ máy đọc mới 81
2.2.4. Từ nhà nước đến thuật toán 82
2.2.5. Hai đại dự án và một bài học 83
2.2.6. Một khác biệt quan trọng của thời hiện tại 85
2.3. PHÂN LOẠI VÀ LÃNG QUÊN: BOWKER, STAR, FOUCAULT 88
2.3.1. Không có phân loại nào là trung tính 88
2.3.2. Hai cơ chế lãng quên 89
2.3.3. Quyền lực tạo ra tri thức 90
2.3.4. Khả đọc luôn là một quyết định chính trị 92
2.3.5. Ai đang giữ quyền mô tả một vùng đất 92
2.3.6. Một lược đồ không có ô cho cái gì 94

2.4. MÔ HÌNH BỐN BỘ LỌC, MỘT VÒNG, MỘT VAN 97
2.4.1. Định nghĩa 97
2.4.2. Bộ lọc thứ nhất, phân giải 98
2.4.3. Bộ lọc thứ hai, chứng thực 101
2.4.4. Bộ lọc thứ ba, xếp hạng 102
2.4.5. Bộ lọc thứ tư, thực thi 102
2.4.6. Cái vòng: ký ức thuật toán 104
2.4.7. Cái van: quyền mờ 105
2.4.8. Bốn bộ lọc nhìn cùng một lúc 107
2.4.9. Mô hình này có thể sai ở đâu 108
2.5. KHI THẾ GIỚI TỰ SỬA MÌNH THEO LƯỢC ĐỒ 111
2.5.1. Phát biểu mệnh đề 111
2.5.2. Bốn bộ lọc và bốn phản ứng 113
2.5.3. Cùng một mệnh đề, năm lĩnh vực 114
2.5.4. Vì sao địa phương là trường hợp khó nhất 115
2.5.5. Ba câu hỏi mệnh đề này đặt ra mà lý thuyết cũ không đặt 116
2.6. PHẢN BIỆN MÔ HÌNH 118
2.6.1. Máy đang đọc dữ liệu phi cấu trúc tốt hơn,
lược đồ còn quan trọng không 118
2.6.2. Danh tiếng có thay thế khả đọc không 119
2.6.3. Khả đọc không đồng nghĩa với dữ liệu mở 119
2.6.4. Giá trị biên của khả đọc 121
2.6.5. Phản biện thứ năm:
mô hình này có mô tả đúng Việt Nam không 122
2.6.6. Sáu phản bác từ phía công nghệ 126
2.6.7. Ba phản bác có thể làm suy yếu luận đề 130
2.7. NGHỊCH LÝ SCOTT VÀ CÁI GIÁ CỦA MỘT TRẢI NGHIỆM KHẢ ĐỌC 134
2.7.1. Mỗi lần được đọc là một lần bị làm phẳng 134
2.7.2. Khả đọc làm nghèo trải nghiệm như thế nào 135
2.7.3. Nghịch lý trung tâm của cuốn sách 136
2.7.4. Chủ quyền thuật toán 136

2.8. BẢN ĐỒ PHẦN CÒN LẠI CỦA CUỐN SÁCH 139
2.8.1. Chương 3 trả lời câu hỏi máy đọc cái gì 139
2.8.2. Chương 4 và 5 trả lời câu hỏi làm thế nào để khả đọc 139
2.8.3. Chương 6 tới 9 chuyển sang tầng quản trị 140
2.8.4. Cách đọc phần còn lại 141

**CHƯƠNG 3: MÁY ĐỌC ĐỊA PHƯƠNG BẰNG KIẾN TRÚC GÌ? 143**
3.1. BỐN RÀNG BUỘC SINH RA BỐN BẤT BIẾN 147
3.1.1. Bốn ràng buộc 147
3.1.2. Vòng sơ loại bằng phản ví dụ 148
3.1.3. Quan hệ tích giữa bốn thao tác 149
3.1.4. Ba làn sóng khả đọc 150
3.2. RÀNG BUỘC THỨ NHẤT: BỘ NHỚ HỮU HẠN 153
3.2.1. Chứng minh 153
3.2.2. Phản ví dụ 153
3.2.3. Phản biện 154
3.2.4. Hệ quả: dữ liệu không phải ngữ liệu 156
3.2.5. Điều không bao giờ vào được bộ nhớ 157
3.2.6. Thí nghiệm với Huế, khâu thứ nhất 158
3.3. RÀNG BUỘC THỨ HAI: YÊU CẦU ĐỒNG ĐO 160
3.3.1. Chứng minh 160
3.3.2. Phản ví dụ 161
3.3.3. Phản biện 162
3.3.4. Hệ quả thứ nhất: mọi thực thể đều bị nén 163
3.3.5. Hệ quả thứ hai: phần dễ nén là phần ít phân biệt 163
3.3.6. Hệ quả thứ ba: địa phương là một tập hợp thực thể 164
3.3.7. Biểu diễn là một vấn đề triết học 165
3.3.8. Thí nghiệm Huế, khâu thứ hai 166
3.4. RÀNG BUỘC THỨ BA: CHÚ Ý KHAN HIẾM VÀ THÔNG TIN BẤT CÂN XỨNG 168
3.4.1. Chứng minh 168
3.4.2. Phản ví dụ 169

3.4.3. Năm nhóm tín hiệu 170
3.4.4. Phản biện 171
3.4.5. Hệ quả thứ nhất: ngưỡng cắt 172
3.4.6. Hệ quả thứ hai: ưu thế tự tích lũy 172
3.4.7. Thí nghiệm Huế, khâu thứ ba 173
3.5. RÀNG BUỘC THỨ TƯ: CHI PHÍ GIAO DỊCH 176
3.5.1. Chứng minh 176
3.5.2. Phản ví dụ 176
3.5.3. Một dự báo kiểm chứng được 177
3.5.4. Phản biện 178
3.5.5. Hệ quả: khả đọc không còn đủ 179
3.5.6. Thí nghiệm Huế, khâu thứ tư 180
3.6. ĐIỀU KIỆN TIỀN ĐỀ: BẢN THỂ LUẬN 182
3.6.1. Chứng minh 182
3.6.2. Phản ví dụ 183
3.6.3. Ai viết bản thể luận 184
3.6.4. Mỗi bản thể luận là một thế giới quan 185
3.6.5. Phản biện 186
3.6.6. Thí nghiệm với Huế, điều kiện tiền đề 187
3.7. NGUYÊN LÝ KHẢ ĐỌC THUẬT TOÁN 188
3.7.1. Hành trình mà cỗ máy trả về 189

**CHƯƠNG 4: NIỀM TIN VÀ KÝ ỨC CỦA MÁY 191**
***Phần A: Máy hình thành tri thức bằng cách nào 196***
4.1. PHÂN GIẢI THỰC THỂ: BÀI TOÁN CỦA MỘT CÁI TÊN 196
4.1.1. Không có thực thể thì không có tri thức 196
4.1.2. Một cái tên không đủ để định danh 198
4.1.3. Thực thể được ghép từ những gì 199
4.1.4. Sai ở bước phân giải sẽ sai toàn bộ chuỗi 203
4.1.5. Địa phương phải tự giải bài toán phân giải 205
4.2. XUẤT XỨ, ĐỐI CHIẾU VÀ THẨM QUYỀN 208
4.2.1. Muốn biết điều gì đúng, trước hết phải biết nó đến từ đâu 209

4.2.2. Một thông tin mạnh vì nhiều nguồn cùng xác nhận 211
4.2.3. Mọi nguồn đều có trọng số 214
4.2.4. Thẩm quyền là tài sản tích lũy 217
4.2.5. Mâu thuẫn giữa các nguồn được giải quyết như thế nào 219
4.3. KHI MÁY TỰ VIẾT CÂU MÔ TẢ 224
4.3.1. Từ truy hồi sang tổng hợp 225
4.3.2. Mô tả là sản phẩm của nhiều nguồn 227
4.3.3. Vì sao AI khác Wikipedia 228
4.3.4. Khi sai lệch nhỏ trở thành mô tả chính thức 230
4.3.5. Địa phương mất quyền kể câu chuyện của mình như thế nào 231
***Phần B: Tri thức ấy biến đổi như thế nào*** ***234***
4.4. VÒNG LẶP THAY CHO ĐƯỜNG THẲNG 234
4.4.1. Mỗi lần được đọc là một lần được củng cố 237
4.4.2. Bộ nhớ thuật toán khác bộ nhớ con người 240
4.4.3. Vì sao vòng lặp luôn khuếch đại 242
4.4.4. Một mô hình động của niềm tin 244
4.5. KÝ ỨC THUẬT TOÁN VÀ HIỆN TƯỢNG KHÓA CHẶT 247
4.5.1. Ký ức thuật toán là gì 248
4.5.2. Đường dẫn tạo nên tương lai 249
4.5.3. Hiệu ứng khóa chặt 251
4.5.4. Địa phương cũng bị khóa chặt 252
4.5.5. Ký ức mạnh hơn sự thật hiện tại 254
4.6. ĐỊNH VỊ LẠI TRONG MỘT HỆ THỐNG CÓ KÝ ỨC 256
4.6.1. Thay đổi thương hiệu dễ hơn thay đổi ký ức 257
4.6.2. Vì sao tái định vị luôn tốn kém 258
4.6.3. Một thuộc tính sai kéo theo hàng trăm mô tả sai 260
4.6.4. Quản trị thương hiệu trở thành quản trị ký ức 261
4.7. SỬA SỚM HAY KHÔNG SỬA ĐƯỢC 263
4.7.1. Sai lúc mới xuất hiện dễ sửa nhất 263
4.7.2. Sau khi được trích dẫn nhiều lần, chi phí tăng phi tuyến 265
4.7.3. Khi AI đã học từ dữ liệu cũ 268
4.7.4. Chiến lược sửa sai của địa phương 269
4.8. KẾT LUẬN CHƯƠNG 273

**CHƯƠNG 5: CUỘC CẠNH TRANH GIÀNH QUYỀN ĐƯỢC ĐỀ XUẤT 275**
***Phần A: Quyền đề xuất được hình thành thế nào 281***
5.1. KINH TẾ HỌC CỦA SỰ ĐỀ XUẤT 281
5.1.1. Từ được nhìn thấy đến được đề xuất 281
5.1.2. Giá trị kinh tế của một vị trí 285
5.1.3. Mọi đề xuất đều là một quyết định phân bổ 288
5.1.4. Khi AI trở thành nhà môi giới lớn nhất 290
5.1.5. Kinh tế học của xác suất 292
5.2. ĐỊA PHƯƠNG CẠNH TRANH TRONG TẬP ỨNG VIÊN 294
5.2.1. Một nhóm nhỏ được thấy, phần lớn bị loại trước khi xuất hiện 295
5.2.2. Không còn cạnh tranh với hàng xóm 296
5.2.3. Địa phương cạnh tranh bằng tín hiệu 298
5.2.4. Danh tiếng trở thành dữ liệu 301
5.2.5. Cuộc đua tích lũy niềm tin 303
5.3. NỀN TẢNG CẠNH TRANH ĐỂ ĐỊNH NGHĨA TIÊU CHUẨN 306
5.3.1. Quyền đề xuất không còn tập trung ở một cổng 307
5.3.2. Mỗi nền tảng có một logic tối ưu 308
5.3.3. Người dùng không còn nhìn thấy quá trình lựa chọn 310
5.3.4. Bốn loại quyền lực trong hệ sinh thái đề xuất 311
5.3.5. Không phải một chuẩn thắng mà nhiều lớp chuẩn 314
***Phần B: Vì sao lợi thế tự khuếch đại? 318***
5.4. HIỆU ỨNG MATTHEW CỦA THUẬT TOÁN 318
5.4.1. Giàu càng giàu trong thế giới số 318
5.4.2. Được chọn tạo ra thêm lý do để được chọn 320
5.4.3. Thiên lệch theo mức độ phổ biến 321
5.4.4. Bốn dạng thiếu dữ liệu của địa phương mới 323
5.4.5. Khi bất bình đẳng tự khuếch đại 326
5.5. TỪ TỐI ƯU CỤC BỘ ĐẾN TẬP TRUNG HỆ THỐNG 328
5.5.1. Tối ưu cục bộ 328
5.5.2. Thất bại ở quy mô hệ thống 329
5.5.3. Hội tụ đề xuất ở cấp hệ thống 330
5.5.4. Cân bằng tự củng cố không tối ưu xã hội 332

5.5.5. Nghịch lý của nền kinh tế tác nhân: Máy càng thông minh, đám đông càng tập trung 335
*Phần C: Khi đề xuất trở thành áp lực lên nơi chốn* *337*
5.6. QUÁ TẢI DU LỊCH ĐƯỢC THUẬT TOÁN KHUẾCH ĐẠI 337
5.6.1. Quá tải du lịch trong điều kiện có tác nhân 338
5.6.2. Cơ chế tác động lên cư dân 339
5.6.3. Suy giảm trải nghiệm và độ trễ của phản hồi thuật toán 341
5.7. ĐIỀU TIẾT Ở ĐÂU TRONG CHUỖI QUYẾT ĐỊNH 344
5.7.1. Venice và bài toán phí vào cửa 345
5.7.2. Kyoto và sự phản ứng của cư dân 347
5.7.3. Giới hạn của giá trong điều tiết đề xuất 349
5.7.4. Ba tầng can thiệp vào chuỗi quyết định 351
5.8. KẾT LUẬN CHƯƠNG 359

**CHƯƠNG 6: AI GIỮ LẠI GIÁ TRỊ** **361**
6.1. CHUỖI GIÁ TRỊ BỊ CẮT Ở QUYỀN LỰA CHỌN VÀ GIAO DỊCH 365
6.1.1. Chuỗi trải nghiệm truyền thống 365
6.1.2. Tác nhân chen vào ở đâu 366
6.1.3. Điểm cắt đầu tiên là quyền lựa chọn 371
6.1.4. Điểm cắt thứ hai là giao diện giao dịch 372
6.1.5. Từ tạo giá trị đến giữ giá trị 373
6.2. TÔ THUẬT TOÁN VÀ QUYỀN LỰC ĐIỂM NGHẼN 376
6.2.1. Điểm nghẽn tạo ra khả năng thu tô 377
6.2.2. Thuật toán trở thành hạ tầng thu phí 378
6.2.3. Phí dịch vụ, lợi nhuận và tô là ba thứ khác nhau 382
6.2.4. Dữ liệu và lãi kép của quyền lực 383
6.2.5. Tô thuật toán khác gì tô đất đai 386
6.3. CHI PHÍ THAM GIA NỀN KINH TẾ TÁC NHÂN 389
6.3.1. Ba lớp chi phí tham gia 391
6.3.2. Không đầu tư cũng mất thị trường 395
6.3.3. Chi phí này mang tính lũy thoái 396
6.3.4. Từ marketing sang hạ tầng dữ liệu 399
6.3.5. Ngưỡng chi phí cố định mới 402

6.4. KHI GIAO DIỆN NÉN NƠI CHỐN THÀNH THUỘC TÍNH 404
6.4.1. Giao diện làm phẳng khác biệt 405
6.4.2. Từ nơi chốn thành một sản phẩm 406
6.4.3. Thuộc tính nào không được biểu diễn sẽ biến mất 408
6.4.4. Khi AI tối ưu theo tiêu chí giống nhau 409
6.4.5. Hàng hóa hóa dẫn tới cạnh tranh giá 411
6.5. ĐƯỜNG CONG LỢI NHUẬN MỚI 414
6.5.1. Giá trị dịch chuyển khỏi nơi tạo trải nghiệm 414
6.5.2. Hình dạng mới của đường cong lợi nhuận 417
6.6. NGUY CƠ OEM HÓA ĐỊA PHƯƠNG 420
6.6.1. Mất giao diện, mất dần dữ liệu khách hàng 421
6.6.2. Quyền định giá bị thu hẹp 423
6.6.3. Từ nhà điều hành thành nhà cung cấp phía sau 424
6.6.4. Từ khách sạn gia công đến địa phương gia công 424
6.7. AI TRẢ TIỀN CHO HẠ TẦNG KHẢ ĐỌC VÀ KHẢ HÀNH ĐỘNG 427
6.7.1. Phần nào mang tính hạ tầng dùng chung, phần nào là đầu tư tư 428
6.7.2. Vì sao thị trường sẽ đầu tư thiếu 430
6.7.3. Ba mô hình chia sẻ chi phí 431
6.7.4. Ai hưởng lợi và ai có nghĩa vụ đóng góp 435
6.7.5. Đo năng lực giữ giá trị 435
6.8. KẾT LUẬN CHƯƠNG 439

**CHƯƠNG 7: THIẾT KẾ ĐỊA PHƯƠNG CHO MÁY 441**
7.1. ĐIỂM ĐẾN SẴN SÀNG CHO MÁY NGHĨA LÀ GÌ 446
7.1.1. Từ "được biết đến" sang "được giao việc" 448
7.1.2. Bốn cấp độ trưởng thành của một điểm đến 450
7.1.3. Bộ lọc cuối cùng 452
7.2. MÔ HÌNH VẬN HÀNH CỦA ĐỊA PHƯƠNG Ở DẠNG DỮ LIỆU 457
7.2.1. Một địa phương không phải trang web mà là mô hình dữ liệu 458
7.2.2. Từ sơ đồ tổ chức sang sơ đồ tri thức 460
7.2.3. Địa phương như một hệ điều hành 462
7.2.4. Vì sao dữ liệu phải phản ánh vận hành thật 464

7.3. ĐỊA PHƯƠNG NHƯ MỘT ĐỒ THỊ 466
7.3.1. Thực thể sống trong mạng lưới 467
7.3.2. Quan hệ tạo ra ngữ nghĩa 469
7.3.3. Khi hành động, quan hệ thường quan trọng hơn thuộc tính 472
7.3.4. Điểm đứt của đồ thị 475
7.3.5. Khi cả địa phương trở thành đồ thị tri thức 480
7.4. TỒN KHO CÓ THỂ GIAO DỊCH 484
7.4.1. Máy không mua trải nghiệm, máy mua tồn kho 486
7.4.2. Điều gì thực sự được bán: phòng, ghế, vé, suất, khung giờ 487
7.4.3. Tồn kho phải sống: khả dụng theo thời gian thực 488
7.4.4. Giá là một trạng thái của tồn kho 489
7.5. TỪ TỒN KHO ĐẾN GIAO DỊCH 494
7.5.1. Từ khả đọc bằng máy đến khả đặt bằng máy 495
7.5.2. API là cánh cửa, không phải mục tiêu 496
7.5.3. Bốn điều kiện để tác nhân dám giao dịch 499
7.5.4. Khi mọi dịch vụ đều trở thành một hàm có thể gọi 500
7.5.5. Địa phương không còn bán thông tin, địa phương bán khả năng thực hiện 503
7.6. BẢNG ĐIỀU KHIỂN KHẢ HÀNH ĐỘNG 506
7.6.1. Bảng điều khiển kết quả khác bảng điều khiển khả hành động 508
7.6.2. Năm tín hiệu phải theo dõi hằng ngày 510
7.6.3. Chỉ số suy giảm khả hành động: khi địa phương bắt đầu biến mất 513
7.6.4. Điều hành bằng tín hiệu, không bằng chiến dịch 516
7.7. KIỂM TOÁN ĐỊA PHƯƠNG BẰNG CHÍNH HỆ THỐNG SUY LUẬN 519
7.7.1. Máy đang mô tả địa phương như thế nào 522
7.7.2. Kiểm toán đa tác nhân 524
7.7.3. Phát hiện khoảng cách bản sắc 527
7.7.4. Từ kiểm toán sang sửa bằng chứng gốc 529
7.7.5. Cơ chế liên thông tín hiệu địa phương 531
7.8. KẾT LUẬN CHƯƠNG 536

**CHƯƠNG 8: VIỆT NAM: DẤU THANH, ĐỊA DANH VÀ NĂNG LỰC CẤP XÃ** **539**
8.1. BÀI TOÁN DẤU THANH VÀ PHÂN MẢNH THỰC THỂ 544
8.1.1. Một địa danh có bao nhiêu tên 547
8.1.2. Dấu thanh không chỉ là dấu thanh 550
8.1.3. Từ viết tắt tạo ra thực thể mới 553
8.1.4. Khi tiếng Anh và tiếng Việt không gặp nhau 554
8.1.5. Kiến trúc tối thiểu của danh tính địa phương 556
8.2. ĐỊA DANH SAU SÁP NHẬP: KHI BẢN ĐỒ ĐỔI MÀ DỮ LIỆU KHÔNG ĐỔI 560
8.2.1. Địa giới đổi nhanh hơn internet 563
8.2.2. Ba lớp dữ liệu của một tỉnh mới 565
8.2.3. Bản đồ, OTA và AI không cập nhật cùng nhịp 568
8.2.4. Một điểm đến sống trong nhiều bản đồ 571
8.2.5. Sau sáp nhập, việc đầu tiên không phải làm logo mà hợp nhất thực thể số 573
8.3. DỮ LIỆU HÀNH CHÍNH VÀ DỮ LIỆU ĐIỂM ĐẾN 577
8.3.1. Dữ liệu quản lý không phải dữ liệu quyết định 578
8.3.2. Từ hồ sơ pháp lý đến hồ sơ trải nghiệm 578
8.3.3. Máy cần gì hơn mã xã 579
8.3.4. Hai lớp dữ liệu, một thực thể 581
8.3.5. Song sinh số phải phản ánh nhịp sống 582
8.4. NĂNG LỰC SỐ Ở CẤP XÃ VÀ CẤP TỈNH 584
8.4.1. Vì sao cấp xã mới là tuyến đầu của AI 585
8.4.2. Người hiểu dữ liệu nhất lại ít được quyền cập nhật 588
8.4.3. Một cán bộ văn hóa có thể làm thay đổi khả đọc của cả địa phương 591
8.4.4. Năng lực số không phải mua phần mềm 593
8.4.5. Một bộ chỉ số tối thiểu cho cấp xã 595
8.5. QUAN HỆ GIỮA CƠ QUAN QUẢN LÝ VÀ DOANH NGHIỆP 601
8.5.1. Chính quyền giữ danh tính 602
8.5.2. Doanh nghiệp giữ tồn kho 603
8.5.3. Nền tảng giữ hành vi 605
8.5.4. Khi không ai là người tích hợp 606
8.5.5. Cơ chế liên thông tín hiệu địa phương 607

8.6. GIÀU DI SẢN, NGHÈO DỮ LIỆU 610
8.6.1. Di sản nhiều không đồng nghĩa được AI biết đến 611
8.6.2. Từ nội dung truy hồi được đến thực thể giao dịch được 612
8.6.3. Địa phương nhỏ có cơ hội mới 613
8.6.4. Việt Nam không thiếu tài nguyên, thiếu lớp biểu diễn 614
8.7. KẾT LUẬN CHƯƠNG 616
*Phụ lục 8a. Danh mục việc cần làm trong mười hai tháng* *617*

**CHƯƠNG 9: CHỦ QUYỀN VÀ TRÁCH NHIỆM 619**
***Phần A: Chủ quyền*** ***625***
9.1. QUYỀN ĐỊNH NGHĨA HÌNH ẢNH ĐỊA PHƯƠNG 625
9.1.1. Hình ảnh địa phương không còn nằm trong brochure 628
9.1.2. Một địa phương, nhiều bản sao trong thế giới của máy 629
9.1.3. Quyền định nghĩa dịch chuyển khỏi chính quyền 630
9.1.4. Từ thương hiệu sang quyền biểu diễn 632
9.1.5. Chủ quyền bắt đầu từ quyền sửa tên mình 633
9.2. CHỦ QUYỀN DỮ LIỆU VÀ CHỦ QUYỀN THUẬT TOÁN LÀ HAI KHÁI NIỆM KHÁC NHAU 636
9.2.1. Có dữ liệu nhưng không có quyền suy luận 639
9.2.2. Chủ quyền dữ liệu trả lời "ai giữ" 639
9.2.3. Chủ quyền thuật toán trả lời "ai quyết định" 640
9.2.4. Quyền lực nằm ở lớp diễn giải 641
9.2.5. Không thể bảo vệ thương hiệu chỉ bằng dữ liệu mở 642
9.3. BA MẶT CỦA CHỦ QUYỀN 644
9.3.1. Chủ quyền biểu diễn 645
9.3.2. Chủ quyền khả chuyển 646
9.3.3. Chủ quyền liên thông 647
9.3.4. Thiếu một mặt, hai mặt còn lại suy yếu 648
9.3.5. Chủ quyền là khả năng tồn tại nhất quán ở mọi hệ thống 650
9.4. QUYỀN MỜ VÀ DỮ LIỆU NHƯ HẠ TẦNG CÔNG 652
9.4.1. Không phải mọi thứ đều cần khả đọc 654
9.4.2. Quyền được giữ mờ 655
9.4.3. Dữ liệu công khác dữ liệu thương mại 657

9.4.4. Khi mở dữ liệu nhưng không mất bản sắc 658
9.4.5. Hạ tầng dữ liệu công giống đường và điện 659
***Phần B: Trách nhiệm*** ***662***
9.5. TÁC NHÂN ĐẶT SAI THÌ AI CHỊU 662
9.5.1. Sai vì dữ liệu 664
9.5.2. Sai vì thuật toán 666
9.5.3. Sai vì nền tảng 667
9.5.4. Sai vì người vận hành 669
9.5.5. Chuỗi trách nhiệm mới 671
9.6. MÔ TẢ SAI VÀ THIỆT HẠI THƯƠNG HIỆU 675
9.6.1. Khi máy tạo ra một sự thật không tồn tại 677
9.6.2. Sai nhưng được lặp lại thành đúng 678
9.6.3. Danh tiếng bị hao mòn âm thầm 679
9.6.4. Chi phí sửa sai tăng theo thời gian 680
9.6.5. Quản trị thương hiệu trở thành quản trị tri thức 681
9.7. HẠ TẦNG KHIẾU NẠI ĐỐI VỚI HỆ SINH THÁI AI 684
9.7.1. Khiếu nại trang web không đủ 686
9.7.2. Khiếu nại AI khác khiếu nại báo chí 687
9.7.3. Thiếu quy trình xác minh 688
9.7.4. Thiếu chuẩn phản hồi 689
9.7.5. Đầu mối bảo vệ quyền biểu diễn số 690
9.8. NĂNG LỰC QUẢN TRỊ NỘI BỘ CỦA ĐỊA PHƯƠNG 693
9.8.1. Giao trách nhiệm cho một đầu mối 694
9.8.2. Kiểm toán tri thức theo chu kỳ 695
9.8.3. Xác lập nguồn chuẩn 696
9.8.4. Sửa sai theo thời hạn 697
9.8.5. Đo chất lượng hiện diện số thay vì đếm lượt khách 698
9.9. KẾT LUẬN CHƯƠNG 703
*Phụ lục 9a. Bộ câu hỏi kiểm toán tri thức* *705*

***Danh mục thuật ngữ*** ***706***
***Danh mục tài liệu tham khảo*** ***714***

# GIỚI THIỆU NỘI DUNG CÁC CHƯƠNG

**CHƯƠNG 1: TỪ TRẢI NGHIỆM ĐẾN ỦY NHIỆM** 31

*Điều gì đã thực sự thay đổi khi con người ngày càng ít tự tìm kiếm. Chương dựng lại bước chuyển từ kinh tế chú ý sang kinh tế ủy nhiệm, phân biệt tác nhân với trợ lý hội thoại và đặt nghịch lý trung tâm của cuốn sách: trải nghiệm càng sâu thì càng khó khả đọc.*

**CHƯƠNG 2: TÍNH KHẢ ĐỌC THUẬT TOÁN** 71

*Phả hệ chính trị của khái niệm khả đọc, từ Scott, Bowker và Star tới thuật toán. Chương đề xuất mô hình bốn bộ lọc, một vòng và một van, mô tả cơ chế tái cấu trúc và đối chất mười một phản bác.*

**CHƯƠNG 3: MÁY ĐỌC ĐỊA PHƯƠNG BẰNG KIẾN TRÚC GÌ?** 143

*Bốn ràng buộc của mọi hệ thống suy luận hữu hạn và bốn thao tác bất biến sinh ra từ chúng. Chương theo một minh hoạ duy nhất, qua bốn khâu để kiểm tra điều kiện khả đọc thuật toán.*

**CHƯƠNG 4: NIỀM TIN VÀ KÝ ỨC CỦA MÁY** 191

*Hai bộ lọc đầu và cái vòng. Máy phân giải thực thể, đối chiếu xuất xứ, tự viết câu mô tả, rồi ghi nhớ. Chương giải thích vì sao sửa sai muộn tốn kém phi tuyến và vì sao quản trị thương hiệu trở thành quản trị ký ức.*

**CHƯƠNG 5: CUỘC CẠNH TRANH GIÀNH QUYỀN ĐƯỢC ĐỀ XUẤT 275**

*Bộ lọc xếp hạng. Kinh tế học của sự đề xuất, tập ứng viên, hiệu ứng Matthew của thuật toán và nghịch lý máy càng thông minh thì đám đông càng dễ tập trung, trong điều kiện không có cơ chế đa dạng hóa hoặc điều tiết. Xem xét ba tầng can thiệp vào chuỗi quyết định.*

**CHƯƠNG 6: AI GIỮ LẠI GIÁ TRỊ 361**

*Bốn bộ lọc đọc như một biểu giá. Chuỗi giá trị bị cắt ở quyền lựa chọn và giao diện giao dịch, tô thuật toán, ba lớp chi phí tham gia và nguy cơ địa phương bị OEM hóa.*

**CHƯƠNG 7: THIẾT KẾ ĐỊA PHƯƠNG CHO MÁY 441**

*Từ khả đọc sang khả hành động: địa phương như một đồ thị, tồn kho giao dịch được, bốn điều kiện để tác nhân dám giao dịch và bảng điều khiển khả hành động thay cho bảng điều khiển kết quả.*

**CHƯƠNG 8: VIỆT NAM: DẤU THANH, ĐỊA DANH VÀ NĂNG LỰC CẤP XÃ 539**

*Bộ lọc thứ nhất trong điều kiện Việt Nam. Phân mảnh thực thể do dấu thanh và viết tắt, địa danh sau sáp nhập, khoảng cách giữa dữ liệu hành chính và dữ liệu quyết định và năng lực số ở tuyến đầu là cấp xã.*

**CHƯƠNG 9: CHỦ QUYỀN VÀ TRÁCH NHIỆM 619**

*Ai sở hữu hình ảnh của một địa phương, chủ quyền dữ liệu khác chủ quyền thuật toán ở đâu, quyền mờ như một lựa chọn quản trị và chuỗi trách nhiệm khi một tác nhân đặt sai.*

## DANH MỤC TỪ VIẾT TẮT

| Từ viết tắt | Từ viết đầy đủ | Nghĩa dùng trong sách |
|---|---|---|
| **AI** | Artificial Intelligence | Trí tuệ nhân tạo |
| **ACP** | Agentic Commerce Protocol | Giao thức thương mại tác nhân do OpenAI và Stripe phát triển. Đặc tả doanh nghiệp, chưa phải tiêu chuẩn ngành |
| **AP2** | Agent Payments Protocol | Giao thức thanh toán cho tác nhân do Google công bố, sau đó chuyển cho FIDO Alliance. Đặc tả mở do doanh nghiệp khởi xướng, chưa phải tiêu chuẩn chính thức |
| **API** | Application Programming Interface | Giao diện lập trình ứng dụng |
| **CARE** | Collective benefit, Authority to control, Responsibility, Ethics | Bộ nguyên tắc quản trị dữ liệu bản địa. Sáng kiến của cộng đồng nghiên cứu, không có hiệu lực pháp lý |
| **CMA** | Competition and Markets Authority | Cơ quan Cạnh tranh và Thị trường Vương quốc Anh |
| **DMA** | Digital Markets Act | Đạo luật Thị trường Số của Liên minh châu Âu |
| **DSA** | Digital Services Act | Đạo luật Dịch vụ Số của Liên minh châu Âu |
| **DOI** | Digital Object Identifier | Mã định danh tài liệu số |
| **EBITDA** | Earnings Before Interest, Taxes, Depreciation and Amortization | Lợi nhuận trước lãi vay, thuế và khấu hao |
| **FAIR** | Findable, Accessible, Interoperable, Reusable | Bộ nguyên tắc quản trị dữ liệu khoa học. Sáng kiến của cộng đồng nghiên cứu, không có hiệu lực pháp lý |
| **FDI** | Foreign Direct Investment | Đầu tư trực tiếp nước ngoài |
| **GDPR** | General Data Protection Regulation | Quy định chung về bảo vệ dữ liệu của Liên minh châu Âu |
| **GDS** | Global Distribution System | Hệ thống phân phối toàn cầu trong ngành lữ hành |

| Từ viết tắt | Từ viết đầy đủ | Nghĩa dùng trong sách |
|---|---|---|
| **LLM** | Large Language Model | Mô hình ngôn ngữ lớn |
| **MCP** | Model Context Protocol | Giao thức kết nối mô hình với công cụ và dữ liệu ngoài. Đặc tả mở do doanh nghiệp công bố, được nhiều bên áp dụng, chưa qua một tổ chức tiêu chuẩn hóa |
| **OCOP** | One Commune One Product | Chương trình mỗi xã một sản phẩm |
| **OEM** | Original Equipment Manufacturer | Nhà sản xuất theo đặt hàng, không sở hữu thương hiệu cuối |
| **OSM** | OpenStreetMap | Cơ sở dữ liệu bản đồ mở do cộng đồng đóng góp |
| **OTA** | Online Travel Agency | Đại lý du lịch trực tuyến |
| **PROV-DM** | Provenance Data Model | Mô hình dữ liệu về xuất xứ. Khuyến nghị chính thức của W3C, ban hành năm 2013 |
| **SEO** | Search Engine Optimization | Tối ưu hóa cho công cụ tìm kiếm |
| **UDRP** | Uniform Domain-Name Dispute-Resolution Policy | Chính sách giải quyết tranh chấp tên miền của ICANN, đang có hiệu lực |
| **UNESCO** | United Nations Educational, Scientific and Cultural Organization | Tổ chức Giáo dục, Khoa học và Văn hóa Liên hợp quốc |
| **W3C** | World Wide Web Consortium | Tổ chức chuẩn hóa công nghệ web |

# DANH MỤC BẢNG

**Bảng 1.1.** Bốn khái niệm thường bị dùng thay cho nhau 45
**Bảng 2.1.** Cơ chế tái cấu trúc trong năm lĩnh vực không liên quan tới nhau 114
**Bảng 4.1.** Sáu lớp bằng chứng cấu thành một thực thể 202
**Bảng 4.2.** Trọng số nguồn thay đổi theo loại câu hỏi 216
**Bảng 5.1.** Ba tầng can thiệp vào chuỗi quyết định 352
**Bảng 6.1.** Mức thu phổ biến trên mỗi giao dịch tại một số điểm nghẽn 380
**Bảng 6.2.** Cấu phần chi phí tham gia nền kinh tế tác nhân 390
**Bảng 6.3.** Khách sạn M, ba lớp chi phí tính trên doanh thu phòng 393
**Bảng 6.4.** Bộ chỉ số đo năng lực giữ giá trị của một địa phương 436
**Bảng 7.1.** Bốn cấp độ trưởng thành của một điểm đến trước máy 453
**Bảng 7.2.** Từ tài sản sang đơn vị tồn kho giao dịch được 490
**Bảng 7.3.** Từ khái niệm đến công cụ 538
**Bảng 9.1.** Chủ quyền dữ liệu và chủ quyền thuật toán so theo bảy tiêu chí 638
**Bảng 9.2.** Ba mặt của chủ quyền: câu hỏi kiểm tra, chỉ báo suy giảm và biện pháp 649
**Bảng 9.3.** Bộ chỉ báo quản trị hiện diện số 700
**Bảng 9A.1.** Bộ câu hỏi kiểm toán tri thức về địa phương 705

# DANH MỤC HÌNH

**Hình 1.1.** Chuỗi ra quyết định 46
**Hình 2.1.** Bốn bộ lọc, một vòng, một van 99
**Hình 2.2.** Cơ chế tái cấu trúc, bốn bộ lọc của máy 113
**Hình 4.1.** Năm nấc chi phí sửa sai 267
**Hình 6.1.** Hai chuỗi giá trị đặt cạnh nhau 367
**Hình 6.2.** Sự dịch chuyển của biên lợi nhuận 415
**Hình 8.1.** Bốn nguồn phân mảnh giữa một địa phương và thực thể mà máy nhìn thấy 546

## LƯU Ý VỀ PHẠM VI, NGUỒN VÀ CÁCH DÙNG THUẬT NGỮ

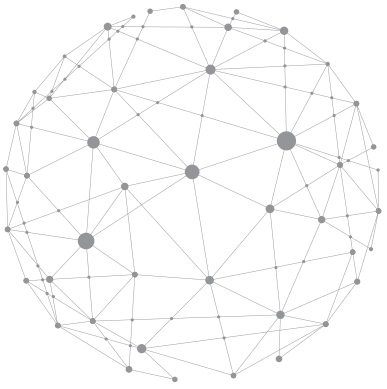
Về chữ "máy". Trong cả cuốn sách, "máy" là cách gọi rút gọn cho một chuỗi hệ thống kỹ thuật và tổ chức nối nhau, gồm công cụ tìm kiếm, mô hình ngôn ngữ, hệ truy hồi, đồ thị tri thức, hệ thống xếp hạng, tác nhân giao dịch, nền tảng đặt chỗ, bộ xử lý thanh toán và các mô hình của bên thứ ba. Cách gọi này không hàm ý có một thuật toán duy nhất hay một chủ thể duy nhất đọc toàn bộ địa phương. Những câu như "máy nhớ", "máy tin" hay "máy dám giao dịch" là phép nhân hóa dùng để rút gọn và ở mỗi chỗ quan trọng, cơ chế kỹ thuật tương ứng đều được nêu kèm. Hệ quả của lưu ý này rất cụ thể ở Chương 9: trách nhiệm phải được phân bổ theo từng mắt xích, không quy cho một "máy" trừu tượng.

Về địa vị của khái niệm trung tâm. "Nền kinh tế tác nhân" được dùng trước hết như một cấu hình kinh tế, tức một cách tổ chức quan hệ giữa cung và cầu trong đó tác nhân máy tham gia dựng tập lựa chọn và thực hiện nhiệm vụ. Đây không phải tuyên bố rằng toàn bộ nền kinh tế đã bước sang một giai đoạn hoàn toàn mới, cũng không phải tên gọi cho một thế hệ công nghệ. Cấu hình ấy hiện diện không đều, đậm ở một số ngành và loại giao dịch, mờ ở nhiều ngành khác. Mọi mệnh đề trong sách cần được đọc trong phạm vi ấy.

Về nguồn và mức tin cậy. Các dẫn chứng trong sách thuộc ba nhóm có giá trị chứng minh khác nhau. Nhóm thứ nhất là nghiên cứu bình duyệt và văn bản pháp luật. Nhóm thứ hai là số liệu thống kê nhà nước và báo cáo có ràng buộc pháp lý như hồ sơ tuân thủ hoặc báo cáo tài chính. Nhóm thứ ba là khảo sát ngành, dữ liệu do doanh nghiệp tự công bố và phát ngôn của bên có lợi ích trực tiếp. Chú thích chân trang ghi rõ nhóm của từng nguồn

cùng giới hạn phương pháp của nó. Các con số thuộc nhóm thứ ba chỉ nên đọc theo bậc độ lớn và theo hướng biến động.

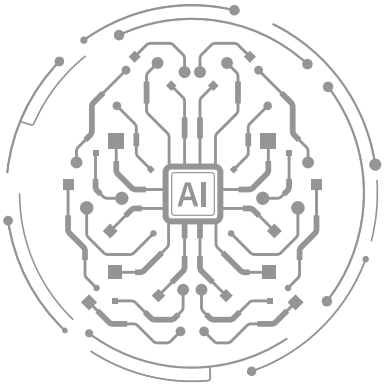
Về thời điểm. Các nền tảng, giao thức, sản phẩm và quy định được trình bày theo trạng thái tại thời điểm bản thảo hoàn thành. Lĩnh vực này thay đổi nhanh, vì vậy độc giả cần kiểm tra phiên bản hiện hành trước khi áp dụng bất kỳ khuyến nghị kỹ thuật hoặc pháp lý nào trong sách.



## CHƯƠNG 1

# TỪ TRẢI NGHIỆM ĐẾN ỦY NHIỆM

*Điều gì đã thực sự thay đổi khi con người ngày càng ít tự tìm kiếm. Chương dựng lại bước chuyển từ kinh tế chú ý sang kinh tế ủy nhiệm, phân biệt tác nhân với trợ lý hội thoại và đặt nghịch lý trung tâm của cuốn sách: trải nghiệm càng sâu thì càng khó khả đọc.*


AI

Trong nền kinh tế trải nghiệm, địa phương cạnh tranh để con người lựa chọn. Trong nền kinh tế tác nhân, địa phương còn phải cạnh tranh để máy lựa chọn thay con người. Sự thay đổi ấy làm xuất hiện một loại tài sản mới là khả đọc.

Một tình huống không khó hình dung. Chủ một homestay ở Hội An dành buổi chiều tháng Tư để tìm hiểu vì sao mùa cao điểm năm nay vắng hơn năm ngoái. Thứ hạng tìm kiếm không đổi. Điểm đánh giá tăng, 4,8 trên 5. Lượng truy cập trang nhà giữ nguyên. Mọi chỉ số ông biết cách đo đều nói rằng năm nay phải đông hơn. Cùng buổi chiều ấy, một du khách Singapore hỏi trợ lý số của mình về chỗ ở tại Hội An và nhận về ba cái tên. Không cái nào là của ông. Cơ sở của ông có mặt trên internet dưới bốn cách viết tên, hai địa chỉ lệch nhau một số nhà, giờ nhận phòng ghi trong bài giới thiệu mâu thuẫn với giờ trên trang đặt phòng và không nơi nào cho phép một phần mềm hỏi xem tối thứ Sáu còn phòng hay không. Trong mô hình mà hệ thống dựng về Hội An, cơ sở lưu trú ấy gần như không tồn tại. Ông thua một cuộc thi mà ông không biết mình có dự. Những cuộc thi như thế đang diễn ra hằng ngày, ở quy mô từ một cơ sở lưu trú tại nhà dân (homestay) đến một tỉnh. Chúng không có thông báo trượt, không có biên bản và không có ai để hỏi.

## 1.1. NƠI CUỐN SÁCH TRƯỚC DỪNG LẠI

Cuốn ***"Địa phương trong nền kinh tế trải nghiệm"*** khép lại ở một thời điểm mà internet đã thay đổi cách con người tiếp cận thông tin, mạng xã hội đã thay đổi cách con người kể lại chuyến đi, các nền tảng đã thay đổi cách doanh nghiệp tìm đến khách hàng. Xuyên suốt cuốn ấy có một giả định không bao giờ bị chất vấn, vì nó chưa từng có lý do để bị chất vấn, người tìm kiếm, người đánh giá và người quyết định cuối cùng đều là con người.

### 1.1.1. Từ phù sa sang bộ lọc

***"Địa phương trong nền kinh tế trải nghiệm"*** mô tả sự phát triển của một địa phương bằng hình ảnh một dòng sông. Tài sản của vùng đất sinh ra trải nghiệm, trải nghiệm lắng lại thành ký ức, ký ức bồi lên thành tình yêu nơi chốn và qua nhiều thế hệ khách, lớp bồi ấy trở thành thương hiệu. Phù sa không bồi theo quý. Đó là toàn bộ luận điểm. Hình ảnh ấy vẫn đúng. Cuốn sách này không thay bằng hình ảnh khác mà đặt thêm một vật thể vào giữa dòng. Giữa thượng nguồn và hạ lưu, một cấu trúc được dựng lên. Nước vẫn chảy, nhưng chảy qua một loạt bộ lọc xếp nối nhau. Mỗi tầng giữ lại một phần và để phần còn lại đi tiếp. Cái không lọt qua không biến mất, nó chỉ nằm lại phía trên. Và ở đâu đó trên cấu trúc ấy có một cái van, tức là một chỗ mà ai đó quyết định mở bao nhiêu và cho cái gì đi qua. Toàn bộ cuốn sách này bàn về bốn bộ lọc và một cái van đó. Chương sau sẽ dựng chúng thành mô hình có định nghĩa và cũng có tiêu chuẩn bác bỏ. Ở đây chỉ cần giữ lấy ba điều mà hình ảnh này hàm ý.

- Bộ lọc không phá dòng sông. Nó quyết định phần nào của dòng sông đến được hạ lưu và hạ lưu là nơi có người ở.

- Cái bị giữ lại vẫn tồn tại. Nó chỉ tồn tại ở nơi không ai đi qua.

- Phù sa chỉ đi qua khi mắt lọc đủ lớn và chỉ lắng lại khi van dòng chảy đổi trạng thái. Chính bộ lọc và van, chứ không phải bản thân hạt phù sa, quyết định số phận của nó.

Hình ảnh này sẽ không được nhắc lại thường xuyên trong các trang tới. Nó cần thiết ở đây vì nó là chỗ nối giữa hai cuốn sách và vì nó cho phép chương sau dựng mô hình mà không phải giải thích lại từ đầu.

### 1.1.2. Khi trải nghiệm trở thành lợi thế cạnh tranh

Luận điểm cốt lõi của tôi trong tác phẩm ***"Địa phương trong nền kinh tế trải nghiệm"*** là một địa phương không phát triển bền vững bằng cách bán tài nguyên hay xây thêm hạ tầng. Bãi biển đẹp có ở khắp nơi. Điều hiếm là năng lực chuyển hóa tài sản sẵn có thành trải nghiệm đủ khác biệt để một con người, sau khi trở về và sống tiếp cuộc đời bận rộn của mình mà vẫn còn nhớ.

Cơ sở tâm lý của lập luận ấy khó bác bỏ. Người ta quên tên khách sạn, quên giá vé, quên lịch trình chi tiết đã mất nhiều tuần để lập. Cái còn lại sau vài năm thường là những mảnh vụn kỳ quặc, như mùi của một khu chợ lúc năm giờ sáng, cách một người bán hàng cười khi ta trả giá sai, ánh sáng hắt qua một khung cửa vào đúng một buổi chiều, người xa quê nhớ mùi rơm nếp, thậm chí cả mùi phân bò vương vãi trên đường quê. Những mảnh vụn ấy không nằm trong tập gấp quảng bá nào và cũng là toàn bộ lý do cho người ta quay lại.

Điều đó đúng nhưng chưa đủ. Lập luận về chiều sâu trải nghiệm giải thích tốt vì sao người ta yêu một nơi. Nó giải thích kém hơn vì sao người ta chọn một nơi. Yêu xảy ra sau khi đã đến. Chọn xảy ra trước, thường trong trạng thái thiếu thông tin và bị giới hạn thời gian, đặc biệt ở trạng thái ấy con người dựa vào những tín hiệu thô sơ đến mức đáng thất vọng. Cuốn trước dành phần lớn dung lượng cho giai đoạn sau. Chỗ trống ấy là nơi cuốn này bước vào.

### 1.1.3. Điều chưa từng được đặt ra

Trí tuệ nhân tạo không bác bỏ giá trị của trải nghiệm. Không hệ thống nào ngửi được mùi bùn sau cơn mưa ở đồng bằng sông Cửu Long hay hiểu vì sao một bát bún ăn đứng ở vỉa hè lại ngon hơn cùng bát bún ấy dọn trong nhà hàng. Trải nghiệm theo nghĩa chủ quan, tức là sống trong một ngữ cảnh và mang ý nghĩa đối với một chủ thể, vẫn là lãnh địa của con người. Nhưng giữa việc trải nghiệm thuộc về con người và việc quyết định thuộc về con người có một khoảng cách và khoảng cách ấy đang mở ra.

Câu hỏi mà cuốn trước không đặt ra không phải liệu máy có thông minh hơn người hay không. Đó là câu hỏi của giới công nghệ và nó ít liên quan tới quản trị địa phương hơn người ta tưởng. Câu hỏi thật sự nhỏ hơn nhiều, điều gì xảy ra khi con người không còn muốn tự mình làm công việc sàng lọc. Không ai bị tước quyền lựa chọn. Người ta tự nguyện giao nó đi, vì lý do giản dị là việc đọc bốn mươi bài đánh giá khách sạn rất mất thời gian. Quyết định cuối cùng vẫn thuộc về con người. Quyết định đầu tiên, tức quyết định về việc những phương án nào được đưa lên bàn cân, đã bắt đầu chuyển chủ. Phù sa không biến mất. Nó chỉ không lọt qua bộ lọc.

## 1.2. KHI TÌM KIẾM BIẾN THÀNH ỦY NHIỆM

### 1.2.1. Bốn phút, ba phương án

Tình huống dưới đây là một kịch bản giả định nhằm minh họa cơ chế, các con số không phải dữ liệu quan sát. Một phụ nữ ngoài ba mươi ngồi ở sân bay Changi chờ chuyến bay nối. Cô có năm ngày nghỉ vào tháng sau và chưa biết đi đâu. Năm năm trước cô sẽ mở mười ba thẻ trình duyệt. Bây giờ cô nói với trợ lý số trên điện thoại với các thông tin: năm ngày, ngân sách vừa, thích ăn uống và văn hóa bản địa, ngại chỗ đông khách phương Tây, bay từ Singapore. Rồi cô đi mua cà phê.

Bốn phút sau, ba phương án. Mỗi phương án gồm chuyến bay có giờ và giá, khách sạn đã kiểm tra còn phòng, ba nhà hàng mỗi ngày kèm lý do chọn, tổng chi phí ước tính. Cô đọc trong lúc uống cà phê, hỏi lại hai câu, đổi một khách sạn, bấm xác nhận. Trong bán kính bay bốn giờ từ Singapore có hàng trăm điểm đến khớp với mô tả của cô. Ba trong số đó xuất hiện. Phần còn lại không xuất hiện trước người dùng và cô không có cách nào biết chúng đã bị xét rồi hạ điểm, không được truy hồi, hay chưa từng đi vào tập ứng viên.

Cảnh ở Changi không phải trải nghiệm phổ quát và sẽ là cẩu thả nếu coi nó như vậy. Phần lớn khách du lịch trên thế giới hiện vẫn đặt phòng theo cách cũ, còn một tỷ trọng lớn của thị trường vẫn sử dụng tour trọn gói do các đại lý lữ hành phân phối. Tỷ lệ hiện tại chưa nói lên nhiều. Đáng chú ý hơn là đặc điểm của nhóm đã chuyển, trẻ hơn, chi tiêu cao hơn, đi xa hơn. Các cuộc chuyển kênh trong lịch sử thương mại hiếm khi bắt đầu từ số đông.

### 1.2.2. Hai hợp đồng khác nhau

*"Tìm kiếm"* là một hợp đồng ba bên. Người dùng nêu câu hỏi, cỗ máy trả về danh sách đã xếp hạng, quyền quyết định nằm ở người dùng. Máy quyết định thứ tự chứ không quyết định lựa chọn. Vị trí thứ bảy vẫn còn cơ may, vì mắt người vẫn lướt qua. Tìm kiếm chừa cho sự tình cờ một khe hở và ba mươi năm marketing số trong giai đoạn vừa qua là nghệ thuật sống trong khe hở ấy.

*"Ủy nhiệm"* là hợp đồng khác. Người dùng nêu mục tiêu kèm ràng buộc, máy trả về phương án, đôi khi trả về kết quả đã làm xong. Con người từ vị trí kẻ lựa chọn lùi về vị trí kẻ phê duyệt. Kẻ lựa chọn đọc mười phương án để chọn một. Kẻ phê duyệt đọc một phương án để gật hoặc lắc.

Trong chế độ tìm kiếm, phễu quyết định rộng và rò rỉ. Trong chế độ ủy nhiệm, máy cân nhắc hàng nghìn ứng viên ở những tầng không ai nhìn thấy rồi để lọt ra ba. Kẻ không nằm trong tập được cân nhắc không thua ở vòng chung kết. Họ thua ở một vòng loại không có khán giả.

Bước lùi ấy đã để lại dấu vết trong dữ liệu, dù dấu vết ấy chưa mang tên ủy nhiệm. Trong bốn tháng đầu năm 2026, 68% lượt tìm kiếm Google tại Mỹ kết thúc mà người dùng không nhấp vào liên kết nào, tăng từ 60,45% của năm 2024.[1] Cả ngành marketing số được xây trên giả định rằng sự quan tâm để lại dấu vết dưới dạng một cú nhấp. Khi hai phần ba số lượt quan tâm

1. Rand Fishkin, "In 2026, Less than One Third of Google Searches Still Send a Click", SparkToro, ngày 8 tháng 6 năm 2026, https://sparktoro.com/blog/in-2026-less-than-one-third-of-google-searches-still-send-a-click/. Chính xác là 68,01%, trên dữ liệu dòng nhấp chuột Similarweb tại Mỹ, tháng 1 tới tháng 4 năm 2026. Mức nền 60,45% năm 2024 lấy từ bộ dữ liệu Datos. Các mốc 2016, 2019, 2024 và 2026 lấy từ ba bảng đo khác nhau là Jumpshot, Datos và Similarweb, nên chuỗi số chỉ để minh hoạ về hướng.

không còn để lại dấu vết ấy, các bảng điều khiển vẫn chạy, vẫn xanh và vẫn đo một thế giới không còn tồn tại.

### 1.2.3. Không chỉ là chuyện du lịch

Cấu trúc vừa mô tả không thuộc riêng ngành du lịch. Nó xuất hiện ở bất cứ đâu có một tập ứng viên lớn và một người ra quyết định thiếu thời gian.

Trong tuyển dụng, phần mềm sàng lọc hồ sơ đã đứng giữa người ứng tuyển và người tuyển dụng từ hơn một thập niên. Một ứng viên bị loại ở tầng ấy không nhận được thư từ chối có nội dung và trong nhiều trường hợp không nhận được gì. Ứng viên đó không biết mình bị loại vì thiếu năng lực hay vì để trống một trường dữ liệu. Ba tình huống khác nhau về bản chất và thường bị gộp làm một. Bị đánh giá thấp nghĩa là hồ sơ đã được đọc và bị chấm điểm kém, đây là thất bại có thể học hỏi từ nó. Không vượt qua tầng phân giải nghĩa là hệ thống không xác định được ứng viên này là ai, chẳng hạn vì tên trường viết theo ba cách khác nhau trong ba phần của cùng một hồ sơ. Còn bị loại do trường dữ liệu trống thì thậm chí không phải một phán quyết về con người, nó là một phép lọc trên một ô rỗng. Chỉ tình huống thứ nhất là cạnh tranh theo nghĩa thông thường. Hai tình huống sau xảy ra trước khi có cạnh tranh và người bị loại không nhận được tín hiệu nào cho phép phân biệt mình rơi vào loại nào. Đây chính là cấu trúc mà một địa phương gặp phải, chỉ khác ở chỗ một ứng viên có thể sửa hồ sơ trong một buổi tối còn một vùng đất thì không.

Trong mua sắm doanh nghiệp, một nhà cung ứng không có hồ sơ năng lực ở dạng máy đọc được sẽ không xuất hiện trong danh sách so sánh mà phần mềm mua hàng tạo ra. Chất lượng

sản phẩm không tham gia vào quyết định đó, vì quyết định đó xảy ra trước khi có ai xem sản phẩm.

Trong bảo hiểm và tín dụng, một người không có lịch sử giao dịch số bị xếp vào nhóm không đủ dữ liệu để đánh giá. Đây là nhóm khó nhất vì nó không phải nhóm bị đánh giá xấu. Nó là nhóm không có điểm. Ở mọi lĩnh vực, hình phạt dành cho việc không khả đọc đều giống nhau. Không bị chấm điểm thấp. Chỉ đơn giản là không có tên trong danh sách được chấm. Ngành du lịch chỉ là nơi hiện tượng này dễ quan sát nhất, vì đối tượng bị đọc là một vùng đất và một vùng đất thì không thể sửa hồ sơ của mình trong một buổi chiều.

### 1.2.4. Tác nhân không trả lời, tác nhân hành động

Ranh giới giữa một trợ lý hội thoại và một tác nhân nằm ở chỗ ai gánh hậu quả. Một trợ lý trả lời sai thì người dùng đọc xong, thấy vô lý, bỏ qua. Một tác nhân làm sai thì đã có một khoản tiền rời khỏi tài khoản. Vì hậu quả là thật, tác nhân đòi hỏi thứ mà trợ lý không cần, một thế giới truy vấn được ở thời gian thực. Muốn giữ chỗ, phải biết chắc còn chỗ. Muốn thanh toán, phải biết chắc giá. Muốn đổi lịch khi chuyến bay trễ, phải có quyền đổi. Ngành khách sạn đã bắt đầu đo khoảng cách này. Một khảo sát công bố tháng Sáu năm 2026 ghi nhận 11% tổ chức khách sạn được nghiên cứu có khả năng để một tác nhân hoàn tất việc đặt phòng và định giá tồn kho theo thời gian thực.[1] Nghiên cứu này do một

1. Aven Hospitality và h2c, sách trắng "The New Search Economy: Rewriting the Rules of Direct Booking with AI", ngày 22 tháng 6 năm 2026, https://skift.com/2026/06/22/ai-search-economy-hospitality-white-paper/. Bài trên Skift được đánh dấu rõ là nội dung tài trợ, không phải bài báo độc lập. Aven là nhà cung cấp giải pháp phân phối khách sạn, có lợi ích trực tiếp, nên con số chỉ nên đọc theo bậc độ lớn.

nhà cung cấp công nghệ phân phối thực hiện cùng một hãng tư vấn và được công bố dưới dạng nội dung có tài trợ. Một con số thấp phục vụ lợi ích của bên bán giải pháp, đúng như một con số cao sẽ phục vụ lợi ích của bên bán sự yên tâm. Con số nên được dùng để chỉ hướng chứ không để chỉ mức.

## 1.3. NỀN KINH TẾ TÁC NHÂN LÀ GÌ

### 1.3.1. Từ internet đến nền kinh tế tác nhân

Lịch sử các nền tảng số thường được kể như lịch sử của công nghệ. Nó được kể chính xác hơn như lịch sử của việc cắt giảm chi phí. Google không phát minh ra internet, họ hạ chi phí tìm thông tin. Amazon không phát minh ra mua bán, họ hạ chi phí tìm một món hàng giữa hàng triệu món. Airbnb không phát minh ra chỗ trọ, họ hạ chi phí xây niềm tin giữa hai người xa lạ. Mỗi lần một loại chi phí vô hình bị hạ xuống, một ngành bị vẽ lại. Sau ba mươi năm, internet vẫn để lại nguyên một loại chi phí. Chi phí ra quyết định.

Một người chuẩn bị đi Nha Trang vẫn phải đọc hàng trăm bài đánh giá. Một gia đình cho con du học vẫn phải mở hàng chục trang so sánh. Một doanh nghiệp chọn nhà cung cấp vẫn phải đọc vô số hồ sơ. Internet làm thông tin dồi dào lên và chính sự dồi dào ấy sinh ra nghịch lý, khi mọi thứ đều tìm được, việc khó không còn là tìm mà là chọn. Herbert Simon đã mô tả cơ chế này từ năm 1971, rằng một thế giới giàu thông tin tất yếu tạo ra sự nghèo đi của thứ mà thông tin tiêu thụ, tức là sự chú ý của người nhận.[1]

Các tác nhân xuất hiện đúng vào chỗ trống ấy. Nếu Google giúp con người tìm nhanh hơn thì tác nhân giúp con người không phải tự tìm nữa. Đó không phải một cải tiến về tốc độ. Đó là sự thay đổi của chủ thể thực hiện việc lựa chọn và trong kinh tế học, thay đổi chủ thể lựa chọn là loại thay đổi làm dịch chuyển toàn bộ đường cầu.

1. Herbert A. Simon, "Designing Organizations for an Information-Rich World", trong Martin Greenberger (chủ biên), Computers, Communication, and the Public Interest, Johns Hopkins Press, Baltimore, 1971, trang 37 tới 72.

### 1.3.2. Tác nhân khác trợ lý hội thoại ở đâu

Nền kinh tế tác nhân là cấu hình kinh tế trong đó các hệ thống máy có khả năng suy luận trở thành chủ thể tìm kiếm, so sánh, thương lượng và giao dịch thay mặt con người hoặc tổ chức, khiến một phần đáng kể của cầu thị trường được biểu đạt và được đáp ứng thông qua tương tác giữa máy với máy. Định nghĩa gọn này là phần cần nhớ. Ba điều kiện dưới đây phân biệt nó với những thứ dễ bị nhầm.

Cần một thang trước khi đi tiếp, vì chữ "ủy nhiệm" dễ bị đọc ở mức mạnh nhất của nó. Trên thực tế có năm mức và chỉ mức cuối mới đúng nghĩa máy chọn thay người. Mức một, hỗ trợ lựa chọn, máy tổng hợp thông tin, con người tự làm mọi việc còn lại. Mức hai, thu hẹp lựa chọn, máy dựng danh sách ngắn, con người chọn trong danh sách ấy. Mức ba, đề xuất lựa chọn, máy đưa ra một phương án, con người phê duyệt. Mức bốn, thực hiện lựa chọn, máy giao dịch trong phạm vi thẩm quyền được giao. Mức năm, tự thích nghi sau lựa chọn, máy điều chỉnh phương án khi điều kiện đổi, không hỏi lại.

Cuốn sách dùng chữ "ủy nhiệm" theo nghĩa rộng, bao trùm cả năm mức và điều đáng chú ý là hệ quả với địa phương xuất hiện từ mức hai chứ không phải mức bốn. Ngay khi máy được giao việc dựng danh sách ngắn, việc có tên trong danh sách ấy đã trở thành một cuộc cạnh tranh riêng, tách khỏi cuộc cạnh tranh giành sự yêu thích của con người. Ủy nhiệm việc dựng tập lựa chọn quan trọng hơn ủy nhiệm cú bấm cuối cùng.

Tác nhân suy luận trong ngữ cảnh mở chứ không chạy kịch bản viết sẵn. Một chương trình đặt vé theo luật cứng không làm nên nền kinh tế tác nhân, cũng như một máy bán nước tự động

không làm nên thương mại điện tử. Tác nhân hành động thay mặt, mang theo mục tiêu, ràng buộc và một mức thẩm quyền có giới hạn của một chủ thể cụ thể. Một phần cầu được biểu đạt qua tương tác máy với máy, nghĩa là ở phía cung, kẻ tiếp đón khách hàng đầu tiên ngày càng không phải nhân viên lễ tân hay trang chủ mà là giao diện dữ liệu mà tác nhân của khách truy vấn.

Cần vạch sẵn ranh giới của khái niệm để khỏi phải tranh luận với những phản bác dễ dãi. Nền kinh tế tác nhân không có nghĩa là mọi giao dịch đều do máy thực hiện, cũng như nền kinh tế số chưa bao giờ có nghĩa là mọi giao dịch đều trực tuyến. Nó có nghĩa là phần giao dịch có máy tham gia đủ lớn để định hình lại cách phần còn lại vận hành.

Cần chốt địa vị của khái niệm ngay từ đây để tránh một hiểu lầm sẽ đi suốt phần còn lại. Trong cuốn sách này, nền kinh tế tác nhân được dùng như một cấu hình kinh tế, tức một cách tổ chức quan hệ giữa cung và cầu, chứ không phải một tuyên bố rằng toàn bộ nền kinh tế đã chuyển sang một giai đoạn hoàn toàn mới. Nó cũng không phải tên gọi cho một thế hệ công nghệ, không phải một xu hướng truyền thông và không thay thế các khái niệm đã có. Bảng 1.1 đặt nó cạnh ba khái niệm mà nó thường bị dùng lẫn.

Trợ lý hội thoại và tác nhân AI khác nhau không chỉ ở mức độ thông minh mà ở bản chất của nhiệm vụ. Trợ lý hội thoại chủ yếu tiếp nhận câu hỏi và cung cấp câu trả lời, tác nhân AI tiếp nhận một mục tiêu kèm theo các ràng buộc rồi tự xây dựng phương án và thực hiện hành động để đạt mục tiêu đó. Trong mô hình này, vai trò của con người cũng thay đổi, thay vì trực tiếp thao tác từng bước, con người chủ yếu phê duyệt hoặc bác bỏ những phương án mà tác nhân đề xuất. Quan hệ với thế giới thực cũng khác biệt. Trợ lý hội thoại chỉ đọc và tổng hợp thông tin,

**Bảng 1.1.** Bốn khái niệm thường bị dùng thay cho nhau

| Khái niệm | Chức năng chính | AI dựng tập lựa chọn |
|---|---|---|
| **Kinh tế số** | Giao dịch dựa trên hạ tầng số | Con người |
| **Kinh tế nền tảng** | Nền tảng điều phối các phía của thị trường | Con người, trong khuôn khổ do nền tảng đặt ra |
| **Kinh tế chú ý** | Cạnh tranh giành thời gian và sự chú ý của con người | Con người |
| **Kinh tế tác nhân** | Tác nhân máy tham gia dựng tập lựa chọn và thực hiện nhiệm vụ | Máy dựng tập lựa chọn, con người phê duyệt trong phần lớn trường hợp |

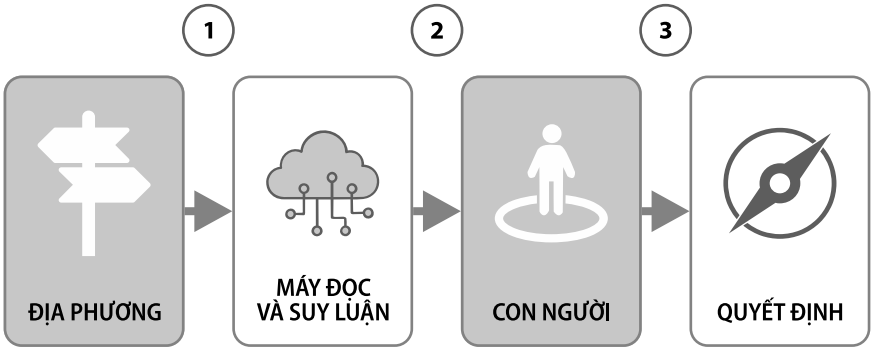
*Nguồn: Tác giả xây dựng*

còn tác nhân AI vừa đọc vừa có thể ghi dữ liệu, cập nhật trạng thái và thực hiện giao dịch trên các hệ thống số. Vì vậy, hậu quả của sai sót cũng lớn hơn nhiều. Một câu trả lời sai của trợ lý hội thoại thường chỉ khiến người dùng bỏ qua hoặc hỏi lại, nhưng một quyết định sai của tác nhân AI có thể dẫn đến tiền đã được chuyển, chỗ đã được đặt hoặc giao dịch đã hoàn tất. Chính vì vậy, điều kiện để chuyển từ mô hình trợ lý hội thoại sang mô hình tác nhân không nằm ở khả năng tạo nội dung hấp dẫn hơn, mà ở việc xây dựng được dữ liệu có cấu trúc, có thể truy vấn, có thể xác minh và có quyền thực hiện giao dịch. Đây lại chính là những năng lực mà phần lớn cơ quan quản lý điểm đến trước nay chưa từng coi là một phần trong nhiệm vụ của mình.

### 1.3.3. Chuỗi ra quyết định mới

Trong nền kinh tế truyền thống, quyết định đi theo một chuỗi quen thuộc, đó là tìm, đọc, so sánh, cân nhắc, chọn. Chuỗi ấy dài vài phút khi mua một cuốn sách, vài ngày khi đặt một chuyến đi,

vài tháng khi chọn một trường đại học và nhiều năm khi một tập đoàn quyết định đặt nhà máy ở đâu. Điều tác nhân làm không phải rút ngắn từng bước. Ở những mức ủy nhiệm cao, chúng bắt đầu thực hiện nhiều bước liền nhau.




**Hình 1.1.** Chuỗi ra quyết định

Hai khâu phía sau không biến mất. Con người vẫn cân nhắc, vẫn đổi ý vào phút chót vì một lý do không giải thích được. Cái đổi là thứ tự, giữa địa phương và con người nay có thêm một khâu và khâu ấy vận hành theo những quy tắc mà địa phương không tham gia soạn thảo, không được thông báo khi thay đổi và không có quyền khiếu nại khi bị áp dụng sai. Mũi tên đầu tiên là mũi tên mới mà phần lớn địa phương chưa tổ chức được năng lực để tác động trực tiếp. Địa phương vẫn tác động trực tiếp vào trải nghiệm thật, giá, hạ tầng, dịch vụ và chính sách. Điều còn thiếu là năng lực tác động vào cách mình được máy đọc. Trong nhiều năm, ngành quản trị điểm đến đổ gần như toàn bộ nguồn lực vào mũi tên thứ ba, tức vào việc thuyết phục con người. Mũi tên thứ nhất chưa từng có ngân sách, chưa từng có phòng ban và ở hầu hết các địa phương, chưa từng có người chịu trách nhiệm.

### 1.3.4. Từ kinh tế chú ý sang kinh tế ủy nhiệm

Nền kinh tế chú ý vận hành trên một tiền đề đơn giản, nguồn lực khan hiếm là mắt người và ai mua được nhiều phút nhìn hơn thì thắng. Toàn bộ ngành quảng cáo, toàn bộ hệ đo lường theo lượt tiếp cận và toàn bộ nghề tối ưu thứ hạng đều là hệ quả của tiền đề ấy.

Nền kinh tế ủy nhiệm không xóa tiền đề của kinh tế chú ý mà đặt thêm một tầng lên trên nó, cạnh tranh giành niềm tin được giao. Nguồn lực khan hiếm không còn là sự chú ý mà là niềm tin được giao và tác nhân chỉ giữ trong tập cân nhắc những thực thể mà nó phân giải được, xác minh được và giao dịch được.

Câu chuyện về việc con người vui vẻ giao quyền cho máy có một hạn chế mà những người kể nó thường bỏ qua. Niềm tin vào tác nhân tự chủ hoàn toàn không tăng đều. Một số khảo sát trong hai năm gần đây ghi nhận tỷ lệ người sẵn sàng để máy hoàn tất giao dịch mà không cần bước xác nhận cuối đã giảm chứ không tăng, khi người dùng va phải những lần máy đặt nhầm và không có ai để gọi. Đường đi của ủy nhiệm nhiều khả năng không phải một đường dốc lên trơn tru mà là một chuỗi tiến hai bước lùi một bước, trong đó mỗi vụ tai tiếng lớn sẽ đẩy thị trường về phía các thiết kế có con người ở giữa. Điều này làm chậm tiến trình. Nó không đảo ngược tiến trình, vì cả những thiết kế có người ở giữa cũng cần máy dựng sẵn tập lựa chọn.

Trong nền kinh tế chú ý, thất bại có nghĩa là bị nhìn thấy ít. Trong nền kinh tế ủy nhiệm, trong phạm vi tập lựa chọn do tác nhân tự dựng, thất bại có thể mang hình thức gần giống sự không tồn tại. Tiền mua được hiển thị chứ không mua được sự hiểu đúng.

## 1.4. KHI MÁY TRỞ THÀNH NGƯỜI ĐỌC ĐẦU TIÊN

### 1.4.1. Ba lần thay đổi của chủ thể đọc

Lịch sử thương hiệu địa phương thường được kể qua các trường phái lý thuyết. Cách kể ấy đúng nhưng che mất một mạch giản dị hơn. Mỗi thời đại của thương hiệu tương ứng với một chủ thể đọc và khi chủ thể đọc đổi, luật chơi đổi theo.

Thời quảng cáo, thông điệp đi thẳng vào mắt và tai con người. New York mua được nửa thế kỷ nhận biết bằng ba ký tự và một trái tim và điều đáng nói là nó rẻ (I ♥ NY). Một ý tưởng sáng đôi khi tạo ra hiệu quả vượt xa quy mô ngân sách ban đầu, dù nó không bao giờ đứng một mình, chiến dịch ấy còn dựa vào chính sách, bối cảnh khủng hoảng và một hệ thống phân phối sẵn có.

Thời nền tảng chèn một tầng trung gian vào giữa. Chủ thể đọc tách làm hai, thuật toán đọc trước và xếp hạng, con người đọc sau và lựa chọn trong phạm vi đã xếp. Địa phương học cách sống hai mặt, một mặt cho người với ảnh đẹp và chuyện hay, một mặt cho máy với từ khóa và siêu dữ liệu. Quyền quyết định cuối vẫn thuộc con người.

Bước chuyển thứ ba hợp nhất hai vai ấy. Hệ thống suy luận không xếp hạng nguồn cho người tự đọc, nó đọc hộ, tổng hợp hộ, so sánh hộ, đề xuất hộ. Một địa phương có thể đứng đầu trang kết quả suốt một thập niên mà vẫn bị mô tả sai trong câu trả lời sinh ra. Bước này đáng lo nhất, vì địa phương không còn biết mình đang được đọc theo tiêu chí nào.

Ba thời đại này chồng lên nhau chứ không thay thế nhau. Quảng cáo vẫn tăng, tìm kiếm vẫn tồn tại và cả hai sẽ còn tồn tại

lâu. Luận điểm ở đây hẹp hơn, chủ thể đọc mới chiếm vị trí quyết định ở đầu phễu, nơi tập lựa chọn được dựng. Mỗi lần chủ thể đọc đổi, tài sản của thời trước không mất giá trị nhưng mất vị trí trung tâm và các địa phương đi sau thường dồn nguồn lực hoàn thiện kỹ nghệ của thời vừa kết thúc đúng lúc trọng tâm cuộc chơi đã dời đi.

### 1.4.2. Máy đọc thế giới như thế nào

Khi ta nói một người đọc một địa phương, ta hình dung họ mở trang thông tin của tỉnh, xem vài bài báo, lướt ảnh, đọc dăm ba đánh giá. Quá trình ấy tuần tự, bị giới hạn bởi thời gian và trí nhớ. Máy không nhất thiết đọc theo cách ấy. Nó không lần lượt lướt từng trang như một người dùng thông thường. Nó dựng một mô hình về thế giới bằng cách liên kết các mảnh thông tin và khi nhận câu hỏi, nó không tìm một câu trả lời có sẵn mà suy ra câu trả lời từ những quan hệ đã có. Máy không chỉ đọc từng tài liệu. Ngày càng nhiều hệ thống còn liên kết các dữ kiện và quan hệ giữa chúng để dựng câu trả lời.

Hãy hình dung một câu hỏi thật: gia đình hai con nhỏ, muốn sang Việt Nam sống ba năm, nên chọn địa phương nào. Để trả lời, máy không tìm nơi nào có cụm từ đáng sống. Nó ráp lại chất lượng không khí, trường quốc tế, bệnh viện, chi phí sinh hoạt, an toàn, cộng đồng người nước ngoài, khoảng cách tới sân bay, cơ hội việc làm cho cha mẹ, thậm chí độ ổn định của chính sách địa phương. Không yếu tố nào riêng lẻ đủ tạo ra khuyến nghị. Cái máy làm là nối chúng lại.

Một mảnh dữ liệu sai không chỉ sai một chỗ. Nó lan theo các mối nối. Một địa chỉ lệch làm hỏng phân giải thực thể, phân giải hỏng làm hỏng đối chiếu nguồn, đối chiếu hỏng làm giảm điểm

tin cậy và điểm tin cậy thấp khiến toàn bộ những gì địa phương ấy nói về mình bị hạ trọng số. Trong một mô hình quan hệ, một lỗi ở tầng định danh có thể lan sang nhiều tầng sau và tạo ra thiệt hại lớn hơn tổng của các lỗi cục bộ.

### 1.4.3. Điều máy đọc được

Máy đọc được tên chuẩn và mã định danh, tọa độ, quan hệ hành chính giữa các cấp, giờ mở cửa, khoảng giá, ngôn ngữ phục vụ, tình trạng còn chỗ, chính sách hủy, thuế phí. Nó đọc được các con số đã được ai đó chịu trách nhiệm công bố. Nó đọc được sự nhất quán giữa nhiều nguồn độc lập và đây là năng lực bị đánh giá thấp nhất, máy không chỉ hỏi thông tin này nói gì, nó hỏi thông tin này có khớp với những chỗ khác không.

Bầu không khí của một khu phố có thể được suy ra từ mật độ quán, giờ hoạt động, phân bố đánh giá và ảnh chụp. Không chính xác, nhưng đủ để xếp hạng.

### 1.4.4. Điều máy không đọc được

Rồi có phần máy không đọc được. Máy không đọc được lý do một bát bún ăn đứng ngon hơn cùng bát bún ấy được phục vụ trong nhà hàng. Nó không đọc được thứ bậc ý nghĩa trong một nghi lễ mà chỉ người lớn lên trong làng mới hiểu. Nó không đọc được sự khác nhau giữa lòng hiếu khách thật và lòng hiếu khách được huấn luyện, dù cả hai đều tạo ra đánh giá năm sao.

Cần phát biểu giới hạn này cho đúng, vì cách nói thông thường rằng máy không đọc được những thứ ấy là quá mạnh và dễ bị bác

bỏ. Máy hoàn toàn có thể mô tả, phân loại và suy đoán về lòng hiếu khách hay về một nghi lễ, qua quá trình học máy với các mẫu ngôn ngữ và hành vi gắn với chúng. Điều nó không làm được là giữ nguyên toàn bộ ý nghĩa của chúng khi tách chúng khỏi ngữ cảnh sống. Mất mát nằm ở phép tách, không ở khả năng nhận biết.

Các hệ thống sinh thường có xu hướng điền vào khoảng trống nếu không được thiết kế hoặc ràng buộc để biểu thị sự bất định. Khi thiếu dữ liệu về một nơi, các hệ thống sinh không được ràng buộc tốt thường không dừng lại ở câu trả lời không biết mà có xu hướng suy ra từ những nơi tương tự. Một xã miền núi chưa từng tự mô tả mình sẽ được mô tả bằng chất liệu vay mượn của xã bên cạnh. Bản dịch ấy không sai hẳn. Nó chỉ không phải là họ.

Một câu tóm tắt do máy sinh ra có độ dài trung vị 67 từ.[1] Sáu mươi bảy từ. Đó là toàn bộ không gian mà một vùng đất có, để tồn tại trước một người chưa từng nghe tên nó. Cùng nghiên cứu ấy ghi nhận rằng khi có tóm tắt do máy sinh ra, chỉ 8% lượt tìm kiếm dẫn tới một cú nhấp vào kết quả thường, so với 15% khi không có. Còn tỷ lệ người nhấp vào chính các nguồn được trích dẫn bên trong bản tóm tắt là 1%.[2] Trong thời nền tảng, một mô tả sai còn sửa được bằng cách kéo người đọc về lãnh thổ của mình. Lãnh thổ ấy giờ hầu như không có ai ghé.

1. Pew Research Center, "Google users are less likely to click on links when an AI summary appears in the results", ngày 22 tháng 7 năm 2025, https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/. Mẫu 900 người trưởng thành tại Mỹ, dữ liệu duyệt web tháng 3 năm 2025, 68.879 lượt tìm kiếm, trong đó 12.593 lượt sinh ra tóm tắt máy.

2. Pew Research Center, cùng nghiên cứu đã dẫn ở chú thích 4. Bên vận hành nền tảng đã công khai phản đối phương pháp luận, xem Liz Reid, "AI in Search is driving more queries and higher quality clicks", blog Google, ngày 6 tháng 8 năm 2025, https://blog.google/products-and-platforms/products/search/ai-search-driving-more-queries-higher-quality-clicks/.

### 1.4.5. Nghịch lý Wikipedia

Tháng Mười năm 2025, Wikimedia Foundation công bố một con số mà tổ chức này rõ ràng không muốn công bố. Sau khi cập nhật hệ thống phân biệt người thật với máy và phân loại lại toàn bộ lưu lượng từ tháng Ba tới tháng Tám, họ phát hiện lượt xem trang do con người thực hiện đã giảm khoảng 8% so với cùng kỳ năm trước. Người phụ trách sản phẩm của tổ chức viết thẳng nguyên nhân trong bài công bố, các công cụ sinh câu trả lời và các nền tảng xã hội đang thay đổi cách con người tìm thông tin, đặc biệt khi công cụ tìm kiếm trả lời trực tiếp cho người dùng, thường là dựa trên chính nội dung của Wikipedia.[1]

Wikipedia không mất độc giả vì bị thay thế bằng một nguồn tốt hơn. Nó mất độc giả vì đã trở thành nguyên liệu cho thứ thay thế nó. Bách khoa toàn thư mở là một trong những nguồn được các hệ thống suy luận dựa vào nhiều nhất. Nó được xây bằng lao động tình nguyện và được nuôi bằng quyên góp cá nhân, mà cả hai dòng ấy đều phụ thuộc vào việc có người ghé thăm. Càng hữu ích cho máy, nó càng ít cần thiết với người. Càng ít cần thiết với người, nó càng ít tình nguyện viên và ít tiền để tiếp tục hữu ích cho máy. Cỗ máy uống nước từ một con sông mà nó đang lấp dần thượng nguồn.

Một địa phương đầu tư nghiêm túc vào khả đọc sẽ đối diện chính xác cùng một cấu trúc. Càng chuẩn hóa dữ liệu về mình, địa phương ấy càng dễ được máy trích dụng và càng ít lý do để

1. Marshall Miller, "New User Trends on Wikipedia", blog Diff của Wikimedia Foundation, ngày 17 tháng 10 năm 2025, https://diff.wikimedia.org/2025/10/17/new-user-trends-on-wikipedia/. Việc phân loại lại lưu lượng khiến số liệu cần diễn giải thận trọng. Cần phân biệt mức giảm 8% này với con số giảm 23% mà báo chí thường dẫn, vốn là ước lượng Similarweb trên ba năm và một định nghĩa khác.

người dùng phải ghé trang thông tin chính thức. Ngân sách cho hạ tầng khả đọc, ở nhiều nơi, lại được biện minh bằng chính lượt truy cập vào trang ấy. Thành công của khoản đầu tư sẽ phá hủy chỉ số dùng để bảo vệ nó. Đây không phải lý do để không đầu tư. Đây là lý do để đổi thước đo trước khi đầu tư và cuốn sách này sẽ trở lại vấn đề thước đo ở chương về hạ tầng.

Một cách nhìn khác. Cần ghi nhận rằng không phải ai cũng đồng ý với bức tranh này và bên không đồng ý không phải là một bên nhỏ. Công ty vận hành công cụ tìm kiếm lớn nhất thế giới đã phản bác công khai, khẳng định tổng lượt nhấp tự nhiên từ tìm kiếm sang các trang web vẫn tương đối ổn định so với cùng kỳ và lập luận rằng các bản tóm tắt sinh ra có thể làm tăng chứ không giảm hoạt động tìm kiếm.[1] Ở chiều ngược lại, một tập đoàn báo chí lớn của Anh khai báo với cơ quan cạnh tranh nước này rằng tỷ lệ nhấp chuột của họ đã giảm 89%. Hai con số ấy không thể cùng đúng theo nghĩa thông thường.[2] Nhiều khả năng cả hai đều đúng trong phạm vi của mình, tổng thể ổn định trong khi phân phối dịch chuyển dữ dội, tức là một số ít nguồn được dẫn lại nhiều hơn còn phần đông mất gần hết. Với một địa phương, kịch bản ấy tệ hơn là một sự sụt giảm đồng đều, vì nó có nghĩa là kết quả phân hóa mạnh chứ không phải mọi nguồn đều suy giảm theo cùng một mức.

1. Liz Reid, "AI in Search is driving more queries and higher quality clicks", blog Google, ngày 6 tháng 8 năm 2025, https://blog.google/products-and-platforms/products/search/ai-search-driving-more-queries-higher-quality-clicks/. Bài viết không kèm số liệu nào để kiểm chứng và không công bố dữ liệu Search Console làm căn cứ. Đây là phát ngôn của bên có lợi ích trực tiếp.

2. DMG Media, chủ sở hữu MailOnline và Metro, văn bản gửi Competition and Markets Authority của Anh tháng 7 năm 2025, tường thuật bởi BBC News, https://www.bbc.com/news/articles/c0mlvryx0exo. Trên máy tính để bàn, tỷ lệ nhấp giảm từ 25,23% xuống 2,79%, tương đương giảm 89%, trên di động giảm 87%. Đây là mức xấu nhất theo nhóm truy vấn, do một bên có lợi ích khai báo.

Từ nghịch lý Wikipedia rút ra một mệnh đề sẽ theo suốt cuốn sách. Trong nền kinh tế tác nhân, được máy sử dụng và được con người ghé thăm là hai chuyện khác nhau và chúng có thể đi ngược chiều nhau. Bất kỳ chiến lược nào đo thành công bằng lượt ghé thăm đều đang đo sai vật thể.

## 1.5. TRẢI NGHIỆM VÀ KHẢ ĐỌC: HAI LOGIC NGƯỢC CHIỀU

### 1.5.1. Cuộc tranh luận về lược đồ

Trước khi đi tiếp, cần dừng lại ở phản bác nghiêm túc nhất nhắm vào toàn bộ lập luận của cuốn sách này. Nó xứng đáng một mục riêng chứ không phải một câu ghi chú.

Phản bác như sau. Các mô hình ngôn ngữ lớn hiểu ngôn ngữ tự nhiên ngày càng tốt. Chúng rút được thông tin từ một bài viết lộn xộn, suy ra được điều không nói thẳng, xử lý được cả mô tả mâu thuẫn. Toàn bộ ngành công nghiệp dữ liệu có cấu trúc, với các lược đồ, bản thể luận và kho tri thức được biên tập thủ công, được xây trong một thời kỳ mà máy chưa đọc được văn xuôi. Thời kỳ ấy đã qua. Vậy tại sao một tỉnh phải bỏ tiền chuẩn hóa dữ liệu khi chỉ cần viết cho hay là máy sẽ hiểu.

Phản bác này không phải quan điểm của người ngoài ngành. Nó là quan điểm được nhiều kỹ sư có kinh nghiệm bảo vệ và nó dựa trên một quan sát đúng về hướng tiến bộ của công nghệ.

Có ba lý lẽ đáng cân nhắc về phía nó. Thứ nhất, chi phí. Duy trì một bộ dữ liệu có cấu trúc đòi nhân sự thường xuyên, trong khi viết nội dung tốt là việc các cơ quan quản lý điểm đến vốn đã làm. Thứ hai, tính bền. Các chuẩn kỹ thuật thay đổi nhanh, còn một bài viết tốt thì không lỗi thời. Thứ ba, bằng chứng thực tế. Nhiều địa phương chưa từng công bố dữ liệu có cấu trúc vẫn được các hệ thống mô tả tương đối chính xác, đơn giản vì có nhiều người viết về chúng.

Trả lời ba lý lẽ ấy đòi hỏi phải tách hai việc mà cuộc tranh luận thường gộp lại, đó là hiểu và làm. *Ở phần hiểu*, phản bác đúng và

cuốn sách này chấp nhận nó không kèm điều kiện. Với các thuộc tính mô tả, tức là không khí của một khu phố, đặc điểm của một món ăn, lý do một nơi đáng đến, khả năng đọc văn bản tự do đã tiến tới mức việc chuẩn hóa mang lại lợi ích biên nhỏ. Một tỉnh dồn tiền để mã hóa những thứ ấy thành trường dữ liệu đang giải một bài toán đã được giải hộ. *Ở phần làm*, phản bác không áp dụng được. Một hệ thống có thể hiểu rất tinh một đoạn văn về một nhà hàng mà vẫn không giữ được bàn cho ai, vì giữ bàn đòi hỏi một câu trả lời đúng ở thời điểm hỏi về việc còn hay hết. Không lượng văn xuôi nào chứa thông tin ấy vì thông tin ấy thay đổi từng giờ.

Sự phân biệt này cho phép rút ra một kết luận hẹp hơn nhưng chắc hơn kết luận thông thường. Tiến bộ trong đọc hiểu không xóa nhu cầu về khả đọc. Nó dịch chuyển nhu cầu ấy từ tầng mô tả sang tầng định danh và tầng giao dịch. Cái cần chuẩn hóa là tên gọi, mã, ranh giới, quan hệ hành chính, tồn kho, giá và quyền, chứ không phải câu chuyện về một vùng đất.

Kết luận trên có một hệ quả cho những người đang bán các dự án nội dung số. Nếu tầng mô tả đang mất giá trị biên, thì phần lớn ngân sách truyền thông số của các địa phương hiện nay đang chảy vào đúng tầng mà máy đã tự làm được, còn hai tầng thực sự khan hiếm thì không ai phụ trách. Đây không phải lập luận đòi thêm tiền cho công nghệ. Đó là lập luận đòi chuyển tiền từ chỗ này sang chỗ khác và loại đề xuất ấy luôn khó hơn.

Có một trường hợp mà phần mô tả vẫn cần chuẩn hóa và bỏ qua nó sẽ dẫn tới lời khuyên sai với chính Việt Nam. Khả năng đọc hiểu văn bản tự do phụ thuộc vào khối lượng ngữ liệu của ngôn ngữ đó. Với các ngôn ngữ ít ngữ liệu hơn tiếng Anh, năng lực ấy thường kém ổn định hơn và phụ thuộc mạnh hơn vào lĩnh

vực, phương ngữ cùng lượng ngữ liệu sẵn có. Lời khuyên đừng chuẩn hóa phần mô tả vì thế đúng ở thị trường Anh ngữ và chưa chắc đúng ở nơi khác.

### 1.5.2. Trải nghiệm là thứ khó số hóa nhất

Đến đây dễ tưởng rằng lời giải chỉ là một dự án dữ liệu gọn gàng, chuẩn hóa định danh, mở giao diện truy vấn, để nguyên phần mô tả cho máy tự đọc. Nếu vậy thì cuốn sách này là một cẩm nang kỹ thuật mỏng. Vấn đề nằm sâu hơn và nó nằm trong di sản của cuốn trước.

Lợi thế bền nhất của một địa phương là những trải nghiệm giàu ngữ cảnh. Bữa cơm trong gian bếp của người bản địa. Nghi lễ chỉ có nghĩa với ai hiểu lịch sử của nó. Nhịp sống của một khu phố không giống bất cứ khu phố nào. Giá trị của chúng nằm ở chỗ không tách được khỏi ngữ cảnh, không lặp lại được, không nói hết được bằng lời. Michael Polanyi gọi đó là tri thức ẩn, chúng ta biết nhiều hơn những gì nói ra được.[1]

Khả đọc, ở mọi tầng của nó, là một phép nén. Muốn vào tập lựa chọn, trải nghiệm phải được đo bằng thang chung để so sánh được với hàng nghìn trải nghiệm khác. Món ăn của bà cụ trong hẻm trở thành một mục nhà hàng với tọa độ, giờ mở cửa, khoảng giá, điểm trung bình. Phần nén được đi qua bộ lọc. Phần không nén được, tức phần làm nên lý do đáng đến, nằm lại phía trên.

1. Michael Polanyi, The Tacit Dimension, Doubleday, New York, 1966. Bản tái bản thường được dẫn của University of Chicago Press, Chicago, 2009.

### 1.5.3. Bản sắc càng sâu càng khó khả đọc

Từ đó rút ra một tương quan, nhìn chung, chiều sâu ngữ cảnh của một tài sản văn hóa càng lớn thì chi phí làm cho nó khả đọc mà không làm nghèo nó càng cao.

Một bãi biển rất dễ khả đọc. Nó có tọa độ, chiều dài, chất lượng nước, mùa đẹp nhất. Một bảo tàng khó hơn một chút nhưng vẫn nằm gọn trong lược đồ. Một lễ hội làng đã bắt đầu chật chội, vì nó rơi vào ô sự kiện, mà ô sự kiện đòi ngày bắt đầu và ngày kết thúc, trong khi nhiều lễ hội không có ngày cố định vì chúng theo lịch âm, theo mùa vụ hoặc theo quyết định của các cụ trong làng. Còn một tập quán sống, một cách ứng xử, một quan niệm về thời gian thì không có ô nào.

Các hệ thống mô tả thế giới vì thế ưu ái những nền văn hóa đã tự chuẩn hóa mình từ trước và các nền văn hóa ấy thường là các nền văn hóa đã từng xây bảo tàng theo mẫu Pháp, lập danh mục di sản theo mẫu Anh. Bất bình đẳng trong khả đọc không phải sản phẩm của trí tuệ nhân tạo. Trí tuệ nhân tạo là lớp khuếch đại mới nhất của một bất bình đẳng đã có từ trước.

Bức tranh này không một chiều. Trong nhiều truy vấn về du lịch, nguồn được các hệ thống dẫn lại nhiều nhất không phải trang chính thức của một quốc gia mà là các nền tảng đánh giá của người dùng và các diễn đàn. Một cộng đồng nhỏ, đông người kể chuyện, có thể tạo ra khối lượng bằng chứng vượt xa một cơ quan nhà nước im lặng. Khả đọc không hoàn toàn là trò chơi của kẻ mạnh. Nó là trò chơi của kẻ để lại nhiều dấu vết nhất quán nhất và dấu vết thì không đắt.

### 1.5.4. Nghịch lý của địa phương

Logic trải nghiệm và logic khả đọc cùng tạo ra giá trị cho địa phương nhưng phục vụ hai đối tượng khác nhau. Logic trải nghiệm hướng tới con người, logic khả đọc hướng tới máy suy luận. Nếu trải nghiệm được đánh giá bằng chiều sâu, cảm xúc và sự khác biệt, thì khả đọc được đánh giá bằng sự rõ ràng, nhất quán và khả năng so sánh. Vì vậy, điều con người yêu thích thường phải được chuyển hóa thành thực thể, thuộc tính và quan hệ trước khi máy có thể hiểu và hành động.

Hai logic này cũng tích lũy giá trị theo những cách khác nhau. Trải nghiệm cần nhiều năm để hình thành qua ký ức và truyền miệng, khả đọc có thể được xây dựng nhanh bằng dữ liệu chuẩn hóa nhưng cũng mất rất nhanh khi thông tin lỗi thời hoặc phân mảnh. Rủi ro vì thế cũng trái ngược. Chỉ đầu tư vào trải nghiệm, địa phương có thể rất hấp dẫn nhưng không được AI tìm thấy. Chỉ tối ưu cho khả đọc, địa phương có thể dễ được đề xuất nhưng đánh mất bản sắc và sức hấp dẫn. Vì vậy, trong nền kinh tế tác nhân, logic khả đọc không thay thế logic trải nghiệm mà bổ sung cho nó. Trải nghiệm khiến con người muốn đến, khả đọc khiến AI có thể tìm thấy, hiểu đúng và lựa chọn. Địa phương chỉ thực sự cạnh tranh khi đồng thời chinh phục được cả con người lẫn máy.

Điều đáng chú ý là hai logic này không chỉ khác nhau về cách chiến thắng mà còn khác nhau về cách thất bại. Logic trải nghiệm thất bại khi con người không bị thuyết phục. Logic khả đọc thất bại khi AI không nhìn thấy, không hiểu hoặc không thể hành động. Một bên thất bại sau khi đã bước vào cuộc cạnh tranh, bên kia thất bại trước cả khi cuộc cạnh tranh bắt đầu. Địa phương chỉ theo lực thứ nhất trở thành viên ngọc không ai tìm thấy. Địa

phương chỉ theo lực thứ hai trở thành hàng hóa hoàn hảo và vô hồn, so sánh được với mọi nơi và không đáng nhớ hơn nơi nào.

Nghịch lý này không có lời giải triệt để. Bất kỳ ai hứa với một tỉnh rằng chỉ cần chuẩn hóa dữ liệu là vừa giữ được bản sắc vừa thắng trong hệ thống suy luận đều đang bán một thứ không tồn tại. Cái có thể làm là quyết định, một cách có ý thức, phần nào đi qua và phần nào giữ lại. Đó là cái van và chương sau sẽ đặt nó vào mô hình.

Có một cách nhìn bi quan hơn về nghịch lý này. Nếu phần khả đọc là phần duy nhất sinh ra doanh thu, thì theo thời gian mọi nguồn lực sẽ chảy về phần ấy, bất kể ban đầu người ta có ý định giữ cân bằng hay không. Không cần ai quyết định bỏ rơi phần sâu. Chỉ cần ngân sách được phân bổ theo chỉ số đo được, năm này qua năm khác, là phần không đo được sẽ teo dần. Đây là lý do các chương sau lập luận rằng cái van phải là một thiết chế có thẩm quyền chứ không phải một nguyên tắc đạo đức nằm trong lời nói đầu.

## 1.6. CHÚNG TA ĐỨNG Ở ĐÂU TRONG BƯỚC CHUYỂN NÀY

### 1.6.1. Những con số tăng trưởng dựa trên hệ thống cũ

Sáu tháng đầu năm 2026, Việt Nam đón gần 12,3 triệu lượt khách quốc tế, tăng khoảng 14,9 phần trăm so với cùng kỳ, hoàn thành gần một nửa mục tiêu 25 triệu lượt của cả năm. Du lịch nội địa phục vụ khoảng 81 triệu lượt, tổng doanh thu toàn ngành đạt khoảng 569 nghìn tỷ đồng.[1] Đó là những con số đạt chuẩn theo mọi tiêu chí đánh giá thông thường. Những chỉ số này ghi nhận các kết quả từ hệ thống vận hành cũ, bao gồm các chính sách thị thực, đường bay thẳng, giá cả cạnh tranh, xúc tiến quảng bá. Tuy nhiên, các chỉ số ngành hiện tại chưa được thiết kế để đo vị thế trước các hệ thống suy luận. Một quốc gia có thể vừa đồng thời phá kỷ lục về lượng khách, vừa có thể mất chỗ đứng trong tập lựa chọn của máy. Hai sự thật ấy ít khi gặp nhau cho tới khi đường cong đầu tiên đi xuống. Đây chính là tình huống của người chủ homestay ở đầu chương, nhân lên quy mô thương hiệu và cạnh tranh của ngành.

### 1.6.2. Nơi quyết định thực sự được đưa ra

Trong nửa đầu năm 2026, khoảng 10,12 triệu lượt khách quốc tế tới Việt Nam bằng đường hàng không, chiếm 82,6 phần trăm tổng lượng khách.[2] Con số ấy thường được đọc như chỉ báo về hạ tầng. Nó cũng là chỉ báo về địa lý của quyết định.

1. Cục Thống kê, Bộ Tài chính, số liệu 6 tháng đầu năm 2026, công bố tháng 7 năm 2026. Số liệu du lịch nội địa và tổng thu toàn ngành theo Cục Du lịch Quốc gia Việt Nam cùng kỳ. Lượt khách nội địa là số lượt phục vụ chứ không phải số người, nên không cộng trực tiếp được với lượt khách quốc tế.

2. Cục Thống kê, Bộ Tài chính, 6 tháng đầu năm 2026. Khách đến bằng đường bộ đạt 1,92 triệu lượt, tăng 37,5 phần trăm, là mức tăng cao nhất trong ba loại hình vận chuyển.

Một người bay tới Việt Nam đã quyết định trước khi lên máy bay. Toàn bộ chuỗi cân nhắc, so sánh, loại trừ và chọn lựa diễn ra ở nơi người ấy sống, bằng ngôn ngữ người ấy dùng, trên hạ tầng thông tin mà Việt Nam không vận hành và phần lớn thời gian cũng không xuất hiện. Việt Nam kiểm soát gần như tuyệt đối phần trải nghiệm sau khi khách hạ cánh, nhưng dường như chưa thể có mặt ở khâu quyết định của du khách xem có hạ cánh hay không.

Bất đối xứng ấy không phải phát minh của thời tác nhân. Cái mới là khâu tuyển chọn ngoài lãnh thổ đang chuyển từ hàng triệu quyết định phân tán sang một số ít hệ thống tập trung. Trước đây, để lọt vào tầm nhìn của khách Đức, một tỉnh cần lọt vào tầm nhìn của hàng nghìn blogger, biên tập viên và đại lý người Đức. Bây giờ nó cần lọt qua bốn bộ lọc của vài hệ thống.

### 1.6.3. Bài học từ thị trường Nga: Tăng trưởng đến từ đường bay

Sáu tháng đầu năm 2026, thị trường Nga đạt khoảng 742,7 nghìn lượt khách, tăng 185,8 phần trăm so với cùng kỳ, vươn lên vị trí thị trường lớn thứ ba của du lịch Việt Nam.[1] Mức tăng trưởng đột biến ấy được các cơ quan quản lý chỉ rõ là nhờ vào việc khôi phục và mở rộng các tuyến bay thẳng kết hợp nhu cầu nghỉ dưỡng biển, lợi thế khí hậu và chi phí cạnh tranh, không liên quan đến nội dung số.

Trường hợp này là một ví dụ thực tế cho toàn bộ lập luận của chương. Nó cho thấy trong nhiều tình huống, một quyết định

1. Cục Thống kê, Bộ Tài chính và Cục Du lịch Quốc gia Việt Nam, sáu tháng đầu năm 2026. Mức tăng tính trên nền so sánh rất thấp của cùng kỳ năm trước.

về hạ tầng vật lý vẫn tạo ra kết quả lớn hơn bất kỳ chương trình dữ liệu số nào. Do đó, bất kỳ ai định thuyết phục một sở du lịch dồn ngân sách vào "giải pháp khả đọc" đều nên chuẩn bị trước câu hỏi vì sao không dùng số tiền ấy để hỗ trợ mở thêm một đường bay. Nhưng trường hợp này cũng chứa một cảnh báo. Tăng trưởng đến từ đường bay thẳng là loại tăng trưởng phụ thuộc vào một biến số duy nhất và biến số ấy chịu chi phối của địa chính trị, giá nhiên liệu và quyết định thương mại của các hãng bay. Khi nó đảo chiều, sẽ không có lớp nền nào bên dưới đỡ lấy. Một thị trường được xây dựng bằng đường bay có thể mất đi cùng đường bay ấy. Nhưng một thị trường được xây bằng vị thế trong tập lựa chọn thì không.

Trên thực tế hai lớp không tách rời như vậy và chúng tương tác theo một chiều đáng chú ý, chính dòng khách lớn từ một thị trường mới sẽ sinh ra khối lượng đánh giá, ảnh và dấu vết giao dịch bằng ngôn ngữ của thị trường ấy, tức là chính nó bồi đắp khả đọc cho các điểm đến liên quan. Đường bay không thay thế dữ liệu. Nó sản xuất ra dữ liệu. Kết luận thực hành vì thế không phải chọn một trong hai mà là nhận ra rằng thời điểm tốt nhất để đầu tư vào khả đọc là ngay khi một dòng khách mới bắt đầu, chứ không phải sau khi nó đã đi qua và các dấu vết đã nằm rải rác không ai thu gom.

### 1.6.4. Vì sao thành công hiện tại làm chậm phản ứng

Có một quy luật quen thuộc trong lịch sử các cuộc chuyển kênh. Tổ chức đang thắng trong luật chơi cũ luôn là tổ chức chậm nhất trong việc chuyển sang luật chơi mới. Không phải vì lãnh đạo kém mà vì mọi tín hiệu nội bộ đều nói rằng chiến lược hiện tại đang hiệu quả. Ngân sách chảy về nơi có kết quả đo được. Người

được thăng tiến là người tạo ra kết quả ấy. Đề xuất đầu tư vào một hạng mục chưa có chỉ số, chưa có tiền lệ và chưa có kết quả trong ba năm đầu sẽ thua trong mọi cuộc họp phân bổ nguồn lực và nó thua một cách hoàn toàn hợp lý theo các tiêu chí đang được dùng.

Rào cản lớn nhất với chương trình khả đọc ở Việt Nam không phải công nghệ, ngân sách hay nhận thức của người quản lý. Đó là bộ chỉ số. Chừng nào ngành còn báo cáo bằng lượt khách và doanh thu, mọi khoản đầu tư vào vị thế trước máy đều là khoản chi không giải trình được. Bổ sung một chỉ số mới vào hệ thống báo cáo ngành nghe như một đề xuất hành chính bình thường. Nó có thể là can thiệp chính sách rẻ nhất và có sức đòn bẩy lớn nhất trong toàn bộ cuốn sách này.

Một điểm cân bằng cần nêu trước khi khép lại mục này. Chính sự khan hiếm tín hiệu lại khiến giá trị biên của mỗi hành động chuẩn hóa dữ liệu tại Việt Nam cao hơn nhiều so với các quốc gia đã bão hòa dữ liệu. Trong môi trường mà phần lớn các thực thể chưa được mô tả định danh ..... vùng trên không gian số thì bên nào chủ động mô tả trước sẽ trở thành nguồn dữ liệu mặc định và vị thế nguồn mặc định có tính tự củng cố theo thời gian. Các điều kiện thể chế để tận dụng cơ hội ấy, cùng những thách thức đi kèm, thuộc chương về hạ tầng và chương về năng lực cạnh tranh.

Chương 7 trở lại chủ đề này với dữ liệu trong nước và các đề xuất tới cấp tỉnh. Ở đây chỉ cần ghi nhận một điều, chi phí giành lấy quyền định nghĩa hiện còn thấp và giai đoạn ấy khép lại không phải khi ta từ bỏ mà khi có bên khác lấp vào chỗ trống.

## 1.7. MỘT CUỘC CẠNH TRANH HOÀN TOÀN MỚI

### 1.7.1. Từ cạnh tranh điểm đến sang cạnh tranh khả đọc

Trong nhiều thập niên, các địa phương cạnh tranh nhau bằng lợi thế sẵn có, gắn với những bãi biển đẹp, lễ hội có quy mô lớn, sân bay mới hơn, chiến dịch quảng bá rầm rộ. Cuộc cạnh tranh mới không diễn ra ở bất cứ chỗ nào trong số đó. Nó diễn ra ở tầng mà không ai chụp ảnh được.

Khi thua trong cạnh tranh điểm đến truyền thống, người ta biết. Lượng khách giảm, hội nghị được triệu tập, tư vấn được thuê. Khi thua trong cạnh tranh khả đọc, mọi chỉ số nội bộ vẫn xanh, vì chúng đều được thiết kế cho thời nền tảng và tiếp tục nhấp nháy bình thường trong lúc tập lựa chọn của máy đã khép lại. Một địa phương có thể đang thắng mọi chỉ số nó theo dõi và thua đúng chỉ số nó không theo dõi. Thông tin về thất bại ở đây chỉ đo được từ bên ngoài, bằng cách chạy hàng nghìn truy vấn mô phỏng và đếm xem tên mình xuất hiện bao nhiêu lần. Rất ít cơ quan quản lý điểm đến làm việc đó thường xuyên.

Phép đo này rẻ. Nó không đòi hạ tầng, không đòi hợp đồng với bên vận hành nền tảng. Nó đòi một người kiên nhẫn, một danh sách câu hỏi mà khách hàng mục tiêu thực sự hỏi và một bảng tính. Việc đầu tiên nên làm cũng là việc rẻ nhất và chính vì rẻ nên nó thường bị bỏ qua để nhường chỗ cho những đề án nghe hoành tráng hơn.

### 1.7.2. Địa phương sẽ phải quản trị điều gì

Để quản trị hiệu quả, hệ thống vận hành cần đảm bảo các nguyên tắc cơ bản sau:

- ***Định danh:*** Thiết lập một mã định danh dữ liệu chuẩn hóa cơ bản mà mọi nguồn khác quy chiếu về. Đây là cơ sở nền tảng để tất cả các nguồn thông tin đến sau đối chiếu và đồng bộ.
- ***Thẩm quyền:*** Cần một nguồn được xác định là nguồn gốc chính thống cho từng loại dữ kiện và cần cơ chế để nguồn ấy được các hệ thống bên ngoài nhận ra là có thẩm quyền, thay vì chỉ là một trang trong hàng nghìn trang khác.
- ***Tính nhất quán:*** Nguyên nhân thất bại lớn nhất của các địa phương không phải do thiếu dữ liệu mà vì có quá nhiều dữ liệu mâu thuẫn do nhiều cơ quan cùng công bố thông tin nhưng không ai chịu trách nhiệm làm sạch.
- ***Khả giao dịch:*** Khả năng đảm bảo để một hệ thống bên ngoài có thể truy vấn thông tin, kiểm tra tính sẵn có và nhận về câu trả lời chính xác tại thời điểm hỏi.
- ***Cơ chế bảo mật:*** Cái van chốt chặn quan trọng nhất Công việc này thuộc về chính sách quản trị hệ thống nội bộ, không phải là vấn đề kỹ thuật. Vì vậy, không thể giao trách nhiệm này cho nhà thầu công nghệ.

Kinh nghiệm của các dự án chính phủ điện tử trong thời gian qua cho thấy phần khó nhất chưa bao giờ là xây dựng hệ thống mà là duy trì nó sau khi đã hoàn thiện. Dữ liệu về giờ mở cửa sai sau mười tám tháng còn tệ hơn không có dữ liệu, vì nó khiến máy tin vào một điều không đúng thay vì thừa nhận là chưa biết. Bất kỳ chiến lược khả đọc nào không kèm ngân sách vận hành thường xuyên đều là một khoản chi một lần, bởi vì người ta gọi nó là chiến lược dài hạn.

### 1.7.3. Khả đọc trở thành hạ tầng mới

Máy hơi nước không làm nên cách mạng công nghiệp vì nó tạo ra nhiều năng lượng hơn, mà vì nó buộc người ta tổ chức lại sản xuất. Điện không quan trọng vì nó thay động cơ hơi nước, mà vì nó cho phép nhà máy được thiết kế lại từ đầu. Trong cả hai trường hợp, khoảng cách giữa lúc công nghệ xuất hiện và lúc năng suất tăng kéo dài hàng chục năm và cái lấp đầy khoảng cách ấy là hạ tầng cùng thói quen tổ chức.

Khả đọc là loại hạ tầng ấy. Nó không cắt băng khánh thành được và nó có đặc tính kinh điển của hạ tầng, chi phí trả trước, lợi ích trả sau, rất khó bảo vệ trước một hội đồng đang phải chọn giữa nó với đầu tư xây một cây cầu.

Giá trị biên của khả đọc cao nhất ở những nơi ít được biết đến nhất. Một thành phố lớn đã có hàng trăm triệu trang viết về mình, nên máy vẫn suy ra được điều cần suy dù dữ liệu chính thức có lộn xộn. Một xã chỉ xuất hiện vài trăm lần trên internet thì mỗi mảnh dữ liệu chuẩn hóa đều làm thay đổi đáng kể xác suất được nhắc tới. Nơi nghèo tín hiệu là nơi mỗi đồng đầu tư vào khả đọc sinh lợi cao nhất và cũng là nơi thường không có ai đề xuất khoản đầu tư ấy.

Tin tốt ấy đi kèm một tin xấu ít được nhắc đến. Thẩm quyền trước máy không chỉ tích lũy chậm mà còn có thể thay đổi nhanh sau các lần cập nhật mô hình. Một số nghiên cứu và báo cáo gần đây cho thấy, việc nâng cấp mô hình có thể làm thay đổi đáng kể cách hệ thống truy hồi và trích dẫn nguồn, ngay cả khi nội dung của các trang web không thay đổi. Chẳng hạn, một phân tích trên cùng một tập câu hỏi ghi nhận số lượng nguồn được AI trích dẫn trung bình giảm gần một nửa sau một lần

nâng cấp mô hình, trong khi mô hình đối chứng không thay đổi. Điều đó cho thấy khả năng được hệ thống lựa chọn không chỉ phụ thuộc vào chất lượng nội dung mà còn chịu ảnh hưởng bởi các thay đổi trong chính kiến trúc của mô hình. Tuy nhiên, các kết quả này cần được hiểu thận trọng vì chủ yếu dựa trên báo cáo thực nghiệm của doanh nghiệp, chưa phải bằng chứng học thuật kết luận.

### 1.7.4. Khi khả đọc là công cụ phòng vệ

Khả đọc thường được bàn như công cụ để được chọn. Trường hợp dưới đây cho thấy nó cũng là một loại công cụ phòng vệ. Venice đi theo hướng ngược lại và vì một lý do ngược lại. Thành phố này không thiếu sự nhìn thấy. Nó thừa. Vấn đề của Venice chưa bao giờ là làm sao để nhiều người biết tới hơn mà là làm sao chịu đựng được số người đã biết tới. Cơ chế thu phí vào thành phố mà Venice áp dụng trong những năm gần đây, với mức phí khác nhau, tùy theo người đến đăng ký sớm hay muộn và chỉ áp dụng trong một số ngày cao điểm nhất định trong khung giờ ban ngày, là một dự án khả đọc hóa, dù không ai gọi nó bằng tên ấy. Muốn thu phí, phải đếm được. Muốn đếm được, phải phân biệt được khách trong ngày với cư dân, với người đi làm, với khách lưu trú. Muốn phân biệt được, phải có đăng ký, mã, cổng kiểm tra và dữ liệu.

Venice không phải trường hợp duy nhất và cũng không phải trường hợp điển hình nhất. Barcelona chọn một tổ hợp công cụ khác, giới hạn số tàu biển được cập bến, siết giấy phép cho thuê căn hộ ngắn hạn và công bố kế hoạch chấm dứt loại giấy phép ấy trong thập niên này. Amsterdam đi xa hơn về phía truyền thông, với một chiến dịch chủ động khuyên một số nhóm khách đừng

đến, song song với việc dừng cấp phép xây khách sạn mới. Ba thành phố, ba điểm can thiệp khác nhau trong cùng một chuỗi. Venice can thiệp ở cửa vào, Barcelona ở nguồn cung lưu trú, Amsterdam ở tầng nhu cầu. Điều đáng chú ý với cuốn sách này là cả ba đều phải làm cho quy định của mình trở nên khả đọc trước khi nó có hiệu lực với một tác nhân. Một mức phí không được máy đọc thì không vào được phép tính của máy và một giới hạn giấy phép không được công bố ở dạng máy tra được thì không ngăn được một đề xuất sai.

Điều đáng chú ý về mặt lý thuyết ở đây là khả đọc không nhằm mục đích phục vụ việc chọn. Nó phục vụ việc điều tiết. Cùng một hạ tầng cho phép một địa phương xuất hiện trong tập lựa chọn của máy cũng cho phép nó đặt điều kiện cho sự xuất hiện ấy. Đây là mầm mống của khái niệm "van" và Venice cho thấy "cái van" không nhất thiết là một ý niệm đạo đức trừu tượng. Nó có thể là một cái cổng, một mã đăng ký và một mức phí.

Hai trường hợp trên đến từ những nơi có năng lực hành chính và ngân sách lớn. Một xã/phường có ba người phụ trách truyền thông kiêm nhiệm không thể vận hành cổng kiểm soát kiểu Venice. Với họ, toàn bộ chương trình khả đọc thực tế có thể chỉ gồm ba việc: thống nhất một cách viết tên, công bố một bộ dữ liệu cơ bản đúng và cập nhật nó mỗi quý. Ba việc ấy nghe có vẻ bình thường đến mức không ai muốn đưa vào trong các cuộc họp quan trọng nhưng chúng lại tạo ra phần lớn giá trị.

### 1.7.5. Dẫn nhập cho các chương tiếp theo

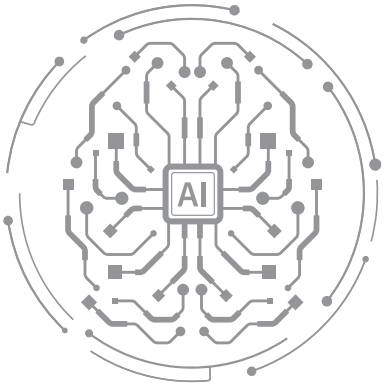
Chương này mô tả hiện tượng và đặt tên tạm cho nó. Việc biến tên gọi ấy thành khái niệm kiểm tra được là công việc của chương sau. Chương 2 truy về phả hệ của khái niệm khả đọc, từ nghiên cứu của James C. Scott về cách nhà nước buộc xã hội tự trình bày theo lược đồ của mình, tới nghiên cứu của Bowker và Star về việc mọi hệ phân loại đều đồng thời là một cỗ máy lãng quên. Từ đó chương dựng mô hình bốn bộ lọc, một vòng và một van, rồi đem mô hình ra đối chất với các phản bác mạnh nhất.

Người chủ homestay ở đầu chương không thiếu tài sản, không thiếu nỗ lực và cũng không thiếu khách hài lòng. Ông thiếu sự tồn tại rõ ràng trong mắt một cỗ máy mà ông chưa từng gặp, chưa từng ký hợp đồng và không có cách nào khiếu nại. Trong nền kinh tế trải nghiệm, địa phương cạnh tranh để được nhớ. Trong nền kinh tế tác nhân, nó phải cạnh tranh để được hiểu và phải làm việc đó trước, bằng một ngôn ngữ không dành cho con người.

## CHƯƠNG 2

# TÍNH KHẢ ĐỌC THUẬT TOÁN

*Phả hệ chính trị của khái niệm khả đọc, từ Scott, Bowker và Star tới thuật toán. Chương đề xuất mô hình bốn bộ lọc, một vòng và một van, mô tả cơ chế tái cấu trúc và đối chất mười một phản bác.*


AI

Câu hỏi của chương này là thế giới đang tự tổ chức lại thế nào để được máy đọc, chứ không phải máy đọc thế giới thế nào. Khả đọc thuật toán là điều kiện tồn tại và cái giá của nó là một cuộc tái cấu trúc mà không ai ra lệnh và không ai chịu trách nhiệm.

## 2.1. MỘT SỰ VẮNG MẶT KHÔNG AI BÁO CÁO

### 2.1.1. Khi quyết định bắt đầu trước khi con người tìm kiếm

Chương trước để lại một hiện tượng chưa có tên. Trước khi đặt tên cho nó, cần nhìn kỹ hình thù của nó, vì nó khác về loại với mọi hình thức mà lý thuyết thương hiệu từng biết cách xử lý. Các thất bại quen thuộc của thương hiệu địa phương đều nhìn thấy được. Nhận biết thấp thì đo bằng khảo sát. Hình ảnh xấu thì đọc trên báo. Khách giảm thì hiện trong thống kê. Với mỗi loại, ngành đã có thước đo, có thuật ngữ, có phác đồ và quan trọng nhất là có một thời điểm mà ai đó trong tổ chức nhận ra rằng có vấn đề.

Sự vắng mặt trước máy không thuộc họ ấy. Nó không phải bị đánh giá thấp mà là không được đánh giá. Khác biệt giữa hai điều đó không phải khác biệt về mức độ. Bị đánh giá thấp nghĩa là đã bước vào phòng thi, đã nộp bài, đã bị chấm điểm kém. Không được đánh giá nghĩa là tên không có trong danh sách dự thi. Người bị điểm kém biết mình cần học gì. Người không có tên không biết là có kỳ thi.

### 2.1.2. Sự biến mất khỏi tập lựa chọn

Hiện tượng này có một tính chất khiến nó gần như miễn nhiễm với các hệ thống giám sát hiện hành, nó không để lại dữ liệu ở phía nạn nhân. Truy vấn diễn ra trên hạ tầng của người khác. Câu trả lời được trả cho một người dùng mà địa phương không bao giờ biết đến sự tồn tại. Giữa hai đầu ấy, ở một tầng xử lý không ai nhìn thấy, tên của một vùng đất đã bị loại khỏi tập ứng viên. Không có bản ghi nào được gửi đi. Không có thông báo từ chối. Sự loại bỏ diễn ra hoàn hảo theo nghĩa nó không tốn của ai một giây nào để thực hiện và không để lại một dấu vết nào để điều tra.

Nhà nghiên cứu truyền thông Taina Bucher đã mô tả cơ chế tâm lý của tình trạng này từ năm 2012, khi phân tích cách các thuật toán phân phối nội dung trên mạng xã hội tạo ra cái mà bà gọi là nỗi đe dọa của sự vô hình. Điểm sắc bén trong lập luận của Bucher không nằm ở chỗ thuật toán làm ai đó vô hình. Nó nằm ở chỗ chính khả năng bị vô hình đã đủ để định hình hành vi của tất cả những người còn lại, kể cả những người chưa bị. Người dùng bắt đầu đăng bài theo cách mà họ đoán là thuật toán ưa thích, không phải vì đã bị phạt mà vì sợ bị phạt.[1]

Với các địa phương, cơ chế ấy còn chưa kịp hình thành và đó mới là điều đáng lo ngại. Nỗi sợ vô hình đòi hỏi người ta phải biết rằng sự vô hình tồn tại. Phần lớn các cơ quan quản lý điểm đến hiện nay chưa nhận diện rõ mối đe dọa vô hình này. Nguyên nhân không phải vì họ chủ quan hay quá tự tin mà do họ chưa xác định được vị trí của mình đang ở đâu trong tiến trình chuyển đổi ấy.

1. Taina Bucher, "Want to be on the top? Algorithmic power and the threat of invisibility on Facebook", New Media and Society, tập 14, số 7, 2012, trang 1164 tới 1180.

### 2.1.3. Ba đặc tính của sự vô hình

Ba đặc tính dưới đây phân biệt hiện tượng này với mọi rủi ro truyền thông thông thường và cả ba đều đi theo hướng bất lợi.

**Không có báo cáo**

Đặc tính thứ nhất đã nêu ở trên nhưng cần được đẩy tới hệ quả tổ chức của nó. Trong một cơ quan hành chính, cái gì không có báo cáo thì không có người phụ trách và cái gì không có người phụ trách thì không có ngân sách. Một giám đốc sở du lịch có thể hoàn toàn hiểu vấn đề, hoàn toàn đồng ý rằng nó nghiêm trọng nhưng vẫn không thể đề xuất một khoản kinh phí để xử lý, đơn giản vì không có con số nào để đưa vào tờ trình.

Điểm bất đối xứng đáng chú ý là hiện tượng này đo được, nhưng chỉ đo được từ bên ngoài. Bất kỳ ai cũng có thể chạy vài nghìn truy vấn mô phỏng và đếm tần suất mỗi địa phương xuất hiện. Đấy là công việc đơn giản, không tốn kém nhiều chi phí, không đòi hỏi sự hợp tác của bên vận hành nền tảng và có thể làm bằng một bảng tính nhưng không được làm vì không ai được giao làm.

**Tích lũy theo thời gian**

Đặc tính thứ hai nguy hiểm hơn nhiều, vì nó biến một vấn đề tĩnh thành một vấn đề động. Mỗi lần một địa phương được chọn, giao dịch diễn ra và sinh ra dấu vết mới, như đánh giá, ảnh, dữ liệu vị trí, lịch sử thanh toán. Các dấu vết ấy quay lại làm đầu vào cho lần đọc sau. Kẻ được chọn hôm nay không chỉ thắng hôm nay. Nó dày thêm một lớp bằng chứng cho ngày mai. Kẻ không được chọn thì không sinh ra gì cả. Sự im lặng của nó tự nuôi chính nó.

Robert Merton đặt tên cho cấu trúc này từ năm 1968 khi nghiên cứu cách uy tín tích lũy trong giới khoa học và gọi đó là hiệu ứng Matthew, theo câu trong Phúc âm Matthew: *"Ai đã có thì sẽ được cho thêm."* Merton chỉ ra rằng, với những đóng góp có giá trị tương đương, các nhà khoa học đã có uy tín thường nhận được sự ghi nhận và mức độ hiển thị lớn hơn, trong khi những nhà khoa học ít tên tuổi dễ bị đánh giá thấp hoặc ít được chú ý hơn. Khoảng cách ấy có xu hướng tự khuếch đại theo thời gian thông qua cơ chế tích lũy lợi thế.[1] Điều mới trong nền kinh tế tác nhân không phải hiệu ứng Matthew. Đó là tốc độ của nó. Trong khoa học, một chu kỳ trích dẫn mất nhiều năm. Trong hệ thống suy luận, chu kỳ ấy mất vài tuần và nó chạy tự động, không cần ai quyết định gì.

**Gần như không đảo ngược**

Đặc tính thứ ba là hệ quả của hai đặc tính đầu và là lý do chương này khẩn thiết hơn giọng điệu thông thường của một cuốn sách học thuật. Khi một mô tả sai đã ổn định trong hệ thống, việc sửa nó không giống việc đính chính một bài báo. Không có tòa soạn để gửi thư. Không có ai chịu trách nhiệm về câu tóm tắt sáu mươi bảy từ kia. Cách duy nhất để sửa là thay đổi khối bằng chứng bên dưới, tức là làm cho các nguồn độc lập cùng nói khác đi và việc đó mất nhiều năm chứ không phải nhiều tuần.

Cụm từ không đảo ngược cần được dùng thận trọng hơn mức mà giọng văn cảnh báo thường cho phép. Các hệ thống suy luận cập nhật liên tục và đã có những trường hợp một nguồn mới, có thẩm quyền rõ ràng, làm thay đổi cách một thực thể được mô tả

1. Robert K. Merton, "The Matthew Effect in Science", Science, tập 159, số 3810, 1968, trang 56 tới 63.

khá nhanh. Chữ chính xác hơn không phải không đảo ngược mà là bất đối xứng, mất uy tín nhanh hơn nhiều so với gây dựng lại và chi phí sửa tăng theo mỗi vòng lặp mà cái sai được nhắc lại. Sự khác nhau giữa hai cách nói không nhỏ. Một cái dẫn tới tuyệt vọng, cái kia dẫn tới việc sửa sớm.

### 2.1.4. Đặt tên cho vấn đề

Điều còn thiếu. Vậy là, không phải một chiến thuật mà là một khái niệm. Cần một cách gọi tên phẩm chất mà kẻ được máy nhìn thấy sở hữu còn kẻ vắng mặt thì không. Phẩm chất ấy không phải chất lượng. Có những nơi tệ hơn mà hiện diện đậm hơn. Không phải danh tiếng, vì có những cái tên lừng lẫy với con người nhưng lại mờ nhạt với máy. Không phải ngân sách quảng bá, vì tiền mua được hiển thị chứ không mua được sự hiểu đúng. Cũng không phải mức độ chuyển đổi số và đây là chỗ nhiều địa phương nhầm lẫn lớn nhất. Một tỉnh có thể sở hữu hàng trăm trang thông tin, hàng nghìn tập tin văn bản, hàng chục cơ sở dữ liệu, mà toàn bộ khối ấy vẫn tồn tại như những hòn đảo rời rạc. Với con người, các mảnh ghép rời vẫn đủ để hiểu một nơi. Với hệ thống suy luận, chúng là hàng triệu mảnh giấy chưa từng được đóng thành sách.

Cuốn sách này gọi phẩm chất ấy là "tính khả đọc thuật toán" và mượn nửa đầu của cái tên từ một truyền thống học thuật đã nghiên cứu việc đọc từ rất lâu trước khi máy biết đọc. Việc đặt tên tuy chưa giải quyết được vấn đề, nhưng nếu một vấn đề không được gọi tên thì sẽ không bao giờ được cấp ngân sách.

## 2.2. PHẢ HỆ CỦA TÍNH KHẢ ĐỌC: SCOTT VÀ BỘ MÁY BIẾT ĐỌC

### 2.2.1. Cái nhìn của nhà nước

Năm 1998, nhà nhân học chính trị James C. Scott xuất bản một cuốn sách về những dự án cải tạo xã hội không thành công của thế kỷ XX. Luận điểm trung tâm xoay quanh một từ, "legibility", tính khả đọc.[1] Nhà nước hiện đại, Scott lập luận, chỉ quản trị được những gì nó đọc được. Vì thế, công việc đầu tiên của mọi nhà nước là làm cho xã hội trở nên khả đọc. Định danh họ cho người dân để phục vụ thu thuế theo hộ. Lập bản đồ địa bạ để xác định quyền sở hữu. Quy hoạch đường phố theo ô bàn cờ để dễ dàng trong việc tuần tra giám sát. Chuẩn hóa cân đong đo đếm để thương mại kiểm soát được. Xã hội tự nhiên vốn phức tạp và đa dạng như một khu rừng già trong khi bộ máy quản lý nhà nước lại cần một khu rừng được quy hoạch trồng thẳng lối. Cái mới trong cuốn sách của Scott không phải phát hiện rằng nhà nước đơn giản hóa mà chính là những phân tích về cái giá phải trả.

Rừng trồng của lâm học Phổ: Cuối thế kỷ XVIII, các nhà lâm học Đức và Phổ xây dựng một ngành khoa học mới nhằm tối đa hóa sản lượng gỗ trên mỗi đơn vị diện tích. Họ chặt sạch rừng hỗn giao, trồng lại một loài duy nhất, xếp thẳng hàng, cùng tuổi, cùng chiều cao. Kết quả của thế hệ đầu tiên rực rỡ tới mức mô hình này được xuất khẩu khắp thế giới. Thế hệ thứ hai chết hàng loạt. Sản lượng sụt tới mức người Đức phải đặt ra một từ mới, Waldsterben, nghĩa là rừng chết. Nguyên nhân không nằm

1. James C. Scott, Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed, Yale University Press, New Haven, 1998.

ở loài cây được chọn. Nó nằm ở tất cả những gì bản đồ đã bỏ đi để dễ đọc. Hệ nấm cộng sinh trong đất. Tầng mùn hình thành từ lá của nhiều loài rụng vào nhiều mùa. Côn trùng ăn sâu bệnh. Chim làm tổ trong cây chết. Không thứ nào trong số đó xuất hiện trong sổ đo sản lượng gỗ, nên không thứ nào được giữ lại và không ai nhận ra chúng quan trọng cho tới khi chúng biến mất.

Scott gọi phần tri thức không lên được bản đồ là metis, tri thức thực hành của người tại chỗ và kết luận rằng các đại dự án thất bại không vì thiếu dữ liệu mà vì xem nhẹ loại tri thức không thể thành dữ liệu. Bản đồ nào cũng bỏ đi phần lớn thực tại để giữ lại phần cần cho mục đích quản trị. Thảm họa bắt đầu khi ai đó tin bản đồ hơn tin lãnh thổ, rồi cưỡng chế lãnh thổ theo bản đồ.

### 2.2.2. Vì sao mọi bộ máy đều cần một thế giới khả đọc

Nhu cầu khả đọc hóa là điều kiện kỹ thuật của bốn chức năng cơ bản và bốn chức năng ấy cho thấy vì sao logic này lặp lại ở mọi bộ máy.

**Thuế: không thu được từ thứ không đếm được**

Muốn thu thuế, phải biết đối tượng thu thuế là ai và thu thuế gì. Điều đó nghe hiển nhiên tới mức người ta quên rằng trong phần lớn lịch sử loài người, nó bất khả thi. Một nông dân không có họ cố định, sống trên mảnh đất không có ranh giới ghi nhận, thu hoạch bằng đơn vị đo địa phương khác với đơn vị đo của làng bên, là một người không thể thu thuế được. Việc phát minh ra họ, địa bạ, hệ đo lường thống nhất không chỉ là tiến bộ của văn minh nói chung mà còn là hạ tầng của việc thu thuế.

**Quản trị: không điều phối được thứ không xác định được**

Muốn phân bổ nguồn lực, phải biết có bao nhiêu đơn vị nhận và mỗi đơn vị ở đâu. Muốn quy hoạch, phải biết ranh giới. Muốn cấp phép, phải biết chủ thể. Mọi hành vi quản trị đều giả định trước một danh mục các thực thể đã được xác định và việc lập danh mục ấy luôn diễn ra trước, thường lặng lẽ và luôn kèm theo quyết định về việc cái gì được xếp vào ô nào.

**Quân sự: không huy động được thứ không truy tìm được**

Chế độ nghĩa vụ quân sự hiện đại không khả thi nếu không biết ai bao nhiêu tuổi và đang ở đâu. Sổ hộ tịch, căn cước và hệ thống đăng ký cư trú đều mang trong mình dấu vết của nhu cầu này. Đây cũng là ví dụ rõ nhất cho thấy khả đọc chưa bao giờ trung tính, cùng một cuốn sổ cho phép nhà nước gửi trợ cấp cũng cho phép nhà nước gửi giấy gọi nhập ngũ.

**Thống kê: không điều hành được thứ chưa thành số**

Từ "statistics" (thống kê) vốn có nghĩa là khoa học về nhà nước. Không phải ngẫu nhiên. Muốn biết nền kinh tế đang lớn hay co, phải định nghĩa cái gì được tính vào sản lượng. Việc lao động nội trợ không được tính vào tổng sản phẩm quốc nội không phải một sơ suất kỹ thuật mà là một quyết định phân loại và nó có hậu quả kéo dài hàng thế kỷ đối với cách xã hội đánh giá công việc của phụ nữ.

Bốn nhu cầu, một logic chung. Bộ máy chỉ hành động được lên những gì nó phân giải được thành đơn vị và mọi phép phân giải đều đòi thế giới tự trình bày theo lược đồ của bộ máy chứ không theo hình dạng vốn có của mình.

### 2.2.3. Khi máy trở thành bộ máy đọc mới

Khung của Scott chuyển sang thời suy luận vừa khớp một cách kỳ lạ vừa lệch ở một khớp then chốt. Khớp ở chỗ hệ thống suy luận cũng chỉ hành động được trên những gì nó đọc được, cũng đòi thế giới tự trình bày theo lược đồ, cũng tạo ra một phiên bản đơn giản hóa của thực tại rồi vận hành trên phiên bản ấy như thể nó là thực tại.

Lệch ở chỗ chủ thể đọc không còn là nhà nước. Khác biệt ấy sâu hơn vẻ ngoài của nó. Nhà nước của Scott đọc để cai trị và chính vì cai trị nên nó có chủ quyền trên lãnh thổ nó đọc, có nghĩa vụ với dân cư trên lãnh thổ ấy, có bộ máy để tiếp nhận khiếu nại và về nguyên tắc có thể bị thay thế bằng bầu cử. Quan hệ giữa kẻ đọc và kẻ bị đọc là một quan hệ chính trị, với đầy đủ những ràng buộc mà từ ấy hàm ý.

Bộ máy biết đọc hôm nay là một nhúm hệ thống tư nhân xuyên biên giới. Chúng đọc để phục vụ truy vấn của người dùng. Chúng không có chủ quyền, không có nghĩa vụ giải trình với vùng đất bị đọc, không có cả ý định cai trị. Địa phương rơi vào một thế lạ trong lịch sử, bị đọc bởi kẻ mình không bầu ra, theo lược đồ mình không góp ý, với hậu quả kinh tế mình phải gánh và không có bàn nào để ngồi vào mà thương lượng.

Bức tranh này bỏ sót một chi tiết làm nó bớt u ám hơn một chút. Không giống nhà nước, các hệ thống suy luận cạnh tranh với nhau và cạnh tranh tạo ra chỗ hở. Một địa phương bị một hệ thống mô tả sai vẫn có thể được hệ thống khác mô tả đúng và vì người dùng ngày càng dùng nhiều hệ thống song song, sự đa dạng ấy có giá trị bảo hiểm thật. Nhưng cần thận trọng với niềm an ủi này. Các hệ thống cạnh tranh nhau ở tầng giao diện trong khi vẫn uống chung

một nguồn nước ở tầng dữ liệu và như phần sau sẽ cho thấy, nguồn nước ấy tập trung tới mức đáng ngại. Cạnh tranh ở nơi không quan trọng không bù được sự tập trung ở nơi quan trọng.

### 2.2.4. Từ nhà nước đến thuật toán

Đặt ba giai đoạn của tính khả đọc cạnh nhau sẽ thấy một quỹ đạo chuyển dịch quyền lực mà không ai chủ ý thiết kế nhưng gần như mọi xã hội số đều đang đi qua. *Ở giai đoạn thứ nhất*, nhà nước xây dựng tính khả đọc để quản trị. Đối tượng bị đọc là dân cư và tài sản trên lãnh thổ, chế tài của việc không khả đọc là xử phạt hoặc cưỡng chế, người bị quản lý biết mình đang bị ghi nhận và có các cơ chế khiếu nại như tòa án, cơ quan hành chính hay bầu cử. Quan hệ quyền lực ở đây là quan hệ chính trị giữa nhà nước và công dân. *Sang giai đoạn thứ hai*, các nền tảng số xây dựng tính khả đọc để phân phối. Thứ được đọc không còn là dân cư mà là nội dung, người dùng và các tín hiệu hành vi. Hình phạt cũng thay đổi, thay vì bị xử phạt hành chính, một thực thể có thể bị giảm hiển thị, giảm tiếp cận hoặc biến mất khỏi các dòng đề xuất. Người dùng chỉ có thể suy đoán điều đó thông qua lượt xem, lượt tương tác hoặc xếp hạng, trong khi cơ chế khiếu nại chủ yếu bị giới hạn trong điều khoản dịch vụ. Quan hệ quyền lực chuyển từ chính trị sang quan hệ hợp đồng giữa nền tảng và người sử dụng. *Đến giai đoạn thứ ba*, khi tác nhân AI trở thành chủ thể thực hiện hành động, tính khả đọc không còn nhằm quản trị hay phân phối mà nhằm ủy nhiệm. Đối tượng bị đọc mở rộng thành mọi thực thể có thể được suy luận và giao dịch: địa điểm, doanh nghiệp, sản phẩm, dịch vụ hay sự kiện. Hình phạt của việc không khả đọc cũng khắc nghiệt hơn, không phải bị cấm hay bị giảm hiển thị, mà đơn giản là không bao giờ xuất hiện trong tập phương án mà AI cân nhắc. Đáng chú ý hơn, phần

lớn các thực thể thậm chí không biết mình đã bị loại khỏi quá trình lựa chọn, cũng chưa có một cơ chế khiếu nại chính thức để yêu cầu AI sửa tri thức hoặc quyết định của mình. Quyền lực vì thế trở nên vô danh hơn, không còn mang hình thức quyền lực chính trị của nhà nước hay quyền lực hợp đồng của nền tảng, mà là quyền lực được thực thi thông qua các mô hình, dữ liệu và thuật toán, nhiều khi không có một chủ thể cụ thể nào có thể chỉ ra và quy trách nhiệm.

Quan hệ giữa một địa phương với hệ thống suy luận đang mô tả nó chưa có tên trong bất kỳ nhánh nào của luật học. Nó không phải quan hệ hợp đồng, vì không bên nào ký gì. Không phải quan hệ hành chính, vì không có thẩm quyền công. Không phải quan hệ dân sự thông thường, vì thiệt hại phát sinh không đo được và bên gây thiệt hại không hề có ý định gây ra nó.

Ai nắm lược đồ, kẻ đó nắm quyền phát biểu chân lý về một vùng đất. Mệnh đề này sẽ trở lại ở mục 2.6 dưới cái tên chủ quyền thuật toán. Ở đây chỉ cần ghi nhận rằng khả đọc chưa bao giờ là một khái niệm kỹ thuật trung tính. Nó là khái niệm về quyền lực, ngay từ trong phả hệ của nó.

### 2.2.5. Hai đại dự án và một bài học

Rừng Phổ là ví dụ cổ điển của Scott nhưng cũng là ví dụ dễ chấp nhận nhất, vì cây cối không phản đối. Hai trường hợp dưới đây khó chịu hơn và chúng cần thiết vì chúng cho thấy điều gì xảy ra khi đối tượng bị khả đọc hóa là con người.

Một thành phố được thiết kế để nhìn từ trên cao. Brasília được xây từ con số không trong khoảng bốn năm và khánh thành năm

1960, theo một bản quy hoạch đoạt giải thi tuyển. Nhìn từ máy bay, nó hoàn hảo. Các khu chức năng tách bạch, khu ở riêng, khu làm việc riêng, khu thương mại riêng, các trục giao thông rộng rãi, không có sự lộn xộn của những thành phố mọc lên qua nhiều thế kỷ. Nhìn từ mặt đất, một điều bất ngờ xảy ra. Người dân bắt đầu than phiền về một cảm giác mà họ gọi bằng một từ mới, "brasilite", tạm dịch là chứng Brasília, mô tả sự trống rỗng của việc sống trong một nơi không có góc phố để dừng lại, không có quán xá mọc tự phát dưới nhà, không có cái mà Jane Jacobs từng gọi là vũ điệu vỉa hè.[1]

Đáng chú ý nhất là phản ứng của thực tại, chứ không phải bản thân các khiếm khuyết ấy. Xung quanh thành phố quy hoạch, các khu định cư tự phát mọc lên, lộn xộn, không có trong bản vẽ và đông dân hơn phần được quy hoạch. Cuộc sống không bị xóa. Nó chỉ dời sang chỗ không được vẽ. Với một địa phương đang chuẩn hóa dữ liệu về mình, đây là bài học có giá trị trực tiếp. Phần không lọt vào lược đồ không biến mất. Nó tiếp tục tồn tại, chỉ là ở ngoài tầm nhìn của hệ thống và theo thời gian, khoảng cách giữa cái được ghi nhận với cái đang thực sự diễn ra ngày càng rộng, cho tới lúc bản mô tả chính thức không còn mô tả nơi nào có thật.

Trường hợp thứ hai nghiêm trọng hơn nhiều và Scott dành cho nó những trang cay đắng nhất trong sách. Khi khả đọc hóa được thực hiện với thiện chí. Đầu thập niên 1970, chính phủ Tanzania triển khai một chương trình quy tụ dân cư nông thôn vốn sống rải rác vào các làng tập trung, với mục tiêu chân thành

1. Jane Jacobs, The Death and Life of Great American Cities, Random House, New York, 1961. Hình ảnh vũ điệu vỉa hè được trình bày ở chương 2. Trường hợp Brasília do James C. Scott phân tích trong Seeing Like a State, chương 4, dựa trên khảo sát dân tộc học của James Holston, The Modernist City: An Anthropological Critique of Brasília, University of Chicago Press, Chicago, 1989.

và tiến bộ là đưa trường học, trạm y tế, nước sạch và khuyến nông tới người dân.[1] Không thể xây trường cho những hộ sống cách nhau nhiều cây số. Muốn phục vụ, phải tập trung. Muốn tập trung, phải biết ai ở đâu. Tuy nhiên, sau đó, chương trình thất bại. Sản lượng nông nghiệp sụt. Đất quanh các làng mới bị khai thác quá mức trong khi đất cũ bỏ hoang. Nhiều gia đình chuyển tới nơi mà tri thức canh tác tích lũy qua nhiều thế hệ về loại đất, hướng gió, mạch nước ngầm và lịch mưa của vùng cũ trở nên vô dụng. Điều cần rút ra ở đây là không phải sai lầm về chính sách của chính phủ Tanzania. Phần lớn những người thực hiện chương trình này tin rằng họ đang làm điều tốt và ở nhiều phương diện, họ đúng. Điều cần rút ra là thiện chí không bảo vệ được ai khỏi hậu quả của việc đơn giản hóa. Bản đồ bỏ đi tri thức tại chỗ vì bản đồ không có ô cho nó và người vẽ bản đồ không hề biết mình đang bỏ đi cái gì.

Mọi thảm họa của khả đọc hóa đều bắt đầu bằng một người thành thật muốn giúp. Ba trường hợp, ba lĩnh vực, một cấu trúc. Bộ máy cần một thế giới đơn giản hơn thế giới thật để hành động được. Nó tạo ra bản mô tả ấy. Rồi nó bắt đầu tin bản mô tả. Rồi nó cưỡng chế thực tại theo bản mô tả. Khâu thứ ba là khâu chết người và điều đáng lo là nó không cần ai ra quyết định, chỉ cần bản mô tả tồn tại đủ lâu và không có ai được giao nhiệm vụ nghi ngờ nó.

### 2.2.6. Một khác biệt quan trọng của thời hiện tại

Đến đây cần dừng lại để nêu một điểm mà nếu bỏ qua sẽ khiến cả chương trở thành một bài luận bi quan thay vì một công cụ làm

1. Chương trình làng tập trung ujamaa của Tanzania, giai đoạn 1973 tới 1976. Trường hợp được phân tích trong James C. Scott, Seeing Like a State, Yale University Press, New Haven, 1998, chương 7.

việc. Trong cả ba trường hợp trên, bộ máy đọc có quyền cưỡng chế. Nhà nước Phổ có thể ra lệnh chặt rừng. Chính phủ Brazil có thể xây một thủ đô. Chính phủ Tanzania có thể di dời hàng triệu người. Bản đồ trở thành hiện thực vì có quyền lực thực thi đứng sau nó. Hệ thống suy luận không có quyền cưỡng chế. Nó không thể di dời ai, không thể chặt cây, không thể ra lệnh cho một tỉnh phải trở thành cái gì. Nó chỉ có thể không nhắc tới. Câu hỏi trung tâm của chương này và có lẽ của cả cuốn sách, là liệu sự khác biệt ấy có ý nghĩa hay không.

Lập luận rằng thiếu quyền cưỡng chế thì bớt nguy hiểm có sức thuyết phục ban đầu và yếu dần khi xét kỹ. Cưỡng chế kinh tế không cần mệnh lệnh. Nếu một địa phương mất dần dòng khách vì không xuất hiện trong tập lựa chọn, nó sẽ tự thay đổi để xuất hiện và nó sẽ thay đổi theo hướng mà nó đoán là hệ thống ưa thích. Không ai ra lệnh cho bà cụ trong hẻm phải mở cửa đúng giờ. Bà tự làm, vì điểm đánh giá. Kết quả cuối cùng có thể giống hệt kết quả của một mệnh lệnh, chỉ khác ở chỗ không có ai để quy trách nhiệm và không có ai để phản đối. Ở phương diện đó, hình thức quyền lực mới khó chống lại hơn hình thức cũ chứ không dễ hơn.

Tuy vậy, sự khác biệt vẫn có ý nghĩa ở một điểm quan trọng và không nên bỏ qua nó vì muốn lập luận gọn. Vì không có cưỡng chế, phần không khả đọc vẫn tiếp tục tồn tại trên thực địa, giống các khu định cư tự phát quanh Brasília. Một làng nghề không xuất hiện trong bất kỳ hệ thống nào vẫn còn đó, vẫn làm nghề, vẫn truyền nghề. Nó nghèo hơn, nhưng nó chưa chết. So với rừng Phổ, nơi hệ nấm bị xóa sạch không thể phục hồi, đây là một khác biệt về khả năng cứu vãn và nó là cơ sở duy nhất cho phép cuốn sách này kết thúc bằng các khuyến nghị thay vì bằng một lời điếu văn.

Điều này có nghĩa gì đối với một địa phương? Bộ máy đọc mới không có quyền cưỡng chế, nên phần không khả đọc vẫn tồn tại trên thực địa và vẫn còn cơ hội được cứu. Nhưng sức ép kinh tế thì vẫn tới, chỉ tới chậm hơn và không kèm một văn bản nào để phản đối. Việc cần làm không phải chờ một chỉ thị mà là tự lập danh sách những tài sản đang tồn tại ngoài mọi hệ thống dữ liệu, xếp theo mức đóng góp cho sinh kế, rồi quyết định đưa phần nào vào lược đồ trước. Danh sách ấy rẻ để lập và hiện gần như không địa phương nào có.

## 2.3. PHÂN LOẠI VÀ LÃNG QUÊN: BOWKER, STAR, FOUCAULT

### 2.3.1. Không có phân loại nào là trung tính

Nếu Scott cho khái niệm này cột sống chính trị thì một truyền thống thứ hai cho nó hệ thần kinh. Năm 1999, Geoffrey Bowker và Susan Leigh Star khảo sát các hệ phân loại lớn, từ danh mục bệnh quốc tế tới bảng phân chủng tộc của chế độ Apartheid.[1] Kết luận của họ nghe hiền lành mà sắc sảo. Mọi hệ phân loại đều đồng thời là một cỗ máy ghi nhớ và một cỗ máy lãng quên. Cái gì có ô để điền thì được hệ thống nhìn thấy, được thống kê, được cấp ngân sách. Cái gì không có ô thì bị dồn vào mục khác và mục khác là nghĩa trang của những thứ có thật nhưng không được thiết kế chỗ đứng. Hệ phân loại tốt thì vô hình. Nó chỉ hiện hình ở chỗ gãy, nơi một hiện tượng thật bị nhét vào ô sai và người bị nhét là người duy nhất cảm thấy có gì đó không ổn.

Khi một cái ô quyết định một cuộc đời. Chế độ Apartheid ở Nam Phi vận hành trên một hệ phân loại chủng tộc bốn nhóm. Toàn bộ quyền lợi của một con người, quyền cư trú, quyền học hành, quyền kết hôn, quyền làm nghề, đều phái sinh từ ô mà người ấy được xếp vào. Vấn đề là thực tại không có bốn nhóm. Nó có hàng triệu người với tổ tiên trộn lẫn qua nhiều thế hệ. Nên bộ máy phải phát minh ra các thủ tục để quyết định trường hợp ranh giới, trong đó có những thủ tục đã đi vào lịch sử như biểu tượng của sự phi lý, chẳng hạn việc thử xem một chiếc lược có mắc lại trong tóc hay không.

1. Geoffrey C. Bowker và Susan Leigh Star, Sorting Things Out: Classification and Its Consequences, MIT Press, Cambridge, Massachusetts, 1999.

Bowker và Star nhấn mạnh rằng cái đáng sợ không phải sự tàn bạo của thủ tục. Đó là tính thường nhật của nó. Hàng nghìn viên chức làm việc ấy như một thao tác hành chính, điền vào biểu mẫu, đóng dấu, chuyển hồ sơ. Bạo lực của hệ phân loại không nằm trong ý định của người vận hành. Nó nằm trong cấu trúc của cái bảng.

Áp vào địa phương, các lược đồ mà máy dùng để đọc thế giới đều là những hệ phân loại như thế, chỉ khác về mức độ hậu quả. Một quần thể tín ngưỡng đang sống có thể bị xếp vào ô điểm tham quan lịch sử, đúng vừa đủ để không ai buồn sửa và sai vừa đủ để mọi câu mô tả sinh ra từ đó lệch giọng. Nghi lễ không có ô riêng thì thành lễ hội. Nghề thủ công không có ô riêng thì thành quà lưu niệm. Một không gian thiêng không có ô riêng thì thành điểm chụp ảnh. Mỗi lần nhét ô như vậy, một phần bản sắc bị dịch sang ngôn ngữ của kẻ khác. Và bản dịch, như mọi bản dịch, có phần mất đi không đòi lại được.

### 2.3.2. Hai cơ chế lãng quên

Bowker và Star tách sự lãng quên của hệ phân loại thành hai cơ chế khác nhau và sự phân biệt này có giá trị thực hành trực tiếp cho các địa phương. *Cơ chế thứ nhất là dọn sạch*. Một hệ phân loại mới được áp lên và những gì không khớp bị loại bỏ có chủ ý, thường nhân danh tính nhất quán. Đây là loại lãng quên nhìn thấy được, tranh cãi được và về nguyên tắc có thể chống lại, vì có một quyết định cụ thể do một người cụ thể đưa ra vào một thời điểm cụ thể. *Cơ chế thứ hai là xóa dần*. Không ai quyết định gì. Chỉ là những hạng mục không có ô thì không được ghi lại, không được ghi lại thì không xuất hiện trong dữ liệu, không xuất hiện trong dữ liệu thì thế hệ sau không biết chúng từng tồn tại. Sau

ba mươi năm, một truyền thống có thể biến mất khỏi ký ức tập thể mà không ai từng ra lệnh cấm nó. Cơ chế thứ hai nguy hiểm hơn nhiều, vì nó không có thủ phạm để phản đối. Không có ai đốt sách. Chỉ là không có kệ nào để xếp chúng lên.

Sẽ là bi thảm hóa nếu coi mọi phép phân loại là một mất mát. Phân loại cũng là điều kiện của công bằng. Danh mục di sản của UNESCO, dù bị phê phán chính đáng vì thiên lệch về kiến trúc đá và về châu Âu, vẫn đã cứu hàng trăm công trình khỏi bị phá và đã cho nhiều cộng đồng ở các nước nghèo một công cụ pháp lý để chống lại chính quyền của họ. Có ô còn hơn không có ô, trong phần lớn tình huống thực tế. Lập luận đúng không phải là chống phân loại mà là đòi quyền tham gia vào việc thiết kế các ô và đó là hai chương trình hành động rất khác nhau.

### 2.3.3. Quyền lực tạo ra tri thức

Không cần theo Foucault đến tận cùng để giữ lại một mệnh đề làm việc, cái được nhìn thấy luôn phụ thuộc vào lược đồ phân loại có trước.[1] Thế giới không tự hiện ra như một tập hợp sự thật trung tính, nó chỉ trở nên khả kiến sau khi được đặt vào một hệ thống các phạm trù, quy tắc và cách phân biệt. Chính lược đồ ấy quyết định điều gì được nhận diện, điều gì bị bỏ qua và điều gì trở thành đối tượng của tri thức. Máy không tìm thấy địa phương rồi phân loại nó. Máy chỉ tìm thấy những gì lược đồ của nó cho phép tồn tại. Ở đây lý thuyết gặp một con số và con số ấy sắc hơn phần lớn lập luận triết học về chủ đề này.

1. Michel Foucault, Les mots et les choses: Une archéologie des sciences humaines, Gallimard, Paris, 1966. Bản tiếng Anh: The Order of Things: An Archaeology of the Human Sciences, Pantheon Books, New York, 1970.

Một phân tích một nghìn câu trả lời sinh ra cho thấy trong khoảng một nghìn hai trăm tên miền từng được dẫn lại ít nhất một lần, chỉ khoảng mười hai tên miền, tức 1% cao nhất, chiếm 47% toàn bộ số lần được dẫn. Nhóm 9% tiếp theo chiếm thêm 31%. Chín mươi phần trăm còn lại chia nhau 22%.[1] Để so sánh, trong kết quả tìm kiếm truyền thống, 1% tên miền cao nhất thường chiếm khoảng 25 tới 30% số vị trí đầu bảng.[2]

Trật tự tri thức mới tập trung gấp đôi trật tự mà nó thay thế. Và nó chưa từng được ai bỏ phiếu thông qua. Với một tỉnh, con số ấy có nghĩa rất cụ thể. Điều mà một hệ thống suy luận biết về tỉnh ấy phần lớn không đến từ cổng thông tin của tỉnh. Nó đến từ một nhúm nền tảng toàn cầu, phần nhiều là các trang đánh giá của người dùng, các bách khoa toàn thư mở và các diễn đàn viết bằng tiếng Anh. Cơ quan có thẩm quyền pháp lý để phát biểu về một vùng đất và nguồn có thẩm quyền thực tế trước máy là hai chủ thể khác nhau, ở hai châu lục khác nhau, không liên lạc với nhau.

1. Digital Applied, "1,000 AI Overviews Analyzed: Citation Pattern Study", ngày 26 tháng 4 năm 2026, https://www.digitalapplied.com/blog/we-analyzed-1000-ai-overviews-citation-pattern-study. Mẫu 1.000 bản tóm tắt sinh ra của Google, thu thập hai tuần tháng 4 năm 2026 trên tìm kiếm tiếng Anh tại Mỹ, 10 nhóm ý định và khoảng 30 lĩnh vực, cho khoảng 1.200 tên miền và 4.200 địa chỉ trang. Nghiên cứu ngành, phương pháp chỉ công bố tóm lược.

2. First Page Sage, "Google Click-Through Rates (CTRs) by Ranking Position", bản cập nhật năm 2026, https://firstpagesage.com/reports/google-click-through-rates-ctrs-by-ranking-position/, và SISTRIX, "Why (almost) everything you knew about Google CTR is no longer valid", ngày 14 tháng 7 năm 2020. Nghiên cứu SISTRIX có trước khi tóm tắt sinh ra xuất hiện nên chỉ dùng làm mốc lịch sử. Cả hai là số liệu ngành, chỉ đọc theo bậc độ lớn.

### 2.3.4. Khả đọc luôn là một quyết định chính trị

Từ ba nguồn trên rút ra kết luận của cả phần này và nó là kết luận mà phần còn lại của cuốn sách dựa vào. Việc quyết định đưa cái gì vào lược đồ không phải một tác vụ kỹ thuật có thể giao cho nhà thầu công nghệ. Nó là việc phân bổ quyền phát biểu chân lý về một vùng đất và vì thế phải được quyết định ở cấp có thẩm quyền chính trị, với thủ tục tham vấn cộng đồng bị ảnh hưởng. Nghe như một nguyên tắc đạo đức. Nó thực ra là một khuyến nghị về tổ chức bộ máy và nó có hệ quả rất cụ thể, hồ sơ dữ liệu di sản của một địa phương không nên do phòng công nghệ thông tin phê duyệt.

Nguyên tắc này có một cái giá mà những người ủng hộ nó thường không nói. Đưa việc thiết kế lược đồ lên cấp chính trị nghĩa là đưa nó vào một quy trình chậm, dễ bị chi phối bởi các nhóm lợi ích địa phương và thường kết thúc bằng những thỏa hiệp làm hài lòng các bên có tiếng nói nhất chứ không phải các bên bị ảnh hưởng nhất. Trong khi đó, các hệ thống suy luận không chờ chúng ta. Một lược đồ được cộng đồng tranh luận trong bốn năm có thể lỗi thời trước khi được thông qua. Đây là một sự đánh đổi thật giữa tính chính danh và tốc độ, không phải một vấn đề kỹ thuật có lời giải tối ưu và ở đây sẽ không giả vờ rằng có cách nào tránh được nó.

### 2.3.5. Ai đang giữ quyền mô tả một vùng đất

Lập luận về tập trung nguồn dẫn ở mục trên còn trừu tượng. Nó trở nên cụ thể một cách khó chịu khi đặt câu hỏi giản dị sau: "Nếu một người ở Berlin hỏi máy về một xã/phường ở Việt Nam, câu trả lời được lắp từ chất liệu của ai?".

Gần như chắc chắn không phải từ cổng thông tin điện tử của xã/phường ấy. Cũng gần như chắc chắn không phải từ một tài liệu tiếng Việt nào. Chất liệu sẽ đến từ vài nguồn quen thuộc, một mục bách khoa toàn thư mở do người không sống ở đó viết, vài bài blog du lịch của khách nước ngoài từng ghé qua trong ba ngày, một số bài đăng trên diễn đàn, các đánh giá trên nền tảng đặt phòng và có thể một bài báo tiếng Anh viết nhân dịp nào đó.

Không nguồn nào trong số ấy sai. Đó không phải vấn đề. Vấn đề là tổng của chúng tạo ra một chân dung có cấu trúc rất đặc thù, nó đầy đủ ở những gì khách quan tâm và trống rỗng ở những gì cư dân quan tâm. Nó biết rõ đường tới điểm chụp ảnh đẹp nhất và không biết gì về việc nghề gì đang nuôi sống mấy nghìn hộ ở đó. Nó ghi lại rằng món ăn địa phương ngon và không ghi lại rằng nguyên liệu chính đang khan hiếm vì một lý do mà cả vùng đều biết.

Một vùng đất được kể lại bằng ký ức của những người đi qua nó, chứ không phải của những người ở lại. Điều này không mới. Toàn bộ lịch sử của ngành du lịch là lịch sử của việc nơi chốn được kể bởi người ngoài. Cái mới là ba thứ.

*Thứ nhất*, quy mô. Trước đây một cuốn cẩm nang du lịch viết sai về một vùng đất thì có ảnh hưởng trong phạm vi số bản in. Nay một mô tả sai được nhân bản trong mọi câu trả lời liên quan, ở mọi ngôn ngữ, không giới hạn số bản.

*Thứ hai*, tính vô danh. Cuốn cẩm nang có tác giả, có nhà xuất bản, có địa chỉ để gửi thư phản đối. Câu trả lời sinh ra không có tác giả theo nghĩa thông thường và người bị mô tả sai không biết gửi thư cho ai.

*Thứ ba*, tính tự củng cố. Cuốn cẩm nang tái bản sau vài năm và có thể sửa. Mô tả sinh ra được đăng lại, trở thành nguồn cho lần sinh sau và mỗi vòng lặp làm nó khó sửa hơn.

Có một chiều ngược lại của hiện tượng này mà việc bỏ qua sẽ khiến bức tranh sai lệch. Trong nhiều trường hợp, chính việc bị người ngoài mô tả đã cứu một di sản khỏi sự thờ ơ của chính quyền sở tại. Không ít làng nghề, tuyến đường cổ và công trình bị bỏ hoang được biết tới rộng rãi nhờ khách nước ngoài viết về chúng trước và sự chú ý ấy sau đó tạo áp lực buộc địa phương phải bảo tồn. Quyền mô tả nằm trong tay người ngoài không tự động là điều xấu. Nó xấu khi nó là quyền duy nhất, nghĩa là khi tiếng nói của người ở lại không có kênh nào để vào cùng bản ghi.

### 2.3.6. Một lược đồ không có ô cho cái gì

Để cụ thể hơn nữa, hãy thử liệt kê những gì các lược đồ phổ biến hiện nay không có chỗ chứa khi mô tả một địa phương của Việt Nam.

- ***Quan hệ dòng họ và cấu trúc làng xã,*** thứ quy định phần lớn cách một cộng đồng ra quyết định về tài sản chung của mình.

- ***Lịch âm và các mốc thời gian không cố định theo dương lịch,*** khiến mọi lễ hội theo lịch âm hoặc bị ép về một ngày sai hoặc bị bỏ trống trường thời gian.

- ***Tri thức về mùa vụ, con nước và thời tiết địa phương,*** loại tri thức mà Scott gọi là metis (tri thức thực hành, tri thức địa phương, kinh nghiệm tích lũy, kỹ năng ngầm và khả

năng ứng biến gắn với hoàn cảnh cụ thể, những thứ rất khó chuẩn hóa thành quy tắc hay dữ liệu) và là thứ quyết định thời điểm tốt nhất để đến một nơi.

- ***Các mức độ thiêng khác nhau của cùng một không gian,*** khiến một ngôi đình vừa là điểm tham quan vừa là nơi mà một số khu vực không nên đặt chân vào và lược đồ chỉ ghi nhận vế thứ nhất.

- ***Quyền sở hữu tập thể đối với tri thức,*** chẳng hạn một bài thuốc hay một kỹ thuật thủ công thuộc về cả làng chứ không thuộc về ai, một quan hệ mà hầu hết hệ thống dữ liệu không biểu diễn được vì chúng giả định mọi thứ đều có một chủ sở hữu xác định.

Danh sách này không nhằm chứng minh rằng lược đồ hiện hành không hiệu quả mà nó nhằm chỉ ra rằng những khoảng trống ấy không ngẫu nhiên. Chúng phản ánh các giả định của nơi mà lược đồ được thiết kế và các giả định ấy gồm sở hữu cá nhân, thời gian tuyến tính theo dương lịch, không gian đồng nhất và tri thức có tác giả. Đây chính là điểm mà Bowker và Star muốn nói khi khẳng định không có phân loại nào trung tính. Không ai cố tình loại bỏ lịch âm khỏi các hệ thống dữ liệu sự kiện. Người thiết kế chỉ đơn giản không nghĩ tới nó, vì trong thế giới của họ, thời gian đã luôn là dương lịch.

Từ nhận xét này rất dễ trượt sang một kết luận sai lầm và tốn kém rằng, chúng ta có cần phải tự xây dựng một lược đồ dữ liệu riêng cho mình. Cám dỗ ấy có thật và đã có những nỗ lực theo hướng đó ở nhiều nước. Vấn đề là một lược đồ mà chỉ mình ta sử dụng sẽ không giải quyết được vấn đề gì, vì giá trị của lược đồ nằm ở chỗ nó được các hệ thống bên ngoài đọc hiểu. Con đường khả thi

hơn, tuy chậm nhưng hiệu quả, là mở rộng các lược đồ hiện hành từ bên trong bằng cách tham gia vào các tổ chức xây dựng chuẩn quốc tế để đề xuất bổ sung trường dữ liệu đặc thù. Đây là công việc thầm lặng, không thể cắt băng khánh thành rầm rộ; nó đòi hỏi sự kiên trì trong nhiều năm, đổi lại chi phí ít tốn kém hơn tổ chức một cuộc hội thảo cấp vùng.

### ĐIỀU NÀY CÓ Ý NGHĨA GÌ ĐỐI VỚI MỘT ĐỊA PHƯƠNG?

*Những gì bản sắc của một nơi thường là những gì không có sẵn để khai báo, và khoảng trống ấy không hề ngẫu nhiên. Nó phản ánh giả định của nơi đã thiết kế lược đồ. Phản ứng đúng không nằm ở việc dựng một lược đồ riêng, vì một lược đồ chỉ mình ta đọc được thì không giải quyết được gì. Phản ứng đúng là lập bảng ánh xạ cho các loại hình bản địa, ghi rõ phần thiếu sót hoặc sai lệch khi ánh xạ, rồi cử người tham gia các nhóm xây dựng chuẩn quốc tế. Đây là công việc của những người kiên trì, không phải của một chiến dịch.*

## 2.4. MÔ HÌNH BỐN BỘ LỌC, MỘT VÒNG, MỘT VAN

### 2.4.1. Định nghĩa

Để khái niệm có ích cho công việc quản trị, nó cần được phát biểu ở dạng kiểm tra được. Tính khả đọc thuật toán (algorithmic legibility) là mức độ mà tài sản, năng lực, tri thức và lời hứa của một địa phương có thể được máy phân giải, xác minh, xếp hạng và hành động lên, mà không cần một trung gian người diễn giải hộ.

Bốn động từ trong định nghĩa không được chọn theo sở thích. Chúng là bốn bước bắt buộc của bất kỳ quy trình lựa chọn được ủy nhiệm nào và chúng xếp theo quan hệ điều kiện cần nghiêm ngặt. Không phân giải được thì không có gì để xác minh. Không xác minh được thì không có căn cứ để xếp hạng. Không xếp hạng được thì không có đối tượng để thực thi. Đó là lý do mô hình có đúng bốn bộ lọc chứ không phải ba hay bảy.

Mệnh đề này bác bỏ được và cuốn sách nêu rõ cách bác bỏ nó. Chỉ cần chỉ ra một trường hợp mà một địa phương giao dịch được với tác nhân số trong khi chưa phân giải được như một thực thể, mô hình sập.

Gọi chúng là bộ lọc, không phải bậc thang, vì chức năng của chúng là loại bỏ. Mỗi tầng giữ lại một phần và để phần còn lại chảy qua và địa phương nào bị chặn ở tầng nào thì đơn giản là không xuất hiện ở tầng dưới. Bậc thang gợi ý rằng leo chậm vẫn tới nơi. Bộ lọc thì không có khái niệm chậm. Chỉ có qua hoặc không.

Một điều cần nói rõ về sơ đồ, vì nếu không nói thì phản bác đầu tiên từ phía kỹ thuật sẽ đúng. Bốn bộ lọc là mô hình chức

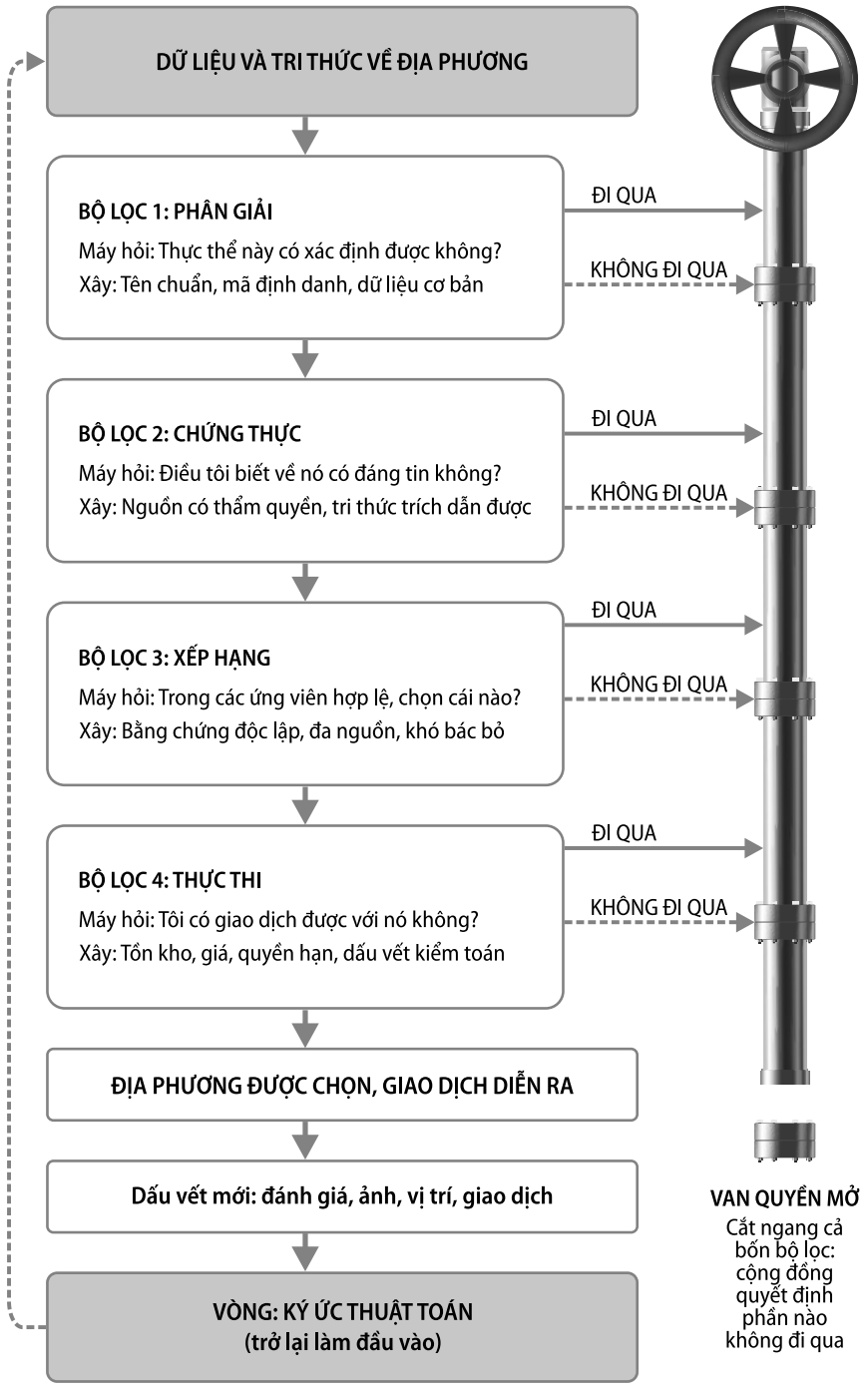
năng của một chuỗi quyết định, không phải sơ đồ của một kiến trúc kỹ thuật duy nhất. Trong thực tế, mỗi tầng do những thành phần khác nhau đảm nhiệm, bộ thu thập và chỉ mục, đồ thị tri thức, hệ truy hồi, mô hình xếp hạng, mô hình ngôn ngữ, hệ gợi ý, nền tảng đặt dịch vụ, cổng thanh toán và lớp điều phối tác nhân. Không có một cỗ máy nào thực hiện cả bốn.

Vì vậy, những phản bác dạng mô hình ngôn ngữ không có ký ức ổn định, mô hình không trực tiếp xếp hạng hay tác nhân không tự giữ dữ liệu nguồn đều đúng ở tầng kỹ thuật mà không bác bỏ mô hình ở tầng chức năng. Điều mô hình khẳng định là bốn thao tác ấy phải xảy ra ở đâu đó trong chuỗi trước khi một địa phương được chọn, không phải chúng xảy ra bên trong cùng một hệ thống.

Và một nguyên tắc phải nói ngay ở đây để tránh cách đọc sai nguy hiểm nhất của cả cuốn sách, khả đọc không thay thế chất lượng thật. Nó chỉ làm cho chất lượng thật có cơ hội được nhận diện và giao dịch. Một địa phương chuẩn hóa hoàn hảo mà sản phẩm kém sẽ chỉ bị loại nhanh hơn và chính xác hơn.

### 2.4.2. Bộ lọc thứ nhất, phân giải

Câu hỏi của tầng này là câu hỏi đơn giản nhất và cũng là câu hỏi nhiều địa phương trượt nhất, thực thể này có xác định được không. Ba việc phải làm ở đây. Thứ nhất là định danh, tức là một cái tên chuẩn cùng một mã ổn định mà mọi nguồn khác quy chiếu về. Thứ hai là bản thể luận, tức là trả lời câu hỏi thực thể này thuộc loại gì và có quan hệ ra sao với các thực thể khác, một xã thuộc tỉnh nào, một di tích nằm trong quần thể nào, một lễ hội gắn với địa điểm nào. Thứ ba là lược đồ, tức là bộ thuộc tính mô tả nó, từ tọa độ tới giờ mở cửa tới ngôn ngữ phục vụ.




**Hình 2.1.** Bốn bộ lọc, một vòng, một van

Bài toán của một cái tên. Một địa danh ở Việt Nam có thể tồn tại trước máy dưới nhiều dạng cùng lúc. Có dấu và không dấu. Viết liền và viết rời. Tên hành chính và tên dân gian. Tên cũ trước sáp nhập và tên mới sau sáp nhập. Phiên âm theo kiểu Pháp còn sót lại trong tài liệu cũ và phiên âm theo kiểu Anh trong các bài viết du lịch gần đây. Với con người, tất cả các dạng ấy đều rõ. Không ai nhầm. Với máy, mỗi dạng có thể là một thực thể riêng, mỗi thực thể có một phần dữ liệu, không thực thể nào có đủ dữ liệu để vượt ngưỡng tin cậy.

Đợt sắp xếp lại đơn vị hành chính ở Việt Nam gần đây làm bài toán này khó lên một bậc và cấp bách lên vài bậc. Hàng loạt địa danh đổi tên, hàng loạt cấp trung gian biến mất, trong khi khối lượng khổng lồ tài liệu cũ trên internet vẫn dùng tên cũ. Một hệ thống suy luận đứng trước tình huống ấy phải tự quyết định xem tên mới và tên cũ có phải cùng một thực thể hay không và nó quyết định dựa trên bằng chứng nó tìm được, chứ không dựa trên các quyết định.

Đây là một cửa sổ chính sách hiếm và nó đang khép. Một bảng đối chiếu chính thức, đầy đủ, có thẩm quyền, công bố ở dạng máy đọc được, ánh xạ từng tên cũ sang tên mới cùng ngày hiệu lực, là loại tài liệu ít tốn kém nhất và có giá trị lâu dài hơn phần lớn chiến dịch quảng bá cấp tỉnh. Điểm cần nhớ về tầng này là tính chất nhị phân của nó. Ở ba tầng sau, làm tốt hơn thì được lợi hơn theo mức độ. Ở tầng này, hoặc máy phân giải được thực thể hoặc không và nếu không thì mọi đầu tư ở các tầng sau đều rơi vào hư không.

### 2.4.3. Bộ lọc thứ hai, chứng thực

Đã xác định được thực thể, câu hỏi tiếp theo là những gì tôi biết về nó có đáng tin không. Máy trả lời câu hỏi này chủ yếu bằng ba tín hiệu. Thẩm quyền của nguồn, tức là nguồn này có phải nơi lẽ ra phải biết điều đó hay không. Khả năng trích dẫn, tức là điều được nói có thể lần ngược về một nguồn cụ thể hay chỉ trôi nổi. Và sự đối chiếu, tức là nhiều nguồn độc lập có nói giống nhau hay không. Tín hiệu thứ ba là tín hiệu bị đánh giá thấp nhất và cũng là tín hiệu mà địa phương kiểm soát được nhiều nhất. Một tỉnh không thể ra lệnh cho hệ thống suy luận tin mình. Một tỉnh hoàn toàn có thể làm cho tất cả các nguồn nói về mình ngừng mâu thuẫn nhau.

Và mâu thuẫn là tình trạng phổ biến. Cùng một di tích có ba năm xây dựng khác nhau trên ba tài liệu chính thức. Cùng một lễ hội có hai khung thời gian tổ chức. Cùng một bãi biển có ba con số chiều dài. Không con số nào sai một cách cố ý. Chúng đến từ các đợt khảo sát khác nhau, các cơ quan khác nhau, các năm khác nhau và chưa từng có ai được giao nhiệm vụ chuẩn hóa chúng.

Có một tình huống mà mâu thuẫn giữa các nguồn lại là tín hiệu tốt và bỏ qua nó sẽ dẫn tới lời khuyên sai. Với các chủ đề đang tranh cãi thật, chẳng hạn tác động môi trường của một dự án hay số liệu về sức chứa của một điểm đến, sự tồn tại của nhiều con số khác nhau phản ánh đúng thực tại chứ không phải sự cẩu thả. Ép chúng về một con số duy nhất để máy dễ đọc là làm giả sự đồng thuận. Nguyên tắc đúng là mọi khác biệt phải có lý do ghi rõ, chứ không phải mọi thứ phải khớp và giữa hai điều đó có một khoảng cách lớn về mặt phương pháp.

### 2.4.4. Bộ lọc thứ ba, xếp hạng

Qua được hai tầng đầu, một địa phương trở thành ứng viên hợp lệ. Bây giờ mới tới cạnh tranh theo nghĩa quen thuộc, trong các ứng viên hợp lệ, chọn cái nào. Điều làm tầng này khác với xếp hạng tìm kiếm là nó không xếp theo mức độ phù hợp với một cụm từ khóa mà theo mức độ phù hợp với một tập ràng buộc do người dùng nêu. Câu hỏi không phải trang nào nói nhiều nhất về ẩm thực bản địa, mà là nơi nào thỏa mãn đồng thời năm ngày, ngân sách vừa, ăn uống tốt, ít khách phương Tây, bay từ Singapore. Một địa phương thắng ở tầng này không phải vì nó nổi bật ở một chiều mà vì nó không bị loại ở chiều nào.

Hệ quả chiến lược của điều đó ngược với trực giác marketing thông thường. Trong nền kinh tế chú ý, sự cực đoan được thưởng, vì cái gì gây ấn tượng mạnh mới được nhớ. Trong xếp hạng theo ràng buộc, sự cực đoan bị phạt, vì mỗi chiều cực đoan là một chiều có nguy cơ vi phạm ràng buộc của ai đó.

Loại bằng chứng được tính ở tầng này cũng khác. Điều địa phương tự nói về mình gần như không có trọng số. Điều các nguồn độc lập nói mới quan trọng và trọng số ấy tăng theo mức độ khó làm giả của nguồn.

### 2.4.5. Bộ lọc thứ tư, thực thi

Tầng cuối cùng là tầng mà toàn bộ lý thuyết trước đó đi vào hiện thực. Câu hỏi ở đây là tôi có giao dịch được với nó không. Không phải nó có hay không. Không phải nó có tốt không. Có đặt được chỗ không, có biết giá không, có xác nhận được không, có đổi được khi hỏng không, có đường hoàn tiền không.

Đây là tầng không có khái niệm gần đúng. Một nhà hàng nổi tiếng khắp thế giới nhưng chỉ nhận đặt bàn qua điện thoại, trong giờ hành chính, bằng tiếng Việt, là một nhà hàng mà tác nhân không thể hoàn tất chuỗi hành động. Nó vẫn có thể được nhắc tới trong câu trả lời. Nó sẽ không có mặt trong phương án được đặt.

Chương trước đã dẫn con số 11% tổ chức khách sạn có khả năng để một tác nhân hoàn tất đặt phòng theo thời gian thực, kèm cảnh báo về nguồn gốc tài trợ của khảo sát ấy.[1] Kể cả khi giảm mạnh con số đó, hình dạng của vấn đề vẫn giữ nguyên và hình dạng ấy mới là điều đáng nói, khoảng cách lớn nhất giữa các doanh nghiệp du lịch hiện nay không phải khoảng cách về chất lượng dịch vụ mà về khả năng được một phần mềm hoàn tất giao dịch. Ba tầng đầu quyết định anh có được nghĩ tới hay không. Tầng thứ tư quyết định anh có được trả tiền hay không. Tầng thực thi thường được trình bày như một bài toán kỹ thuật thuần túy và cách trình bày ấy che mất phần khó thật. Mở giao diện cho tác nhân truy vấn tồn kho và giá đồng nghĩa với việc công khai cấu trúc giá cho mọi bên so sánh tức thời, mất khả năng phân biệt giá theo kênh và chấp nhận một tầng trung gian mới có thể thu phí. Đây chính xác là cuộc chiến mà ngành khách sạn đã trải qua với các sàn đặt phòng trực tuyến hai thập niên trước và bên thua trong cuộc chiến ấy có lý do rất chính đáng để không muốn lặp lại. Sự chậm chạp của 89% còn lại không hoàn toàn là quán tính.[2] Một phần là ký ức.

1. Khảo sát của Aven Hospitality và h2c, tháng 6 năm 2026.

2. Phần bù của khảo sát Aven Hospitality và h2c ngày 22 tháng 6 năm 2026. Nơi nguồn này được dẫn đầy đủ kèm cảnh báo rằng đây là nội dung tài trợ.

### 2.4.6. Cái vòng: ký ức thuật toán

Cần định nghĩa ngay để tránh một cách hiểu sai. Ký ức thuật toán không phải một kho lưu trữ duy nhất mà máy mở ra khi được hỏi về một địa phương. Nó là tổng hợp phân tán của trọng số mô hình, chỉ mục tìm kiếm, đồ thị tri thức, dữ liệu nền tảng, bộ nhớ đệm và các dấu vết số có khả năng được truy hồi. Không bên nào sở hữu toàn bộ ký ức ấy và cũng không bên nào xóa được nó bằng một lần yêu cầu. Bốn bộ lọc mô tả một lần chọn. Nhưng lựa chọn của máy không diễn ra một lần rồi thôi. Thuật toán xây dựng một không gian vector có trọng số cho từng nhân tố mà các nhân tố hay trọng số này lại được tích lũy điều chỉnh theo thời gian nhờ các vòng lặp.

Ký ức thuật toán (algorithmic memory) là cơ chế theo đó mỗi giao dịch hoàn tất sinh ra dấu vết mới và các dấu vết ấy quay lại làm đầu vào cho lần đọc sau, khiến vị thế của một thực thể trước máy tích lũy theo thời gian và phụ thuộc vào đường đi đã qua.

Ba hệ quả của cái vòng cần được nêu riêng. *Thứ nhất*, nó biến khả đọc từ một trạng thái thành một quá trình. Không có mức khả đọc nào đạt được rồi giữ mãi. Có một dòng chảy và một địa phương hoặc đang bồi thêm hoặc đang bị xói. *Thứ hai*, nó tạo ra hiện tượng khóa chặt. Sau đủ nhiều vòng lặp, mô tả về một địa phương trở nên khó thay đổi không phải vì ai bảo vệ nó mà vì nó đã được lặp lại quá nhiều lần ở quá nhiều nguồn. Một tỉnh đã mười năm chuyển mình thành trung tâm công nghiệp vẫn có thể bị mô tả bằng ba thuộc tính nông nghiệp của những năm 2010, đơn giản vì khối văn bản nói về nông nghiệp lớn gấp trăm lần khối văn bản nói về công nghiệp. *Thứ ba* và đây là hệ quả tàn nhẫn nhất, chi phí sửa sai tăng theo mỗi vòng. Sửa một mô tả sai trong năm đầu tiên là công việc của một người trong vài tháng.

Sửa cùng mô tả ấy sau bảy năm là công việc của một chương trình truyền thông nhiều năm với ngân sách không nhỏ và ngay cả khi đó cũng không chắc thành công.

Trở lại với dòng sông. Phù sa bồi vào chỗ dòng chảy chậm và chỗ đã bồi thì dòng càng chậm, nên càng bồi thêm. Cù lao lớn dần không phải vì nó tốt hơn phần sông xung quanh mà vì nó đã bắt đầu trước. Sau vài trăm năm, không ai còn hỏi vì sao cù lao ở đó.

### 2.4.7. Cái van: quyền mờ

Còn lại bộ phận dễ bị đọc lướt nhất của mô hình và quan trọng nhất về mặt chuẩn tắc. Quyền mờ (right to opacity) là quyền của một cộng đồng quyết định phần nào trong tri thức, thực hành và đời sống của mình không đưa vào lược đồ, không phơi ra cho máy đọc, kể cả khi việc phơi ra sẽ mang lại lợi ích kinh tế đo được.

Thuật ngữ này bắt nguồn từ Édouard Glissant, người từ năm 1990 đã đòi cho các nền văn hóa cái quyền không bị làm cho trong suốt trước con mắt kẻ khác. Lập luận của Glissant đi xa hơn quyền riêng tư theo nghĩa thông thường. Ông cho rằng đòi hỏi phải hiểu được là một hình thức của quyền lực, vì nó buộc bên bị hiểu phải tự dịch mình sang các phạm trù của bên hiểu và bản dịch ấy luôn diễn ra theo điều kiện của kẻ mạnh hơn.[1]

Cuốn sách này dùng thuật ngữ theo nghĩa hẹp hơn Glissant, ở tầng quản trị dữ liệu chứ không ở tầng triết học. Nhưng lý do giữ

1. Édouard Glissant, Poétique de la Relation, Gallimard, Paris, 1990. Khái niệm quyền mờ được trình bày trong tiết "Pour l'opacité".

lại cái van thì không hẹp chút nào. Nó nằm trong mô hình vì hai lý do độc lập nhau. *Lý do thứ nhất* đã được chứng minh ở chương trước, có những giá trị chỉ tồn tại khi chưa bị nén. *Lý do thứ hai* là bài học của Scott, mọi dự án khả đọc hóa toàn phần đều kết thúc tồi và chúng kết thúc tồi theo cùng một cách, tức là bản đồ được tin hơn lãnh thổ cho tới lúc lãnh thổ chết.

Một mô hình khả đọc không có van là bản thiết kế cho chính loại thảm họa mà phả hệ của khái niệm này được lập ra để cảnh báo. Cái van vận hành được ở ba mức và cả ba đều là quyết định của cộng đồng chứ không phải của kỹ thuật viên. Mức chọn lọc, quyết định cái gì vào lược đồ với đầy đủ chi tiết và cái gì chỉ hiện diện ở mức chỉ dẫn, như một tấm biển chỉ đường về một ngôi làng mà tấm biển không mô tả ngôi làng. Mức trì hoãn, thừa nhận rằng có những thứ chỉ nên khả đọc sau khi cộng đồng đã đủ vững để giữ nhịp của mình trước dòng khách mới. Mức mặc cả, dùng phần đã khả đọc để tài trợ và che chắn cho phần giữ mờ, theo đúng logic mà các bảo tàng lớn vẫn dùng khi trưng một phần bộ sưu tập để nuôi phần nằm trong kho.

Khái niệm quyền mờ có một nguy cơ bị lạm dụng mà người đề xuất nó cần nói ra trước khi bị người khác nói hộ. Trong tay một chính quyền địa phương muốn tránh giám sát, nó là một lý lẽ hoàn hảo để không công bố dữ liệu về chất lượng nước, về sức chứa thật của một điểm đến, về số vụ tai nạn hay về các khoản thu từ du lịch. Ranh giới giữa bảo vệ tri thức cộng đồng và che giấu trách nhiệm giải trình không tự vạch ra và nó không thể được vạch bằng nguyên tắc chung. Cuốn sách đề xuất một phép thử thực dụng, quyền mờ chính đáng khi phần được giữ mờ thuộc về cộng đồng và do cộng đồng quyết định, không chính đáng khi phần được giữ mờ là thông tin về hành vi của chính cơ quan đang quyết định giữ mờ nó.

### 2.4.8. Bốn bộ lọc nhìn cùng một lúc

Sau khi đi qua từng bộ lọc riêng lẻ, có thể thấy chúng không chỉ khác nhau về chức năng mà còn khác nhau về logic đầu tư, chủ thể thực hiện và tốc độ tạo ra kết quả. Chính những khác biệt này giải thích vì sao nhiều địa phương đầu tư rất nhiều nhưng vẫn không cải thiện được vị thế của mình trong nền kinh tế tác nhân.

Điều đầu tiên cần nhận ra là hai bộ lọc đầu tiên có chi phí thấp nhưng tạo ra giá trị nền tảng lớn nhất. Chuẩn hóa tên gọi, thống nhất thực thể, công bố dữ liệu gốc hay xác lập nguồn thông tin có thẩm quyền đều là những việc không đòi hỏi công nghệ phức tạp hay nguồn lực lớn. Quan trọng hơn, đây là những điều kiện tiên quyết. Nếu chúng chưa được hoàn thành thì mọi nỗ lực ở các bước sau đều đứng trên một nền móng thiếu ổn định.

Ngược lại, bộ lọc xếp hạng lại không thể được tạo ra bằng mệnh lệnh hành chính hay ngân sách truyền thông. Không một địa phương nào có thể tự quyết định vị trí của mình trong lựa chọn của máy. Thứ quyết định vị trí ấy là hàng nghìn tín hiệu độc lập được tích lũy từ người dùng, doanh nghiệp, báo chí, nền tảng và các nguồn tri thức khác. Địa phương không tạo ra các tín hiệu đó. Việc của địa phương là tạo điều kiện để chúng xuất hiện và luôn phản ánh đúng thực tế.

Bộ lọc thực thi cũng có một đặc điểm riêng. Khả năng giao dịch không nằm chủ yếu trong tay cơ quan quản lý mà nằm ở năng lực số của từng doanh nghiệp. Một điểm đến chỉ thật sự "thực thi được" khi khách sạn có thể mở tồn kho, nhà hàng có thể nhận đặt bàn, bảo tàng có thể bán vé và dịch vụ vận tải có thể xác nhận giao dịch bằng máy. Chính quyền khó có thể làm thay những công việc ấy, nhưng có thể làm cho chúng trở nên

dễ dàng hơn thông qua tiêu chuẩn chung, hạ tầng chung và các chương trình hỗ trợ chuyển đổi số.

Từ góc nhìn quản trị, điều quan trọng không phải là hoàn thành bốn bộ lọc theo một trình tự cứng nhắc. Hai bộ lọc đầu cần được ưu tiên vì chúng tạo ra nền móng với chi phí thấp và hiệu quả nhanh. Bộ lọc thực thi nên phát triển song song với quá trình chuyển đổi số của doanh nghiệp. Riêng bộ lọc xếp hạng không phải là một dự án có ngày nghiệm thu, mà là kết quả tích lũy của niềm tin và uy tín qua thời gian. Một địa phương có thể đầu tư để trở nên đọc được, có thể xây dựng để trở nên đáng tin, nhưng chỉ có thể được xếp hạng cao khi toàn bộ hệ sinh thái liên tục tạo ra những bằng chứng khiến máy ngày càng tin tưởng lựa chọn nó.

### 2.4.9. Mô hình này có thể sai ở đâu

Một mô hình được trình bày mà không kèm điều kiện sụp đổ của nó thì không phải một mô hình mà là một niềm tin. Mục này nêu ba cách mà mô hình bốn bộ lọc có thể sai, xếp theo mức độ mà tác giả cho là đáng lo.

*Cách thứ nhất* và là cách đã nêu ở phần định nghĩa, là nếu xuất hiện những trường hợp mà một thực thể giao dịch được với tác nhân số trong khi chưa được phân giải rõ ràng. Điều này có thể xảy ra nếu các hệ thống chuyển sang một kiến trúc mà việc nhận diện được thực hiện ngay tại thời điểm giao dịch thay vì dựa trên tri thức có sẵn. Nếu vậy, quan hệ điều kiện cần giữa các tầng bị phá và mô hình phải được viết lại.

*Cách thứ hai* là nếu hóa ra tầng xếp hạng bị chi phối áp đảo bởi các thỏa thuận thương mại giữa nền tảng với nhà cung cấp lớn,

tới mức bằng chứng độc lập chỉ còn vai trò trang trí. Trong kịch bản ấy, mô hình vẫn mô tả đúng cơ chế kỹ thuật nhưng mô tả sai cơ chế kinh tế và lời khuyên rút ra từ nó sẽ dẫn các địa phương đầu tư vào một cuộc chơi mà kết quả đã được định trước ở nơi khác. Đây là kịch bản mà tác giả cho là đáng lo nhất và nó không kiểm chứng được từ bên ngoài, vì các thỏa thuận ấy không công khai.

*Cách thứ ba* là nếu tốc độ chuyển sang ủy nhiệm chậm hơn nhiều so với giả định, chậm tới mức trong vòng đời của một chu kỳ đầu tư công, các địa phương đầu tư sớm không thu được gì trong khi các địa phương chờ đợi không mất gì. Bằng chứng hiện có không loại trừ kịch bản này. Một giao diện hoàn toàn sinh ra đã vượt một tỷ người dùng hằng tháng theo công bố của bên vận hành, trong khi cùng thời kỳ, các đo lường độc lập ghi nhận nó chỉ chiếm khoảng một phần ba phần trăm tổng số lượt tìm kiếm.[1]

Một tỷ người dùng và ba phần nghìn số lượt tìm kiếm. Cả hai con số cùng đúng và khoảng cách giữa chúng là toàn bộ sự bất định của thập niên này. Cách đọc thận trọng nhất với cặp số liệu ấy là như sau. Quy mô người dùng cho biết hạ tầng và thói quen đang được hình thành. Tỷ trọng lượt truy vấn cho biết hành vi hằng ngày chưa dịch chuyển. Khoảng cách giữa hai chỉ số là khoảng thời gian mà một địa phương còn có thể chuẩn bị mà chưa bị phạt vì chậm.

Lập luận về cửa sổ chuẩn bị có một cách đọc ngược mà người lập kế hoạch nên cân nhắc trước khi trình đề án. Nếu tỷ trọng thực tế còn nhỏ như vậy, thì trong ba tới bốn năm tới, mọi khoản

1. Con số một tỷ người dùng hằng tháng do Google công bố tại hội nghị lập trình viên I/O tháng 5 năm 2026. Ở chiều đo lường độc lập, SparkToro ghi nhận chỉ 0,34% số lượt tìm kiếm chuyển sang giao diện này trong bốn tháng đầu năm 2026. Hai con số đo hai thứ khác nhau, một bên là số người, một bên là tỷ trọng truy vấn.

đầu tư vào khả đọc sẽ không tạo ra doanh thu đo được, trong khi vẫn phải giải trình hằng năm. Điều đó có nghĩa là rủi ro lớn nhất của một chương trình khả đọc không phải rủi ro kỹ thuật mà là rủi ro bị cắt ngân sách ở năm thứ hai vì chưa thấy kết quả. Cách phòng vệ hợp lý là chọn các hạng mục có giá trị độc lập với kịch bản tác nhân, chẳng hạn một bộ dữ liệu gốc chuẩn hóa vốn đã cần cho công tác quản lý thường ngày, thay vì các hạng mục chỉ có nghĩa nếu dự đoán đúng.

### ĐIỀU NÀY CÓ Ý NGHĨA GÌ ĐỐI VỚI MỘT ĐỊA PHƯƠNG?

*Mô hình bốn bộ lọc cho biết thứ tự đầu tư chứ không chỉ cho biết cách gọi tên vấn đề. Tối ưu bộ lọc xếp hạng khi bộ lọc phân giải chưa qua là vô nghĩa, vì bốn khâu nối nhau bằng phép nhân. Trong ba tới bốn năm tới, phần lớn khoản chi cho khả đọc sẽ chưa tạo ra doanh thu đo được, trong khi vẫn phải giải trình hằng năm. Vì vậy nên chọn trước những hạng mục có giá trị độc lập với kịch bản tác nhân, chẳng hạn bộ dữ liệu gốc chuẩn hóa vốn đã cần cho công tác quản lý thường ngày.*

## 2.5. KHI THẾ GIỚI TỰ SỬA MÌNH THEO LƯỢC ĐỒ

Tháng Tư năm 1956, một tàu chở dầu hoán cải rời cảng Newark mang theo 58 chiếc thùng thép hình hộp.[1] Người đứng sau chuyến đi ấy là Malcom McLean, một chủ hãng xe tải, không phải một người của ngành hàng hải. Ý tưởng của ông không có gì tinh vi, thay vì bốc dỡ từng kiện hàng rời, hãy đóng tất cả vào những chiếc hộp cùng kích thước. Chiếc hộp không phải phát minh. Cái mới là sự đồng ý về kích thước. Trong ba thập niên sau đó, thế giới tự sắp xếp lại quanh chiếc hộp ấy. Cần cẩu được thiết kế lại. Tàu được đóng lại. Toa xe lửa, rơ moóc, kho bãi, thậm chí kích thước của chính hàng hóa bên trong đều được điều chỉnh cho vừa. Các cảng cổ nằm sâu trong lòng thành phố, nơi lực lượng bốc xếp thủ công đã làm việc qua nhiều thế hệ, không có chỗ cho cần cẩu và bãi chứa. Khu cảng của New York chuyển sang bờ Newark. Các bến tàu ở London đóng cửa và lưu lượng dời về Felixstowe. Không ai ra lệnh cho ngành vận tải biển phải tái cấu trúc. Không có quy định nào bắt các cảng phải chết. Một chiếc hộp thép không có quyền lực nào. Nó chỉ có một kích thước và thế giới đã tự đo mình theo kích thước ấy.

### 2.5.1. Phát biểu mệnh đề

Câu chuyện container không phải một ẩn dụ về trí tuệ nhân tạo.[2] Nó là một trường hợp của cùng một quy luật và quy luật ấy là đóng góp lý thuyết trung tâm của cuốn sách này. Cơ chế tái

1. Chuyến đi của tàu Ideal X, rời cảng Newark ngày 26 tháng 4 năm 1956 với 58 thùng thép. Marc Levinson, The Box, Princeton University Press, Princeton, 2006, chương 1.

2. Marc Levinson, The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger, Princeton University Press, Princeton, 2006. Chuyến tàu container đầu tiên năm 1956 chở 58 thùng thép.

cấu trúc khả đọc khi một bộ máy có quyền phân bổ nguồn lực bắt đầu đọc một lĩnh vực, lĩnh vực ấy tự tổ chức lại theo lược đồ của bộ máy. Tốc độ tái cấu trúc tỷ lệ thuận với tỷ trọng nguồn lực đi qua bộ máy đó và không phụ thuộc vào việc bộ máy có quyền cưỡng chế hay không.

Mệnh đề này bác bỏ được nếu tìm được một lĩnh vực trong đó một tỷ trọng lớn nguồn lực đi qua một bộ máy đọc mà lĩnh vực ấy không thay đổi cấu trúc của mình, mệnh đề sai. Nếu tìm được một lĩnh vực tái cấu trúc mạnh trong khi tỷ trọng nguồn lực đi qua bộ máy vẫn nhỏ, quan hệ tỷ lệ trong phát biểu sai.

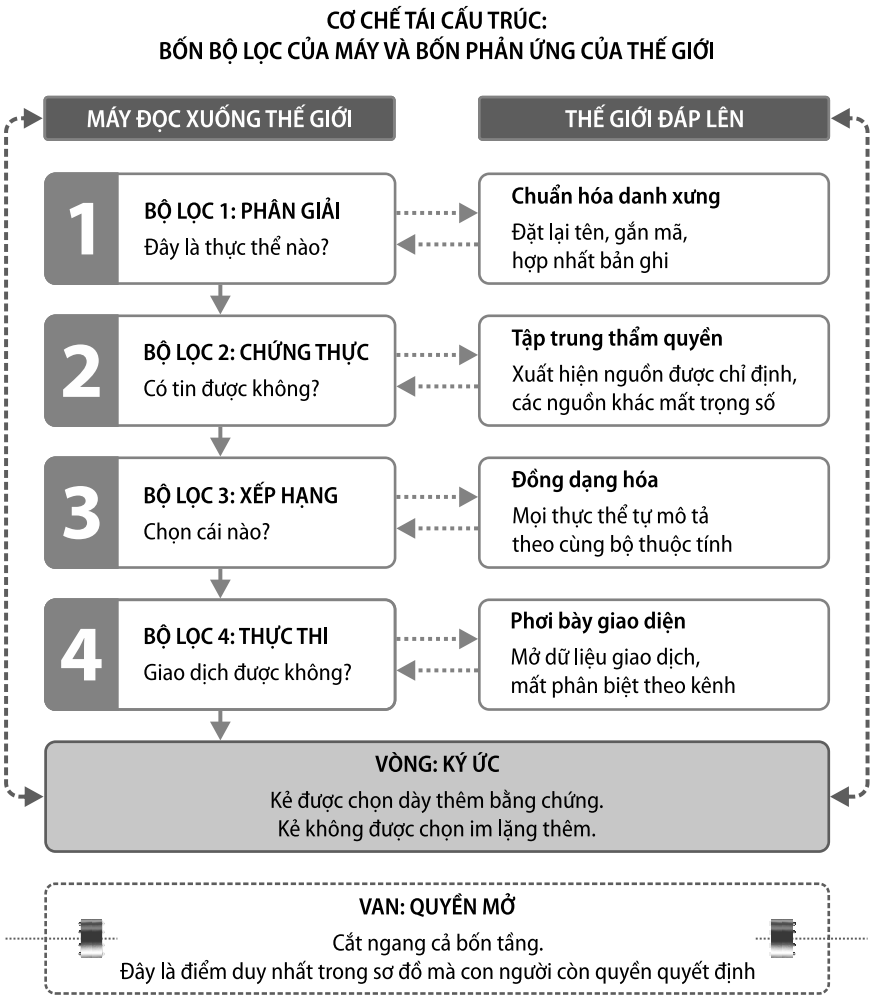
Bốn hệ luận đi kèm và chúng mới là phần có ích cho công việc quản trị.

- ***Tái cấu trúc không cần mệnh lệnh.*** Nó diễn ra qua hàng nghìn quyết định hợp lý của từng chủ thể riêng lẻ đang cố sống sót và vì thế không có ai để quy trách nhiệm.
- ***Phần không được lược đồ ghi nhận không bị xóa.*** Nó teo đi. Sự khác nhau giữa xóa và teo là sự khác nhau giữa một vụ án và một xu hướng và chỉ vụ án mới có bị cáo.
- ***Kẻ tái cấu trúc sớm hưởng chênh lệch.*** Kẻ muộn trả chi phí chuyển đổi cao hơn, vì đến lúc đó chuẩn đã cứng lại và không còn thương lượng được.
- ***Sau khi tái cấu trúc hoàn tất, lược đồ trở thành hiện thực.*** Không ai còn nhớ rằng đã từng có lựa chọn khác và câu hỏi vì sao lại là kích thước này thôi được đặt ra.

Hệ luận thứ tư là hệ luận đáng lo nhất. Nó có nghĩa rằng cửa sổ để tranh luận về lược đồ luôn khép lại trước khi phần lớn những người bị ảnh hưởng kịp nhận ra rằng có một cuộc tranh luận.

### 2.5.2. Bốn bộ lọc và bốn phản ứng

Mô hình ở phần trước mô tả cách máy đọc. Áp cơ chế tái cấu trúc vào mô hình ấy, ta có mặt còn lại, cách thế giới đáp lại việc bị đọc. Mỗi bộ lọc tạo ra một áp lực và mỗi áp lực tạo ra một phản ứng dự đoán được.




**Hình 2.2.** Cơ chế tái cấu trúc, bốn bộ lọc của máy và bốn phản ứng của thế giới

Cột phải là phần mà các nghiên cứu về trí tuệ nhân tạo hầu như không bàn tới, vì chúng đứng ở phía cỗ máy. Nó cũng là phần quan trọng hơn đối với người làm quản trị, vì cột trái nằm ngoài tầm với còn cột phải là những gì sẽ xảy ra trong chính tổ chức của họ.

Bốn phản ứng ấy đều đã được quan sát trong các lĩnh vực khác, ở những thời điểm khác nhau, dưới những cái tên khác nhau. Điều mà mô hình này khẳng định là chúng là bốn biểu hiện của cùng một áp lực chứ không phải bốn hiện tượng rời rạc và chúng xuất hiện theo đúng thứ tự đó.

### 2.5.3. Cùng một mệnh đề, năm lĩnh vực

Nếu mệnh đề đúng, nó phải mô tả được những lĩnh vực không liên quan gì tới du lịch. Bảng 2.1 kiểm tra điều đó.

**Bảng 2.1.** Cơ chế tái cấu trúc trong năm lĩnh vực không liên quan tới nhau

| Lĩnh vực | Bộ máy đọc | Phản ứng đã quan sát được |
|---|---|---|
| **Tuyển dụng** | Phần mềm sàng hồ sơ | Hồ sơ được viết cho máy đọc trước, cho người đọc sau. Xuất hiện cả một nghề dạy cách viết hồ sơ qua được máy |
| **Y tế** | Hệ mã bệnh và hệ thanh toán bảo hiểm | Chẩn đoán được diễn đạt theo mã có mức chi trả. Bệnh không có mã trở thành bệnh không được điều trị |
| **Tín dụng** | Mô hình chấm điểm | Người vay tạo lịch sử giao dịch để có điểm. Nhóm không có dữ liệu bị loại trước khi được đánh giá |
| **Học thuật** | Cơ sở dữ liệu trích dẫn | Đề tài nghiên cứu dịch về phía những gì đo được bằng chỉ số trích dẫn. Ngành không có tạp chí trong danh mục teo dần |
| **Thương mại điện tử** | Bộ lọc và xếp hạng của sàn | Sản phẩm được thiết kế lại cho vừa ô danh mục. Tên hàng hóa viết cho thuật toán, không cho người mua |

Không lĩnh vực nào trong bảng cần tới trí tuệ nhân tạo để hiện tượng xảy ra. Hệ mã bệnh có từ giữa thế kỷ trước. Chấm điểm tín dụng có từ thập niên 1950. Cơ sở dữ liệu trích dẫn có từ 1964. Đây là điểm mà cuốn sách muốn nhấn mạnh nhất và nó ngược với cách phần lớn các cuốn sách về chủ đề này được viết. Trí tuệ nhân tạo không tạo ra cơ chế tái cấu trúc. Nó chỉ làm ba việc: mở rộng phạm vi các lĩnh vực bị đọc, rút ngắn chu kỳ phản ứng từ hàng chục năm xuống vài năm và xóa mất chủ thể chịu trách nhiệm về lược đồ. Việc thứ ba là việc nghiêm trọng nhất. Hệ mã bệnh do một tổ chức quốc tế soạn thảo, có phiên bản, có quy trình sửa đổi và có địa chỉ để khiếu nại. Lược đồ mà một hệ thống suy luận dùng để đọc thế giới không có ba thứ đó.

### 2.5.4. Vì sao địa phương là trường hợp khó nhất

Nếu mệnh đề là phổ quát, câu hỏi tự nhiên là vì sao dành cả một cuốn sách cho địa phương thay vì cho một lĩnh vực nào khác. Câu trả lời không phải vì địa phương quan trọng hơn. Vì địa phương là trường hợp mà cả bốn điều kiện làm cho tái cấu trúc trở nên khó khăn đều xuất hiện cùng lúc. Một doanh nghiệp thương mại điện tử muốn sửa cách mình được máy đọc thì họp một buổi và sửa. Một địa phương thì không, vì bốn lý do:

- ***Không có chủ sở hữu duy nhất.*** Hình ảnh của một tỉnh được tạo bởi hàng nghìn chủ thể độc lập mà chính quyền không có thẩm quyền ra lệnh, từ chủ homestay tới người viết đánh giá.

- ***Tốc độ sửa chậm hơn tốc độ áp lực.*** Một chuỗi bán lẻ đổi cấu trúc dữ liệu trong một quý. Một tỉnh cần một quy trình phê duyệt, một khoản ngân sách và thường là một nhiệm kỳ.

- ***Phần giá trị nhất là phần không nén được.*** Với hàng hóa, gần như mọi thuộc tính đáng giá đều đo được. Với một vùng đất, phần đáng giá nhất lại là phần kháng lại phép đo.

- ***Hậu quả không đảo ngược được.*** Một sản phẩm bị mô tả sai thì đổi bao bì. Một cộng đồng đã tự sửa mình theo lược đồ trong hai thế hệ thì không có bao bì nào để đổi.

Địa phương, nói cách khác, không phải một ví dụ minh họa cho mệnh đề. Nó là trường hợp giới hạn, nơi mệnh đề lộ ra rõ nhất chính vì đối tượng chống lại nó mạnh nhất. Bốn đặc điểm trên được trình bày như bất lợi và ba trong bốn đúng là bất lợi. Đặc điểm thứ nhất thì không hẳn. Việc không có chủ sở hữu duy nhất cũng có nghĩa là không có điểm đơn nào để bị bắt bí. Một chuỗi khách sạn phụ thuộc vào một nền tảng phân phối có thể bị nền tảng ấy siết. Một vùng đất mà hàng nghìn người kể chuyện về nó thì khó bị siết hơn nhiều, vì không ai sở hữu nguồn bằng chứng. Trong ba bộ lọc đầu, sự phân tán là điểm yếu. Ở bộ lọc thứ tư, nó lại là hình thức bảo hiểm rẻ nhất mà một địa phương có.

### 2.5.5. Ba câu hỏi mệnh đề này đặt ra mà lý thuyết cũ không đặt

Scott mô tả nhà nước làm cho xã hội khả đọc và cảnh báo về cái giá. Bowker và Star mô tả cách phân loại ghi nhớ và lãng quên. Cả hai đứng ở phía người bị đọc và cả hai giả định rằng có một bộ máy xác định đứng ở phía bên kia, với một địa chỉ, một thẩm quyền và một lịch sử. Ba câu hỏi sau không nằm trong khung của họ và cuốn sách này lấy chúng làm chương trình.

*Thứ nhất*, điều gì xảy ra khi bộ máy đọc không có chủ quyền và cũng không muốn có. Toàn bộ kho công cụ chính trị mà xã hội tích lũy để đối phó với nhà nước, từ bầu cử tới toà án, tới báo chí, đều giả định một đối tượng có địa chỉ. Chưa có công cụ nào cho một đối tượng chỉ đơn giản là không nhắc tới ta.

*Thứ hai*, điều gì xảy ra khi lược đồ thay đổi nhanh hơn chu kỳ thích nghi của xã hội. Hệ mã bệnh sửa vài năm một lần và ngành y có thời gian điều chỉnh. Nếu trọng số của một nguồn thay đổi trong vài tuần, không thiết chế xã hội nào theo kịp và hệ quả sẽ là dao động chứ không phải thích nghi.

*Thứ ba*, ai giữ cái van khi không ai sở hữu bộ lọc. Câu hỏi này chưa có lời giải trong bất kỳ nhánh nào của luật học hay lý thuyết quản trị và cuốn sách sẽ không giả vờ giải được nó ở chương này.

Scott hỏi nhà nước nhìn thấy gì. Câu hỏi của thời này khác, xã hội đang tự sửa mình thành cái gì, để được nhìn thấy.

### ĐIỀU NÀY CÓ Ý NGHĨA GÌ ĐỐI VỚI MỘT ĐỊA PHƯƠNG?

*Cơ chế tái cấu trúc không chờ trí tuệ nhân tạo và nó đã vận hành ở nhiều ngành từ trước. Điều cần đề phòng là địa phương tự sửa mình theo một lược đồ mà mình không hiểu và không được thông báo khi lược đồ đổi. Cách phòng vệ rẻ nhất là ghi lại. Mỗi lần thay đổi cách mô tả một tài sản để hợp với một hệ thống bên ngoài, ghi rõ đã đổi gì, vì hệ thống nào và mất gì. Sau hai năm, tập ghi chép ấy cho biết địa phương đang đi về hướng nào.*

## 2.6. PHẢN BIỆN MÔ HÌNH

Bất kỳ một mô hình nào cũng mang sẵn trong nó tính có thể sai của giả thuyết vì chấp nhận giả thiết. Một mô hình chỉ đáng tin sau khi chịu được các đòn mạnh nhất nhắm vào nó. Mục này trình bày bốn đòn, theo thứ tự tăng dần về độ nặng và trả lời từng đòn mà không che giấu chỗ mô hình yếu.

### 2.6.1. Máy đang đọc dữ liệu phi cấu trúc tốt hơn, lược đồ còn quan trọng không

Đây là phản bác được nêu nhiều nhất và cũng là phản bác hợp lý nhất. Các mô hình ngôn ngữ ngày càng đọc tốt văn bản tự do. Chúng rút được thông tin từ một bài viết lộn xộn, hiểu được ẩn ý, xử lý được cả những mô tả mâu thuẫn. Nếu vậy, việc bắt các địa phương chuẩn hóa dữ liệu chẳng phải là lo lắng của ngày hôm qua hay sao.

Phản bác này đúng ở tầng đọc hiểu và sai ở tầng hành động và sự phân biệt ấy là toàn bộ câu trả lời. Một mô hình có thể hiểu rất tinh một đoạn văn mô tả không khí buổi sớm ở chợ Đông Ba (Huế) mà vẫn không giữ được bàn cho ai, vì giữ bàn đòi hỏi một giao diện trả lời được câu hỏi còn hay hết ngay tại thời điểm hỏi. Càng giỏi đọc văn bản tự do, máy càng nhanh chóng đi tới giới hạn thật và giới hạn thật nằm ở bộ lọc thứ tư.

Có một hệ quả ngược đời của chính sự tiến bộ ấy mà những người nêu phản bác thường không tính tới. Khi khả năng đọc hiểu tăng, lợi thế của việc viết hay giảm, vì mọi bên đều được đọc hiểu tốt như nhau. Cái không giảm là lợi thế của việc giao dịch được. Tiến bộ trong đọc hiểu, nói cách khác, đang dịch chuyển giá trị từ tầng nội dung sang tầng hạ tầng, tức là theo đúng hướng mà mô hình này dự đoán.

### 2.6.2. Danh tiếng có thay thế khả đọc không

Phản bác thứ hai đến từ những nơi đã nổi tiếng. Nếu Paris vẫn xuất hiện trong mọi gợi ý du lịch châu Âu bất kể dữ liệu chính thức của nó lộn xộn thế nào, thì phải chăng danh tiếng đủ lớn khiến khả đọc thành chuyện thứ yếu. Câu trả lời là có, với một điều kiện và điều kiện ấy hẹp hơn nhiều so với người ta tưởng.

Danh tiếng thay thế được khả đọc ở ba bộ lọc đầu. Khi có hàng trăm triệu trang viết về một thành phố, máy vẫn phân giải được thực thể, vẫn tìm được nguồn đối chiếu, vẫn có căn cứ xếp hạng, dù cổng thông tin của thành phố ấy có tệ đến đâu. Khối lượng bù cho chất lượng.

Danh tiếng không thay thế được khả đọc ở bộ lọc thứ tư. Không lượng bài viết nào cho máy biết tối thứ Sáu còn bàn trống hay không. Đây là lý do một thành phố nổi tiếng vẫn có thể thua một thị trấn vô danh ở đúng khoảnh khắc giao dịch và cũng là lý do các tập đoàn khách sạn lớn đang chi tiền vào tầng này chứ không vào tầng quảng bá.

### 2.6.3. Khả đọc không đồng nghĩa với dữ liệu mở

Phản bác thứ ba đến từ phía ngược lại, từ những người lo ngại rằng toàn bộ chương trình này là một cách nói khác của việc bắt các địa phương giao nộp dữ liệu miễn phí cho các tập đoàn công nghệ nước ngoài. Lo ngại ấy chính đáng và tác giả xin chia sẻ điều đó. Nhưng nó dựa trên một sự nhầm lẫn giữa hai khái niệm khác nhau.

Dữ liệu mở và tính khả đọc có liên quan nhưng không giống nhau. Mục tiêu của dữ liệu mở là để mọi người đều có thể sử

dụng thông tin, còn mục tiêu của tính khả đọc là để máy hiểu đúng và hành động đúng. Vì vậy, dữ liệu mở hướng tới "ai cũng dùng được", còn tính khả đọc chỉ phơi bày phần thông tin cần thiết để máy suy luận và giao dịch. Khác biệt tiếp theo nằm ở phạm vi và cơ chế tiếp cận. Dữ liệu mở thường theo nguyên tắc càng công khai càng tốt, trong khi tính khả đọc chỉ mở những gì cộng đồng quyết định biểu diễn và có thể đi kèm giấy phép, giới hạn truy cập hoặc thu phí.

Chủ thể quyết định cũng khác nhau. Dữ liệu mở chủ yếu do chính sách công khai thông tin quy định, còn tính khả đọc phản ánh lựa chọn của cộng đồng sở hữu tri thức về những gì nên và không nên để máy đọc. Rủi ro của hai cách tiếp cận cũng khác nhau. Dữ liệu mở có thể bị khai thác mà giá trị không quay trở lại cộng đồng. Tính khả đọc nếu được xây dựng nửa vời sẽ khiến máy hiểu sai và đưa ra các quyết định sai.

Một địa phương hoàn toàn có thể khả đọc mà không mở dữ liệu, bằng cách cấp quyền truy cập có điều kiện, có ghi nhận nguồn, có hạn mức và trong một số trường hợp có thu phí. Mô hình mà một số tổ chức tri thức lớn đang áp dụng, tức là giữ nội dung miễn phí cho người đọc nhưng thu phí truy cập quy mô lớn của các hệ thống thương mại, cho thấy hướng đi này khả thi về mặt kỹ thuật lẫn pháp lý.

Khả năng mặc cả ấy không phân bố đều và đây là chỗ lập luận trên dễ trở thành lời an ủi rỗng. Một tổ chức nắm nguồn tri thức mà mọi mô hình đều cần có thể đặt điều kiện. Một xã miền núi không có gì để mặc cả, vì không hệ thống nào cần dữ liệu của xã ấy đủ nhiều để chấp nhận điều kiện. Quyền thương lượng tỷ lệ thuận với mức độ không thể thay thế và phần lớn địa phương thì thay thế được. Điều này gợi ý rằng đơn vị đàm phán

đúng là cấp quốc gia hoặc một liên minh khu vực chứ không phải từng địa phương và đó là một luận điểm mà chương cuối sẽ trở lại.

### 2.6.4. Giá trị biên của khả đọc

Đòn thứ tư không phải một phản bác mà là một câu hỏi phân phối và trả lời được nó là điều kiện để mô hình có ích cho chính sách: "Nếu khả đọc quan trọng, mọi địa phương có nên đầu tư như nhau không?". Không. Giá trị biên của khả đọc cao nhất ở những nơi nghèo tín hiệu nhất. Hãy đặt cạnh nhau hai trường hợp cực đoan. Paris có hàng trăm triệu trang, sách, video và đánh giá viết về mình qua nhiều thế kỷ. Cải thiện dữ liệu chính thức của Paris làm tăng chất lượng mô tả một chút, nhưng nó không thay đổi xác suất Paris xuất hiện, vì xác suất ấy vốn đã gần một. Đầu tư ở đây có lợi suất giảm dần rất nhanh. Một xã miền núi có thể chỉ xuất hiện vài trăm lần trên toàn bộ internet, phần lớn trong các văn bản hành chính không ai đọc. Ở đó, một bộ dữ liệu chuẩn hóa gồm điểm tham quan, tuyến đường, cơ sở lưu trú, sản phẩm đặc trưng, lịch lễ hội và thông tin song ngữ có thể làm thay đổi xác suất xuất hiện từ gần không lên một con số có nghĩa. Cùng một khoản tiền, hai kết quả không so sánh được. Khả đọc là khoản đầu tư hiếm hoi mà người càng ít có thì càng được lợi nhiều. Cũng vì thế nó luôn được đề xuất bởi những nơi ít cần nó nhất.

Câu trên là một quan sát về kinh tế chính trị của ngân sách, không phải một lời châm biếm. Các thành phố lớn có phòng chuyên trách, có tư vấn, có người theo dõi xu hướng quốc tế, nên họ là nơi đầu tiên nghe về vấn đề và đầu tiên xin được tiền. Các xã nghèo tín hiệu thì không có ai làm việc đó. Kết quả dự đoán được

là nguồn lực chảy ngược chiều với giá trị biên, làm rộng thêm khoảng cách mà lẽ ra nó có thể thu hẹp.

Chỗ này cần dừng lại. Luận điểm về giá trị biên cao có một mặt tối cần được nói thẳng với các địa phương nhỏ trước khi họ vay tiền làm dự án. Xác suất tăng từ gần không lên một con số có nghĩa vẫn có thể là một con số rất nhỏ. Chẳng hạn, trong một kịch bản minh họa, nếu một xã xuất hiện trong 3% số câu trả lời thay vì 0,2%, đó là mức tăng gấp mười lăm lần theo tỷ lệ và gần như không có ý nghĩa thương mại theo giá trị tuyệt đối, trong khi chi phí là thật và phải trả trước.[1] Trình bày lợi suất theo bội số thay vì theo giá trị tuyệt đối là một thủ thuật quen thuộc của các bản đề án và người đọc cuốn sách này nên cảnh giác với nó ngay cả khi nó được dùng để ủng hộ luận điểm của chính cuốn sách.

Còn một giới hạn phải tự khai trước khi bị người khác chỉ ra. Mô hình bốn bộ lọc mô tả điều kiện để được máy chọn, không mô tả điều kiện để đáng được chọn. Một địa phương có thể vượt cả bốn tầng bằng dữ liệu chỉn chu về một sản phẩm bình thường. Khả đọc là điều kiện cần của sự hiện diện, không phải điều kiện đủ của giá trị.

### 2.6.5. Phản biện thứ năm: mô hình này có mô tả đúng Việt Nam không

Bốn phản bác trên đều nhắm vào tính đúng đắn của mô hình. Phản bác thứ năm nhắm vào phạm vi áp dụng của nó và với người đọc Việt Nam thì đây mới là phản bác quan trọng nhất. Toàn bộ

1. Giả định của tác giả. Con số 3% và 0,2% là kịch bản dựng để làm rõ chênh lệch giữa bội số và giá trị tuyệt đối, không phải số liệu quan sát của một địa phương có thật.

bằng chứng dùng trong chương này đến từ thị trường nói tiếng Anh. Các nghiên cứu về tập trung nguồn dẫn đo trên nội dung tiếng Anh. Các số liệu về hành vi tìm kiếm đo trên người dùng Mỹ. Các nền tảng được nhắc tới phần lớn không phải nền tảng mà người Việt dùng hằng ngày. Vậy có lý do gì để tin rằng mô hình bốn bộ lọc mô tả đúng những gì đang xảy ra với một tỉnh ở Việt Nam. Câu trả lời trung thực gồm ba phần, trong đó chỉ phần đầu là tin tốt.

**Phần thứ nhất: cơ chế thì phổ quát**

Bốn bộ lọc không phải một đặc điểm văn hóa. Chúng là bốn khâu bắt buộc của bất kỳ quy trình lựa chọn được ủy nhiệm nào và chúng vận hành như nhau bất kể ngôn ngữ. Một hệ thống nào cũng phải xác định thực thể trước khi xác minh về nó, phải xác minh trước khi xếp hạng, phải xếp hạng trước khi giao dịch. Điều này đúng với một trợ lý tiếng Anh, đúng với một trợ lý tiếng Trung và sẽ đúng với bất kỳ hệ thống nội địa nào được xây trong tương lai.

**Phần thứ hai: mức độ thì khác và khác theo hướng bất lợi**

Ở đây bức tranh xấu đi. Ba đặc điểm của bối cảnh Việt Nam làm bài toán khó hơn so với bối cảnh mà các nghiên cứu quốc tế mô tả.

*Thứ nhất* là vấn đề khối lượng ngữ liệu. Lượng văn bản chất lượng cao bằng tiếng Việt trên internet nhỏ hơn nhiều lần so với tiếng Anh và phần nói về các địa phương ngoài vài trung tâm lớn thì nhỏ tới mức có thể đếm được. Điều này có nghĩa là các hệ thống suy luận biết về phần lớn địa phương Việt Nam chủ yếu qua nguồn tiếng Anh, tức là qua con mắt của người ngoài, tức là qua một lớp diễn giải mà không ai ở địa phương ấy tham gia.

*Thứ hai* là vấn đề danh xưng. Đợt sắp xếp lại đơn vị hành chính quy mô lớn tạo ra một tình huống hiếm gặp trong đó hàng loạt thực thể đổi tên trong khi toàn bộ khối tài liệu lịch sử vẫn dùng tên cũ. Với con người, chuyện này gây phiền nhưng không gây nhầm. Với hệ thống phân giải thực thể, nó là một bài toán khó thật sự và cách nó được giải sẽ định hình cách các địa phương Việt Nam được biểu diễn trong nhiều năm.

*Thứ ba* là vấn đề thẩm quyền. Ở nhiều nước, một cơ quan du lịch quốc gia được các hệ thống nhận diện là nguồn có thẩm quyền cho một số loại dữ kiện. Điều đó đòi hỏi cơ quan ấy phải công bố dữ liệu ở dạng máy đọc được, ổn định qua thời gian, có định danh rõ ràng. Phần lớn dữ liệu chính thức ở Việt Nam hiện được công bố dưới dạng tập tin văn bản dành cho người đọc, thường là bản quét, thường không có cấu trúc và thường được thay thế mà không giữ phiên bản cũ.

Ba đặc điểm ấy cộng lại tạo ra một tình trạng gọi là nghèo tín hiệu có hệ thống. Đây không phải hậu quả của việc thiếu đầu tư công nghệ. Nhiều địa phương Việt Nam đã đầu tư đáng kể vào cổng thông tin, ứng dụng du lịch và các nền tảng số. Vấn đề là gần như toàn bộ khoản đầu tư ấy nhắm vào việc phục vụ mắt người, trong khi bốn bộ lọc thì không có mắt.

**Phần thứ ba: nhưng cửa sổ thì đang mở**

Phần cuối là phần mà một cuốn sách phân tích không nên bỏ qua chỉ vì nó nghe lạc quan. Chính vì nghèo tín hiệu mà giá trị biên của mỗi hành động chuẩn hóa ở Việt Nam cao hơn nhiều so với ở một nước đã bão hòa dữ liệu. Trong một môi trường mà mọi thực thể đều đã được mô tả kỹ, việc mô tả thêm không thay đổi gì. Trong một môi trường mà phần lớn thực thể chưa được

mô tả, bên nào mô tả trước sẽ trở thành nguồn mặc định và như cái vòng ký ức thuật toán đã chỉ ra, vị thế nguồn mặc định có tính tự củng cố.

Nói cách khác, đây là giai đoạn mà chi phí giành lấy quyền định nghĩa còn thấp. Giai đoạn ấy không kéo dài, vì nó khép lại không phải khi một địa phương từ bỏ mà khi một bên khác lấp vào chỗ trống.

Cần nói rõ ai là bên khác ấy, vì để mơ hồ thì lập luận mất sức. Trong thực tế hiện nay, các nguồn đang lấp vào khoảng trống mô tả về địa phương Việt Nam chủ yếu là các nền tảng đánh giá quốc tế, các trang blog du lịch nước ngoài và các nội dung do chính các mô hình sinh ra rồi được đăng lại. Nhóm thứ ba đáng lo nhất và ít được nói tới nhất, khi máy sinh ra một mô tả về một xã, mô tả ấy được đăng lên một trang tổng hợp, rồi trở thành nguồn cho lần sinh sau, ta có một vòng lặp không có bất kỳ con người nào ở địa phương ấy tham gia vào bất kỳ khâu nào. Đây không phải viễn cảnh tương lai. Nó đang diễn ra và tốc độ của nó nhanh hơn tốc độ mà bất kỳ quy trình phê duyệt hành chính nào có thể theo kịp.

Lời khuyên hành động sớm cũng cần được kiểm tra ngược. Hành động sớm trong một lĩnh vực mà các chuẩn kỹ thuật còn đang biến động có rủi ro thật là đầu tư vào một chuẩn sẽ bị bỏ. Đã có nhiều địa phương trên thế giới xây nền tảng dữ liệu du lịch theo một kiến trúc mà vài năm sau không còn ai dùng. Cách hóa giải không phải chờ cho tình hình rõ ràng, vì cửa sổ sẽ khép trước khi điều đó xảy ra, mà là tách bạch hai lớp. Lớp nội dung, tức là bản thân dữ liệu đúng về địa phương, có giá trị bất kể chuẩn nào thắng. Và lớp giao thức, tức là cách phơi dữ liệu ấy ra, nên được coi là hạng mục thay thế được và không nên chiếm

phần lớn ngân sách. Đầu tư vào dữ liệu đúng thì không bao giờ lỗi thời. Đầu tư vào một cách phơi dữ liệu cụ thể thì luôn có hạn sử dụng.

Tổng kết phản bác thứ năm, mô hình mô tả đúng cơ chế nhưng đánh giá thấp mức độ nghiêm trọng khi áp vào bối cảnh Việt Nam. Chương 7 sẽ dành trọn cho việc hiệu chỉnh ấy, với dữ liệu trong nước và với các đề xuất cụ thể tới cấp tỉnh.

### 2.6.6. Sáu phản bác từ phía công nghệ

Năm phản bác trên đến từ phía quản trị. Sáu phản bác dưới đây đến từ phía kỹ thuật và chúng nguy hiểm hơn vì mỗi phản bác đều dựa trên một hướng phát triển đang thực sự diễn ra. Nếu một trong sáu đúng hoàn toàn, phần lớn khuyến nghị của cuốn sách này mất giá trị. Trình bày chúng theo đúng sức mạnh vốn có là nghĩa vụ của tác giả trước người đọc. Bác bỏ chúng bằng cách làm chúng yếu đi trước khi bác bỏ là một hình thức gian lận.

**Phản bác một: tác nhân tự thu thập dữ liệu**

Nếu tác nhân có thể tự truy cập trang web, tự điều hướng giao diện dành cho người, tự điền biểu mẫu, thì việc một địa phương có mở giao diện dữ liệu hay không trở nên vô nghĩa. Máy sẽ tự lấy thứ nó cần, giống như một người dùng kiên nhẫn vô hạn.

Phản bác này đúng về mặt kỹ thuật và sai về mặt kinh tế. Một tác nhân điều hướng giao diện người mất nhiều bước hơn, tốn nhiều tài nguyên tính toán hơn và sai nhiều hơn so với một tác nhân gọi thẳng dữ liệu có cấu trúc. Trong một phiên, tác nhân phải cân nhắc hàng trăm ứng viên. Chi phí của việc được

xem xét, vốn xưa nay bằng không, nay là một khoản chi thật mà bên vận hành tác nhân phải trả. Khi chi phí xem xét khác không, thứ tự xem xét không còn ngẫu nhiên. Bên nào rẻ để đọc sẽ được đọc trước và trong phần lớn trường hợp, được đọc trước là được chọn.

**Phản bác hai: mô hình đa phương thức**

Nếu máy đọc được ảnh, video và âm thanh, thì một địa phương chỉ cần để lại nhiều hình ảnh đẹp và bài toán khả đọc tự tan. Đây là phản bác được nêu nhiều nhất trong giới làm truyền thông và nó chứa một sự lẫn lộn giữa nhận biết với phân giải. Một mô hình đa phương thức nhìn ảnh và biết đó là một ngôi chùa cổ ở Đông Nam Á. Nó không biết đó là chùa nào, ở xã nào, mở cửa mấy giờ và có thu phí không. Nhận biết cho ra loại. Phân giải cho ra thực thể. Bốn bộ lọc bắt đầu từ phân giải, không từ nhận biết.

Có một hệ quả ngược đời đáng ghi nhận. Càng nhiều ảnh đẹp mà không có siêu dữ liệu vị trí đáng tin, khối bằng chứng càng lớn và càng mơ hồ, tức là càng nhiều dữ liệu nhưng độ phân giải không tăng.

**Phản bác ba: mô hình thế giới**

Đây là phản bác mạnh nhất trong sáu phản bác và cuốn sách thừa nhận không bác bỏ được nó hoàn toàn. Lập luận như sau. Nếu các hệ thống tiến tới việc dựng được một mô hình nội tại về thế giới, với quan hệ không gian, nhân quả và thời gian được biểu diễn bên trong, thì chúng sẽ suy ra được rất nhiều thứ. Chúng có thể suy rằng một xã miền núi phía Bắc nhiều khả năng có mùa lúa chín vào tháng Năm mà không cần ai công bố điều đó.

Cuốn sách chấp nhận rằng năng lực suy luận ấy sẽ tăng và sẽ làm giảm nhu cầu chuẩn hóa ở tầng mô tả. Nhưng suy luận là xác suất, còn giao dịch là nhị phân. Một hệ thống có thể suy rất giỏi rằng nơi ấy đẹp vào tháng Năm và vẫn không có cơ sở để giữ chỗ cho ai, vì việc còn phòng hay hết phòng không suy ra được từ mô hình thế giới. Nó chỉ đọc được từ một nguồn có thẩm quyền tại thời điểm hỏi.

Nói cách khác, mô hình thế giới càng mạnh thì giá trị của tầng mô tả càng giảm và giá trị của tầng giao dịch càng tăng. Kết luận của cuốn sách không đảo, nó bị đẩy sâu hơn về phía bộ lọc thứ tư.

**Phản bác bốn: dữ liệu tổng hợp**

Nếu máy có thể tự sinh ra dữ liệu mô tả cho những nơi thiếu dữ liệu, khoảng trống sẽ được lấp mà không cần địa phương làm gì. Khoảng trống đúng là sẽ được lấp. Đó chính là vấn đề.

Dữ liệu tổng hợp về một nơi được suy ra từ các nơi tương tự, nên nó tái tạo mẫu hình trung bình của vùng. Một xã được mô tả bằng chất liệu vay mượn sẽ trở nên giống hệt xã bên cạnh trong biểu diễn, đúng lúc mà điều duy nhất có giá trị kinh tế của nó là chỗ nó khác xã bên cạnh. Và vì mô tả tổng hợp ấy được đăng lại rồi trở thành nguồn cho lần sinh sau, sai lệch được củng cố qua mỗi vòng. Phản bác bốn không bác bỏ mô hình. Nó làm cho mô hình khẩn cấp hơn.

**Phản bác năm: máy tự dựng bản thể luận**

Nếu hệ thống tự sinh ra các phạm trù để phân loại thế giới thay vì dùng lược đồ do người soạn, thì toàn bộ lập luận về việc ai thiết kế các ô sẽ lỗi thời. Điều này đang xảy ra và nó không giải

quyết vấn đề mà chuyển vấn đề sang chỗ khó hơn. Một lược đồ do người soạn có tác giả, có phiên bản, có thể tranh luận và có thể sửa. Một tập phạm trù nổi lên từ quá trình huấn luyện không có ba thứ đó và nó phản ánh phân bố của dữ liệu huấn luyện, tức là phản ánh việc ai đã viết nhiều nhất về cái gì.

Chuyển từ lược đồ do người soạn sang phạm trù tự sinh không phải chuyển từ thiên lệch sang trung lập. Đó là chuyển từ thiên lệch tranh luận được sang thiên lệch không tranh luận được.

**Phản bác sáu: mô hình mở và trí nhớ cá nhân hóa**

Nếu mô hình có trọng số mở, ai cũng chạy được và nếu mỗi người dùng có một trợ lý ghi nhớ sở thích riêng của mình, thì sẽ không còn một bộ lọc tập trung nào để phải vượt qua. Thị trường sẽ phân mảnh trở lại và quyền lực sẽ tán đi. Đây là phản bác mà cuốn sách mong là đúng và cũng là phản bác có ít bằng chứng ủng hộ nhất cho tới nay.

Lý do nằm ở chỗ trọng số mở không đi kèm dữ liệu mở. Một mô hình ai cũng chạy được vẫn phải lấy dữ liệu về thế giới từ đâu đó và mục 2.3 đã cho thấy nguồn dữ liệu ấy tập trung tới mức nào. Phân mảnh ở tầng mô hình cộng với tập trung ở tầng dữ liệu không cho ra một thị trường phân tán. Nó cho ra nhiều giao diện khác nhau cùng đọc chung một nguồn, tức là ảo giác về sự đa dạng.

Trí nhớ cá nhân hóa thì lại là chuyện khác và nó có thể thay đổi thật. Khi một tác nhân biết rõ chủ của nó từng thích những gì, trọng số của bằng chứng chung giảm xuống và trọng số của lịch sử riêng tăng lên. Với một địa phương, điều này mở ra một đường khác, thay vì cạnh tranh để được xếp hạng cao trong một

bảng chung, cạnh tranh để lọt vào ký ức của những người đã từng đến, vì chính ký ức ấy sẽ được tác nhân của họ dùng làm căn cứ. Nếu xu hướng này thắng, lợi thế sẽ quay trở lại với chiều sâu trải nghiệm, tức là quay về đúng luận điểm của "Địa phương trong nền kinh tế trải nghiệm". Đây là kịch bản duy nhất trong sáu kịch bản mà hai cuốn sách của bộ này hòa giải được với nhau và cuốn này không dám đặt cược vào nó.

Tổng kết sáu phản bác: một phản bác bị bác bỏ bằng lập luận kinh tế, hai phản bác bị bác bỏ bằng phân biệt giữa suy luận và giao dịch, hai phản bác thực ra củng cố mô hình và một phản bác chưa ngã ngũ. Người đọc có quyền cho rằng tỷ lệ ấy quá thuận lợi cho tác giả. Cách kiểm tra tốt nhất là theo dõi phản bác thứ ba và thứ sáu trong vài năm tới, vì đó là hai chỗ mà mô hình này sẽ vỡ nếu nó phải vỡ.

### 2.6.7. Ba phản bác có thể làm suy yếu luận đề

Mười một phản biện và phản bác trên có một đặc điểm mà người đọc cẩn thận sẽ nhận ra và nên nghi ngờ, phần lớn kết thúc bằng việc củng cố mô hình. Một mô hình hấp thụ được mọi dữ kiện trái chiều thì không còn là mô hình, nó là một cách nhìn không thể sai. Ba phản bác dưới đây khác ở chỗ nếu chúng đúng thì luận đề của cuốn sách yếu đi thật và tác giả không có câu trả lời trọn vẹn cho cả ba.

***Phản bác thứ nhất:*** giá trị kinh tế của khả đọc có thể rất nhỏ. Khả đọc chỉ là điều kiện để được đưa vào so sánh. Nếu một địa phương thiếu đường bay thuận lợi, giá cao, sản phẩm yếu, dịch vụ không đáng tin hoặc chính sách nhập cảnh bất lợi, thì việc trở nên khả đọc chỉ khiến nó bị loại nhanh hơn và chính xác hơn.

Trong trường hợp ấy, chi phí khả đọc là một khoản đầu tư có tỷ suất gần bằng không và toàn bộ chương trình mà cuốn sách đề nghị sẽ là một cách phân bổ ngân sách sai. Cuốn sách chỉ có thể trả lời rằng khả đọc là điều kiện cần chứ không phải điều kiện đủ, nhưng đó là một câu trả lời không giúp gì cho việc quyết định có nên chi tiền hay không.

***Phản bác thứ hai:*** các nền tảng có thể tự giải quyết phần lớn vấn đề. Nếu các công cụ tìm kiếm, nền tảng đặt dịch vụ và dự án bản đồ mở đã tự hợp nhất thực thể, tự thu tồn kho và tự chuẩn hóa dữ liệu vì lợi ích thương mại của chính họ, thì vai trò của chính quyền địa phương nhỏ hơn nhiều so với mức cuốn sách giả định. Lịch sử đứng về phía phản bác này, phần lớn hạ tầng khả đọc hiện có của ngành du lịch đều do bên tư nhân dựng, không do bên công. Câu trả lời khả dĩ là nền tảng chỉ chuẩn hóa phần sinh ra giao dịch cho họ và bỏ lại phần không sinh giao dịch, nhưng đây là một dự đoán chưa được kiểm chứng.

***Phản bác thứ ba:*** người dùng có thể phản ứng ngược với ủy nhiệm. Trong các quyết định giàu bản sắc, quá trình tự khám phá là một phần của giá trị chứ không phải chi phí cần cắt. Nếu điều đó đúng ở quy mô lớn, thị trường sẽ tách đôi lâu dài, phần giao dịch tiện ích bị tác nhân hóa gần như hoàn toàn, phần khám phá có ý nghĩa vẫn do con người thực hiện và có thể còn tăng giá trị chính vì sự khan hiếm của nó. Đây không phải một ngoại lệ nhỏ. Đó có thể là cấu trúc bền của thị trường và nếu vậy thì phần quan trọng nhất của một địa phương lại chính là phần cuốn sách này nói ít nhất.

Ba phản bác trên chưa ngã ngũ và cuốn sách không giả vờ đã trả lời được chúng. Chúng được đặt ở đây để người đọc có căn cứ không đồng ý và để mô hình có thể sai chứ không chỉ có thể giải thích.

Phản bác thứ nhất có một hệ quả thực hành mà cuốn sách nợ người đọc, một tiêu chí để biết nên đầu tư vào khả đọc trước hay vào chất lượng thật trước. Có thể trả lời bằng bốn ô. Chất lượng thật thấp và khả đọc thấp thì sửa sản phẩm trước, vì khả đọc lúc này chỉ làm cho việc bị loại diễn ra nhanh hơn. Chất lượng thật cao và khả đọc thấp là ô duy nhất nên đầu tư khả đọc ngay và đây cũng là ô mà phần lớn các địa phương của Việt Nam đang ở. Chất lượng thật thấp và khả đọc cao là ô nguy hiểm nhất, nên tạm dừng quảng bá mạnh và sửa thực chất, vì khả đọc đang khuếch đại một điểm yếu. Chất lượng thật cao và khả đọc cao thì chuyển trọng tâm sang khả hành động và chủ quyền.

Phản bác thứ hai cũng cần một câu trả lời cụ thể hơn, có những trường hợp chính quyền không nên tự xây. Có ba trường hợp như vậy. Khi nền tảng đã cung cấp dữ liệu đủ mở và đủ khả chuyển. Khi không có thất bại thị trường rõ ràng. Khi chi phí duy trì vượt lợi ích công. Vai trò công chỉ nên mở rộng khi có khoảng trống lợi ích công, có rủi ro bị khóa vào một nền tảng, có loại dữ liệu không sinh giao dịch nhưng có giá trị xã hội, có nhu cầu xác lập thẩm quyền pháp lý, hoặc có nhu cầu liên thông giữa nhiều nền tảng mà không bên tư nhân nào có động cơ làm.

Phản bác thứ ba, nếu đúng, không dẫn tới việc luận đề sai mà dẫn tới việc thị trường phân tầng. Giao dịch tiện ích sẽ bị tác nhân hóa gần như hoàn toàn. Trải nghiệm biểu tượng vẫn do con người chủ động tìm. Du lịch đại trà ủy nhiệm cao, du lịch ngách và du lịch sáng tạo ủy nhiệm thấp. Khách đến lần đầu phụ thuộc vào đề xuất của máy, khách quay lại phụ thuộc vào ký ức cá nhân của chính họ. Một địa phương cần biết mình bán cho phân khúc nào trước khi quyết định đầu tư bao nhiêu vào khả đọc.

### ĐIỀU NÀY CÓ Ý NGHĨA GÌ ĐỐI VỚI MỘT ĐỊA PHƯƠNG?

*Trước khi chi tiền cho khả đọc, cần xác định mình đang ở ô nào trong bốn ô. Chất lượng thật cao và khả đọc thấp là ô duy nhất nên đầu tư ngay và đó cũng là ô mà phần lớn địa phương Việt Nam đang ở. Chất lượng thật thấp và khả đọc cao là ô nguy hiểm nhất, vì khả đọc chỉ khuếch đại một điểm yếu. Sau đó mới tới câu hỏi về phân khúc, bởi mức ủy nhiệm của khách đại trà và khách ngách rất khác nhau.*

## 2.7. NGHỊCH LÝ SCOTT VÀ CÁI GIÁ CỦA MỘT TRẢI NGHIỆM KHẢ ĐỌC

### 2.7.1. Mỗi lần được đọc là một lần bị làm phẳng

Bây giờ là lúc trả món nợ mà chương trước đã vay. Nếu Scott đúng rằng mọi phép khả đọc hóa đều có giá, thì địa phương đi theo mô hình này phải trả giá gì. Gọi đó là nghịch lý Scott của thương hiệu địa phương. Mỗi lần một trải nghiệm được làm cho khả đọc, nó mất một phần thứ bậc ý nghĩa vốn làm nên giá trị của nó.

Cơ chế của mất mát ấy không bí ẩn. Nó là ba bước rất bình thường. Bước một, để vào lược đồ, trải nghiệm phải được tách thành các thuộc tính đo được. Bước hai, các thuộc tính đo được ấy là những gì máy dùng để so sánh, nên chúng là những gì được thưởng. Bước ba, cái gì được thưởng thì được tối ưu và tối ưu lâu thì trở thành bản chất.

*Bà cụ trong hẻm, bảy năm sau:* Một quán ăn trong hẻm, không biển hiệu, chỉ bán buổi sáng, món ăn nấu theo cách mà người nấu học từ mẹ mình. Người tìm đến đó thường được một người bản địa dẫn đi và phần lớn giá trị của trải nghiệm nằm ở chính việc phải được dẫn. *Năm thứ nhất*, quán được đưa vào một bộ dữ liệu chuẩn hóa. Nó có tọa độ, giờ mở cửa, khoảng giá, ảnh, hạng mục ẩm thực đường phố. Khách tăng. *Năm thứ ba*, để giữ điểm đánh giá, quán phải mở cửa đúng giờ đã ghi kể cả những hôm trước đây nghỉ. Phải nhận thanh toán điện tử. Phải có thực đơn tiếng Anh, mà thực đơn tiếng Anh buộc phải đặt tên cho những món trước đây không có tên. *Năm thứ năm*, quán thuê thêm người. Người mới học công thức trong hai tuần thay vì hai mươi năm. *Năm thứ bảy*, quán vẫn đông. Nó vẫn đúng mọi thuộc tính trong cơ sở dữ liệu. Nó không còn là lý do người ta từng tìm

đến nó và điều đáng chú ý là không ai trong chuỗi này làm gì sai. Mỗi bước đều là một quyết định hợp lý của một người đang cố giữ cho quán sống.

Ngành du lịch có sẵn từ cho giai đoạn cuối của quá trình này. Từ ấy là mất chất. Mô hình bốn bộ lọc cho phép nói chính xác hơn, mất chất là khi vòng phản hồi của khả đọc chạy mà không có van.

### 2.7.2. Khả đọc làm nghèo trải nghiệm như thế nào

Cần tách ba loại mất mát khác nhau, vì chúng đòi ba cách xử lý khác nhau và việc gộp chúng lại là lý do phần lớn các cuộc thảo luận về chủ đề này đi vào ngõ cụt:

- ***Mất mát do nén:*** Phần không đưa được vào lược đồ thì không tồn tại trước máy. Loại này không tránh được, chỉ quản trị được và công cụ quản trị là cái van.
- ***Mất mát do đồng dạng hóa:*** Khi mọi nơi đều mô tả mình theo cùng một bộ thuộc tính, các nơi bắt đầu giống nhau trong biểu diễn trước cả khi giống nhau trong thực tế. Loại này giảm được bằng cách mở rộng lược đồ và đó là một cuộc vận động chính sách chứ không phải một dự án kỹ thuật.
- ***Mất mát do phản hồi:*** Cái được thưởng thì được tối ưu và sau đủ nhiều vòng, thực tế bị kéo về phía biểu diễn của nó. Loại này nguy hiểm nhất vì nó không cần ai đồng ý và cũng khó nhất vì nó chỉ hiện ra sau nhiều năm.

Loại thứ ba là loại mà cuốn sách này lo nhất và cũng là loại mà không địa phương nào có công cụ đo. Không ai đo được tốc độ một cộng đồng đang trở nên giống bản mô tả của chính mình.

### 2.7.3. Nghịch lý trung tâm của cuốn sách

Hai lực kéo đã hiện đủ và cần được phát biểu ở dạng cuối cùng của nó. Nền kinh tế trải nghiệm đòi chiều sâu, nền kinh tế tác nhân đòi khả đọc, còn địa phương phải sống với cả hai. Nghịch lý này không có lời giải triệt để. Tác giả nhắc lại điều này lần thứ hai vì kinh nghiệm cho thấy người đọc thường bỏ qua nó ở lần thứ nhất và vì bất kỳ ai hứa với một tỉnh rằng chỉ cần chuẩn hóa dữ liệu là vừa giữ được bản sắc vừa thắng trong hệ thống suy luận đều đang bán một thứ không tồn tại.

Không có lời giải triệt để khác với không có gì để làm. Cái có thể làm là bốn việc: quyết định có ý thức phần nào đi qua bộ lọc, giữ tốc độ thay đổi chậm hơn tốc độ mà cộng đồng thích nghi được, đo cả những thứ không sinh ra doanh thu để chúng còn tồn tại trong báo cáo và đặt quyền quyết định về cái van ở nơi có thẩm quyền chứ không ở nơi có thẩm quyền kỹ thuật.

### 2.7.4. Chủ quyền thuật toán

Việc thứ tư trong danh sách trên dẫn tới khái niệm cuối cùng của chương. Chủ quyền thuật toán (algorithmic sovereignty) là quyền và năng lực của một cộng đồng trong việc định nghĩa cách mình được biểu diễn trước các hệ thống suy luận, mang biểu diễn ấy đi nơi khác và liên thông nó với các hệ thống khác mà không phụ thuộc vào một bên trung gian duy nhất.

Ba mặt trong định nghĩa tương ứng với ba câu hỏi khác nhau. Biểu diễn hỏi ai viết bản mô tả về chúng ta. Khả chuyển hỏi nếu chúng ta rời khỏi một nền tảng, chúng ta mang theo được gì. Khả liên thông hỏi bản mô tả ấy có được hệ thống khác đọc hiểu không hay phải làm lại từ đầu.

Điều nguy hiểm không phải là thuật toán thay đổi. Điều nguy hiểm là nhiều địa phương hành xử như thể nó sẽ không bao giờ thay đổi. Một vị thế phải mất nhiều năm mới xây dựng được có thể suy giảm chỉ sau một lần cập nhật mô hình hoặc điều chỉnh cơ chế xếp hạng. Khi khả năng được tìm thấy phụ thuộc vào hạ tầng của người khác, khả đọc không còn là một tài sản sở hữu mà chỉ là một lợi thế được duy trì tạm thời. Chủ quyền không có nghĩa là miễn nhiễm trước biến động. Chủ quyền là khả năng tiếp tục tồn tại sau biến động, nhờ dữ liệu chuẩn hóa, thương hiệu độc lập và sự hiện diện trên nhiều hệ sinh thái thay vì phụ thuộc vào một cổng truy cập duy nhất.

Chủ quyền thuật toán, như phát biểu ở trên, còn thiếu một mặt thứ tư và nó là mặt duy nhất có thể kiểm chứng được từ bên ngoài. Một cộng đồng tuyên bố có quyền định nghĩa hình ảnh của mình mà không có kênh nào để khiếu nại khi hình ảnh ấy bị dựng sai thì quyền đó chưa tồn tại trên thực tế. Ba mặt biểu diễn, khả chuyển và khả liên thông đều là năng lực kỹ thuật. Kênh khiếu nại là một thiết chế và hiện chưa có ai xây nó, không phải vì khó về kỹ thuật mà vì chưa rõ ai có thẩm quyền tiếp nhận. Chương cuối của cuốn sách sẽ lập luận rằng đây là khoảng trống chính sách cấp bách nhất mà nền kinh tế tác nhân tạo ra và cũng là khoảng trống ít được bàn nhất. Chủ quyền không được đo bằng những gì ta tuyên bố. Nó được đo bằng những gì xảy ra khi ta bị mô tả sai.

### ĐIỀU NÀY CÓ Ý NGHĨA GÌ ĐỐI VỚI MỘT ĐỊA PHƯƠNG?

*Chủ quyền thuật toán không đo bằng tuyên bố mà đo bằng ba năng lực kỹ thuật cộng một thiết chế. Ba năng lực là quyền biểu diễn, khả chuyển và khả liên thông. Thiết chế là kênh khiếu nại khi bản mô tả bị dựng sai. Một địa phương có thể tự kiểm tra rất nhanh bằng ba câu hỏi. Ai đang viết câu mô tả về chúng ta. Nếu rời khỏi một nền tảng, chúng ta mang theo được gì. Khi máy mô tả sai, chúng ta gửi văn bản cho ai và trong bao lâu thì được trả lời.*

## 2.8. BẢN ĐỒ PHẦN CÒN LẠI CỦA CUỐN SÁCH

Chương này dựng xong bộ khung, một khái niệm có định nghĩa kiểm tra được, một mô hình bốn bộ lọc kèm một vòng và một van, một mệnh đề về cách thế giới đáp lại việc bị đọc, mười bốn phản bác đã được đối chất, trong đó ba phản bác cuối chưa ngã ngũ và một khái niệm chuẩn tắc để các chương sau dựa vào.

### 2.8.1. Chương 3 trả lời câu hỏi máy đọc cái gì

Chương này mở nắp cỗ máy. Nó trình bày bốn chức năng tối thiểu mà mọi hệ thống suy luận hữu hạn phải thực hiện dưới hình thức này hay hình thức khác, tức là chọn lọc thông tin khả dụng, biểu diễn, xếp hạng và ủy nhiệm, theo nguyên lý chứ không theo tên công nghệ. Lý do của lựa chọn ấy vì tên các chuẩn kỹ thuật hiện hành nhiều khả năng sẽ lỗi thời trước khi cuốn sách này được tái bản, trong khi bốn chức năng tối thiểu thì không, vì chúng là bốn khâu không nén được của bất kỳ quy trình lựa chọn được ủy nhiệm nào.

### 2.8.2. Chương 4 và 5 trả lời câu hỏi làm thế nào để khả đọc

Chương 4 chuyển từ mô hình sang thực hành. Nó triển khai chi tiết bốn bộ lọc thành các bước công việc cụ thể, với thứ tự ưu tiên, ước tính nguồn lực và các dấu hiệu để biết mình đã qua tầng nào trong việc xây dựng niềm tin và ký ức của máy.

Chương 5 xử lý câu hỏi khó hơn và là câu hỏi mà cả bộ sách tồn tại vì nó, làm sao vừa khả đọc vừa giàu trải nghiệm, cuộc tranh giành quyền được đề xuất diễn ra như thế nào. Đây là chương của

cái van, nơi nghịch lý Scott được đưa từ tầng nhận thức xuống tầng quy trình, với các cơ chế tham vấn cộng đồng và các phép thử để phân biệt quyền mờ chính đáng với việc né tránh trách nhiệm giải trình.

### 2.8.3. Chương 6 tới 9 chuyển sang tầng quản trị

Từ đây, khả đọc thôi là một khái niệm và trở thành nền tảng của năng lực cạnh tranh.

Chương 6 bàn về kinh tế khả đọc. Như đường sá, điện và internet từng là điều kiện phát triển của các nền kinh tế trước, hạ tầng tri thức là điều kiện phát triển của nền kinh tế tác nhân, với đầy đủ những đặc tính của hạ tầng với chi phí trả trước, lợi ích trả sau và chi phí duy trì khó bảo vệ hơn vốn đầu tư ban đầu.

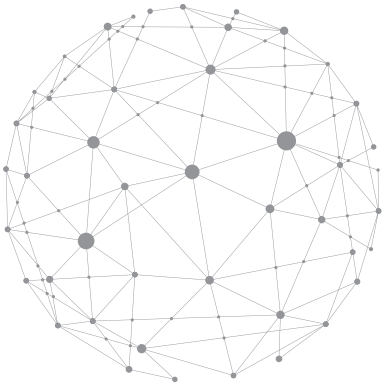
Chương 7 thiết kế địa phương cho máy, trong đó khả năng được máy nhận diện, tin cậy và ưu tiên trở thành một cấu phần ngang hàng với hạ tầng vật chất và chất lượng thể chế.

Chương 8 bàn về thương hiệu địa phương của Việt Nam trong nền kinh tế tác nhân với các đặc điểm riêng về dấu thanh, địa danh và năng lực cấp xã.

Chương 9 mở rộng lên cấp quốc gia, nơi chủ quyền và trách nhiệm về dữ liệu, tri thức và thuật toán trở thành các thành tố không tách rời của năng lực cạnh tranh và nơi khoảng trống về kênh khiếu nại.

### 2.8.4. Cách đọc phần còn lại

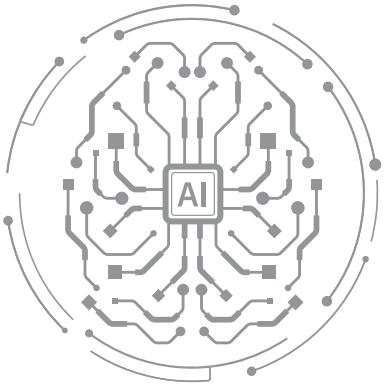
Người đọc vội có thể đi thẳng từ đây tới chương 6 bàn về kinh tế tác nhân. Người đọc kỹ nên đi tuần tự, vì các chương giữa là nơi bộ khung bị thử lửa chứ không phải phần minh họa cho nó và một mô hình chưa bị thử lửa thì chưa đáng tin. Dù đi đường nào, xin giữ theo mình ba mệnh đề. Khả đọc là điều kiện tồn tại. Điều kiện tồn tại chưa bao giờ đồng nghĩa với lý do đáng sống. Và câu hỏi thật sự không phải máy đang đọc chúng ta thế nào, mà chúng ta đang tự viết lại mình thành cái gì để được đọc.



CHƯƠNG 3

# MÁY ĐỌC ĐỊA PHƯƠNG BẰNG KIẾN TRÚC GÌ?

*Bốn ràng buộc của mọi hệ thống suy luận hữu hạn và bốn thao tác bất biến sinh ra từ chúng. Chương theo một minh hoạ duy nhất, qua bốn khâu để kiểm tra điều kiện khả đọc thuật toán.*


AI

Chương này không mô tả cách trí tuệ nhân tạo hoạt động. Nó lập luận cho một mệnh đề về mọi hệ thống suy luận hữu hạn, bất kỳ cỗ máy nào phải hành động trong một thế giới lớn hơn bộ nhớ của nó đều buộc phải truy hồi, biểu diễn và xếp hạng, rồi sẽ ủy nhiệm ngay khi chi phí giao dịch cho phép. Bốn thao tác ấy không phải đặc điểm của công nghệ hôm nay. Chúng là hệ quả của bốn ràng buộc không ai gỡ được. Địa phương không cạnh tranh với công nghệ. Nó cạnh tranh với bốn ràng buộc ấy.

Tháng Ba năm 2026, một gia đình ở Lyon giao cho trợ lý số một nhiệm vụ: dựng một kế hoạch kỳ nghỉ mười ngày ở Việt Nam cho hai người lớn và một đứa trẻ chín tuổi, ngân sách bốn nghìn euro, ưu tiên lịch sử và ẩm thực, tránh đi bộ nhiều dưới nắng, đặt trọn phòng và vé trước khi bay. Pháp là thị trường quốc tế lớn nhất của Huế. Gia đình này nằm đúng giữa tệp khách mà thành phố đang theo đuổi.[1]

1. Giả định của tác giả. Gia đình ở Lyon là nhân vật hư cấu làm sợi chỉ dẫn dắt Chương 3. Nhiệm vụ, ngân sách và ràng buộc trong đoạn này là giả định. Riêng mệnh đề về thị trường Pháp lấy theo Sở Du lịch thành phố Huế, vốn xếp Pháp trong nhóm thị trường truyền thống cùng Mỹ, Đức, Anh và Úc.

Chương này theo dõi duy nhất một câu hỏi, chuyện gì xảy ra với Huế trong mười giây tiếp theo, khi cỗ máy bắt đầu làm việc. Bốn mục lớn tương ứng bốn thao tác mà nó buộc phải đi qua. Ở mỗi thao tác, Huế được đưa lên bàn mổ một lần. Cuối chương, bản hành trình mà cỗ máy thực sự trả về sẽ được đặt cạnh bản hành trình mà Huế xứng đáng có và khoảng cách giữa hai bản ấy là toàn bộ nội dung của cuốn sách.

Huế không được chọn vì yếu. Sáu tháng đầu năm 2026, thành phố đón hơn 4,3 triệu lượt khách, doanh thu du lịch đạt 10.300 tỷ đồng, tăng 56,2% so với cùng kỳ.[1] Cổng thông tin du lịch chạy năm ngôn ngữ. Huế được chọn chính vì mạnh. Nếu một địa phương thành công theo mọi thước đo truyền thống vẫn thất bại ở một số khâu của chuỗi suy luận, thì vấn đề nằm ở cấu trúc chứ không nằm ở nỗ lực.

Chương 2 mô tả bốn bộ lọc mà địa phương phải đi qua: phân giải, chứng thực, xếp hạng, thực thi. Chương này mô tả người gác cửa và chứng minh vì sao người gác cửa ấy không thể làm khác.

1. Sở Du lịch thành phố Huế, số liệu 6 tháng đầu năm 2026, dẫn theo Báo Nhân Dân ngày 27 tháng 6 năm 2026, https://nhandan.vn/du-lich-hue-buoc-tien-moi-trong-ky-nguyen-so-post971847.html. Khách quốc tế khoảng 1,5 triệu lượt, tăng 25,1%, khách nội địa gần 2,8 triệu lượt, tăng 30,1%. Số ước tính theo phương pháp thống kê du lịch, không so sánh trực tiếp được với số lượt lưu trú của Kyoto.

## 3.1. BỐN RÀNG BUỘC SINH RA BỐN BẤT BIẾN

### 3.1.1. Bốn ràng buộc

Mọi lập luận trong chương đều quy về bốn ràng buộc. Chúng không thuộc về trí tuệ nhân tạo. Chúng thuộc về điều kiện tồn tại của bất kỳ hệ thống nào phải quyết định trong một thế giới lớn hơn chính nó.

Ràng buộc thứ nhất là giới hạn bộ nhớ. Không hệ thống nào giữ đồng thời toàn bộ thông tin liên quan. Ràng buộc thứ hai là yêu cầu đồng đo. Hai đối tượng chỉ so sánh được khi cùng nằm trên một hệ đo. Ràng buộc thứ ba gồm hai mặt của cùng một vấn đề, giới hạn chú ý và bất cân xứng thông tin với kênh đầu ra luôn hẹp hơn tập ứng viên, còn chất lượng thật thì không quan sát trực tiếp được. Ràng buộc thứ tư là chi phí giao dịch. Việc giao một nhiệm vụ cho bên khác chỉ xảy ra khi chi phí mô tả, giám sát và khắc phục sai sót đủ thấp.

Từ bốn ràng buộc ấy suy ra bốn thao tác, một đối một. Giới hạn bộ nhớ buộc hệ thống chọn tập con, đó là truy hồi. Yêu cầu đồng đo buộc hệ thống chuyển tập con ấy về dạng chung, đó là biểu diễn. Giới hạn chú ý cùng bất cân xứng thông tin buộc hệ thống áp một trật tự, đó là xếp hạng. Chi phí giao dịch quyết định hệ thống dừng lại ở khuyến nghị hay đi tiếp tới hành động, đó là ủy nhiệm.

Ba ràng buộc đầu là bắt buộc. Ràng buộc thứ tư thì không. Sự khác hạng ấy cần được giữ nguyên chứ không được che đi, bởi một khung giải thích được mọi thứ thì không giải thích được gì. Bốn mục lớn còn lại của chương triển khai đúng bốn ràng buộc này, mỗi mục một ràng buộc.

### 3.1.2. Vòng sơ loại bằng phản ví dụ

Một mệnh đề phổ quát chỉ đáng tin khi đã sống sót qua những nỗ lực bác bỏ nó. Dưới đây là bốn ứng viên phản ví dụ mạnh nhất.

- ***Ứng viên thứ nhất chống lại truy hồi:*** Một mô hình chứa toàn bộ tri thức trong trọng số, không cần tra cứu bên ngoài. Mô hình ấy tồn tại. Nhưng mỗi lần suy luận nó chỉ kích hoạt một phần rất nhỏ của trọng số. Kích hoạt có chọn lọc là một phép chọn tập con. Phản ví dụ thất bại, vì nó chỉ dời truy hồi từ ngoài vào trong.

- ***Ứng viên thứ hai chống lại biểu diễn:*** Một bộ điều nhiệt. Nó hành động mà không cần khái niệm nào về thế giới. Đúng, nhưng nó cũng không chọn giữa các phương án khác loại. Thêm cho nó nhiệm vụ chọn giữa bật lò sưởi, mở cửa sổ và nhắc chủ nhà mặc thêm áo, lập tức phải có một thang chung để ba phương án ấy so được với nhau. Phản ví dụ thất bại, vì nó chỉ đúng với hệ thống không phải lựa chọn.

- ***Ứng viên thứ ba chống lại xếp hạng:*** Một cơ sở dữ liệu trả về toàn bộ bản ghi khớp điều kiện, không sắp thứ tự. Hệ thống ấy có thật và rất phổ biến. Nó không xếp hạng vì kênh đầu ra không hẹp hơn tập kết quả và vì nó đẩy toàn bộ việc chọn lọc sang cho con người. Ngay khi một tác nhân phải trả lời thay con người, kênh đầu ra thu hẹp và trật tự xuất hiện trở lại. Phản ví dụ này không bị bác bỏ, nó chỉ ra điều kiện mà mệnh đề ba đúng: khi và chỉ khi đầu ra hẹp hơn tập ứng viên.

- ***Ứng viên thứ tư chống lại ủy nhiệm:*** Hầu hết hệ thống số của ba mươi năm qua. Chúng trả về danh sách và dừng. Phản ví dụ này thành công. Đó chính là lý do ủy nhiệm được xếp riêng, như một hệ quả kinh tế chứ không phải một hệ quả logic.

Ba mệnh đề đầu sống sót. Mệnh đề thứ tư có điều kiện. Bốn phép thử trên chỉ là vòng sơ loại. Mỗi mục lớn còn lại sẽ tự đưa mệnh đề của mình ra thử lại bằng những hệ thống có thật và trong hai trường hợp phản ví dụ sẽ thắng, buộc mệnh đề phải được phát biểu chặt hơn.

### 3.1.3. Quan hệ tích giữa bốn thao tác

Bốn thao tác nối tiếp nhau theo quan hệ nhân, không phải quan hệ cộng.

$$\mathbf{P}(\text{được chọn}) = \mathbf{P}(\text{truy hồi được}) \times \mathbf{P}(\text{biểu diễn đúng}) \times \mathbf{P}(\text{vượt ngưỡng}) \times \mathbf{P}(\text{thực thi được})$$

Trong một chuỗi cộng, một khâu xuất sắc bù cho một khâu yếu. Trong một chuỗi nhân, một khâu bằng không xóa sạch ba khâu còn lại. Một điểm đến có kho ảnh đẹp nhất nước nhưng không cho phép máy đặt vé nhận xác suất bằng không ở khâu cuối và do đó bằng không cho toàn chuỗi. Cơ hội của một địa phương bằng tích xác suất của bốn khâu. Đầu tư đúng không phải là làm tốt nhất một khâu mà là không để khâu nào bằng không. Mệnh đề này bác bỏ trực tiếp cách phân bổ ngân sách phổ biến hiện nay, vốn dồn vào khâu đã mạnh nhất vì khâu ấy dễ báo cáo nhất.

### 3.1.4. Ba làn sóng khả đọc

Bốn thao tác không sinh ra cùng máy tính. Mỗi khi một nền văn minh phải xử lý lượng tri thức lớn hơn trí nhớ một người, nó phát minh lại đúng bốn thao tác ấy bằng vật liệu của thời đại mình.

Thư viện Alexandria không thiếu sách, nó thiếu khả năng tìm. Callimachus giải bài toán bằng bộ danh mục Pinakes vào thế kỷ III trước Công nguyên.[1] Từ đó, một cuốn sách không có trong danh mục không tồn tại đối với người tra cứu, dù vẫn nằm trên giá. Linnaeus (năm 1735) quy hàng vạn sinh vật về một hệ danh pháp và chỉ khi ấy hai nhà tự nhiên học ở hai châu lục mới biết chắc họ đang bàn về cùng một loài.[2] Dewey (năm 1876) ép toàn bộ tri thức vào mười lớp số.[3] Shepard (năm 1873) rồi Garfield (năm 1955) biến trích dẫn thành phiếu tín nhiệm, bốn thập niên trước khi nguyên lý ấy được đem áp cho các trang web dưới một cái tên khác.[4]

Việt Nam có phiên bản của mình và nó rất gần. Từ năm 1805, triều Nguyễn lập địa bạ trên quy mô toàn quốc, đo đạc từng thửa ruộng theo hạng đất, chủ sở hữu, ranh giới, song song với sổ

1. Callimachus xứ Cyrene, bộ Pinakes, Thư viện Alexandria, thế kỷ III trước Công nguyên. Nguyên bản thất truyền, nội dung chỉ được dựng lại qua các trích dẫn về sau. Xem Rudolf Blum, Kallimachos: The Alexandrian Library and the Origins of Bibliography, bản dịch của Hans H. Wellisch, University of Wisconsin Press, Madison, 1991.

2. Carl Linnaeus, Systema Naturae, Theodorus Haak, Leiden, 1735. Bản in đầu chỉ gồm mười hai trang khổ lớn. Hệ danh pháp hai phần cho động vật mà giới phân loại học lấy làm mốc khởi đầu nằm ở lần in thứ mười, Laurentius Salvius, Stockholm, 1758.

3. Melvil Dewey, A Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library, Amherst, Massachusetts, 1876.

4. Frank Shepard khởi lập bộ Shepard's Citations tại Chicago năm 1873. Eugene Garfield, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas", Science, tập 122, số 3159, 1955, trang 108 tới 111. Science Citation Index bắt đầu xuất bản năm 1964.

đinh ghi nhân khẩu chịu thuế.[1] Trước địa bạ, làng là một cộng đồng có ký ức, tập quán và những thỏa thuận bất thành văn về ruộng công. Sau địa bạ, làng là một tập hợp thửa đất có mã số. Nhà nước thu được nhiều hơn. Những quyền sử dụng theo mùa vụ không lọt vào ô của cuốn sổ thì mất chỗ đứng pháp lý dù vẫn tồn tại trong đời sống.

James C. Scott gọi đó là "legibility", tính khả đọc của xã hội đối với nhà nước và rút ra hai vế mà chương này kế thừa nguyên vẹn. Muốn quản trị một thứ thì trước hết phải làm cho nó khả đọc. Và mọi phép làm cho khả đọc đều là một phép giản lược, trong đó phần bị lược bỏ không biến mất khỏi thực tại mà chỉ biến mất khỏi tầm nhìn của người quản trị.

Tính khả đọc thuật toán là làn sóng thứ ba. Thứ nhất với nhà nước, kéo dài nhiều thế kỷ. Thứ hai với thị trường, khi hàng hóa phải có mã vạch, nhãn và chứng nhận. Thứ ba với máy. Ba làn sóng cùng cấu trúc, khác nhau ở ba điểm và cả ba đều bất lợi cho địa phương, bởi lược đồ thay đổi theo tháng chứ không theo thế kỷ, một sai sót được nhân bản sang mọi hệ thống kế thừa và không còn cơ quan nào nhận đơn khiếu nại.

Nhà nước từng đọc làng bằng địa bạ. Thị trường đọc sản phẩm bằng mã vạch. Máy đọc địa phương bằng thực thể. Ba lần cùng một câu chuyện và lần nào phần không lọt vào lược đồ cũng là phần trả giá.

1. Về công cuộc lập địa bạ triều Nguyễn, xem Nguyễn Đình Đầu, bộ Nghiên cứu địa bạ triều Nguyễn, Nhà xuất bản Thành phố Hồ Chí Minh, từ năm 1994, và Phan Huy Lê chủ biên, bộ Hệ thống tư liệu địa bạ Việt Nam, từ năm 1995. Khối địa bạ gốc bảo quản tại Trung tâm Lưu trữ quốc gia I. Công cuộc đạc điền cho khoảng 18.000 xã thôn kéo dài từ năm 1805 tới năm 1836, nên 1805 là mốc khởi đầu chứ không phải mốc hoàn tất.

***Nguyên tắc quản trị:*** Không đầu tư vào khâu mạnh nhất khi còn khâu bằng không. Chuỗi là phép nhân, nên ngân sách phải chảy về khâu yếu nhất. Không tối ưu cho một sản phẩm công nghệ, chỉ tối ưu cho ràng buộc sinh ra nó. Sản phẩm có vòng đời vài năm, ràng buộc có vòng đời nhiều thập niên.

Nếu là Chủ tịch tỉnh. Việc đầu tiên không phải là phê duyệt một đề án chuyển đổi số mà là yêu cầu trả lời bốn câu hỏi trong một trang giấy: hai mươi thực thể quan trọng nhất của tỉnh có được máy tìm thấy không, có được gán đúng loại không, có lọt vào tập ứng viên khi khách hàng mục tiêu hỏi trợ lý số không và có giao dịch nào hoàn tất được bằng máy không. Khâu nào trả lời không thì đó là nơi ngân sách phải đi tới, bất kể ba khâu còn lại đang tốt tới đâu.

## 3.2. RÀNG BUỘC THỨ NHẤT: BỘ NHỚ HỮU HẠN

Nếu tập thông tin liên quan lớn hơn dung lượng xử lý đồng thời, mọi hệ thống buộc phải chọn một tập con trước khi suy luận. Phép chọn ấy là truy hồi.

### 3.2.1. Chứng minh

Gọi tập thông tin liên quan tới một quyết định là W, dung lượng xử lý đồng thời là M. Với mọi hệ thống thực, W lớn hơn M. Suy luận chỉ diễn ra trên nội dung đang nằm trong M. Vậy tồn tại một ánh xạ chọn từ W vào M. Ánh xạ ấy chính là truy hồi. Không có nó, M rỗng và suy luận không khởi động được. Hệ quả trực tiếp là phần của W không được ánh xạ vào M không tham gia bất kỳ bước nào phía sau. Nó không bị đánh giá thấp. Nó không được đánh giá.

### 3.2.2. Phản ví dụ

Chỉ đáng tin nếu sống sót qua những hệ thống thật có vẻ vi phạm nó.

- ***Phản ví dụ 1A. Trợ lý của một khách sạn.*** Một khách sạn nạp trọn bốn mươi trang tài liệu nội bộ vào ngữ cảnh, không truy hồi gì thêm, vẫn trả lời chính xác mọi câu hỏi về chính nó. Ở đây W nhỏ hơn M thật, nên không tồn tại phép chọn tập con nào. Phản ví dụ thành công và nó khoanh đúng phạm vi của mệnh đề, truy hồi chỉ bắt buộc khi W lớn hơn M. Hệ quả cho địa phương lại đáng giá hơn cả phản ví dụ. Nếu một kho ngữ liệu đủ nhỏ và đủ sạch thì nó lọt trọn vào bộ nhớ làm việc và không còn chịu rủi

ro bị lọt. Xây một kho tối thiểu vài trăm thực thể cốt lõi, chuẩn hóa kỹ, có sức mạnh lớn hơn nhiều so với việc số hóa hàng vạn hồ sơ rời rạc.

• ***Phản ví dụ 1B. Ảo giác.*** Một mô hình mô tả tỉ mỉ một lễ hội không có thật, đầy đủ ngày tháng và nghi thức, mà không truy hồi bất kỳ nguồn nào. Đây là bằng chứng rõ ràng rằng hệ thống có thể sinh đầu ra mà bỏ qua bước truy hồi. Phản ví dụ thất bại và thất bại theo cách củng cố mệnh đề. Nó cho thấy điều gì xảy ra khi bước truy hồi vắng mặt, đầu ra mất liên hệ với thế giới. Ảo giác không phải ngoại lệ của mệnh đề 1 mà là minh chứng cho nó. Với địa phương, hệ quả rất trực tiếp. Nơi có ngữ liệu mỏng chính là nơi máy dễ bịa nhất, nên rủi ro không chỉ là bị bỏ sót mà còn là bị mô tả sai.

• ***Phản ví dụ 1C. Bộ nhớ tham số.*** Một mô hình rất lớn trả lời đúng về Huế mà không tra cứu gì, vì tri thức nằm sẵn trong trọng số. Phản ví dụ thất bại. Mỗi lần suy luận, mô hình chỉ kích hoạt một phần rất nhỏ trọng số của nó và kích hoạt có chọn lọc là một phép chọn tập con. Truy hồi bị dời từ ngoài vào trong chứ không biến mất. Điều thay đổi là ai kiểm soát nó khi truy hồi nằm ngoài, địa phương còn tác động được bằng dữ liệu công bố. Khi nó nằm trong trọng số, địa phương chỉ tác động được ở chu kỳ huấn luyện kế tiếp, tức chậm hơn nhiều năm.

### 3.2.3. Phản biện

• ***Phản bác thứ nhất.*** Cửa sổ ngữ cảnh của các mô hình hiện nay đã lên tới hàng trăm nghìn từ, đủ chứa trọn một hồ sơ di sản, nên bài toán chọn lọc mất ý nghĩa. Bác bỏ: mở rộng M không thay đổi quan hệ W lớn hơn M, vì W là toàn bộ tri thức có liên

quan trên thế giới chứ không phải một tài liệu. Quan trọng hơn, việc quyết định đưa hồ sơ nào vào cửa sổ ngữ cảnh lại là một phép chọn tập con thực hiện trước đó. Cửa sổ lớn hơn chỉ dời truy hồi lùi một nấc, không xóa nó.

• ***Phản bác thứ hai.*** Các trợ lý hiện đại tìm kiếm trực tuyến theo thời gian thực, nên không thực thể nào bị khóa vĩnh viễn ngoài bộ nhớ. Bác bỏ: quyền tra cứu không xóa bài toán, nó chuyển bài toán sang tiêu chí truy vấn. Tác nhân vẫn phải quyết định tra cứu gì, tin nguồn nào, dừng ở đâu. Một làng nghề không có tên chuẩn, không định danh, mô tả mâu thuẫn giữa các nguồn thì vẫn không xuất hiện. Cánh cửa rộng hơn không giúp ích cho người không có địa chỉ.

• ***Phản bác thứ ba, mạnh nhất.*** Emily Bender, Timnit Gebru và cộng sự lập luận rằng dữ liệu huấn luyện của các mô hình quy mô lớn phản ánh quá mức tiếng nói của những nhóm đã có quyền lực và rằng khuyên các cộng đồng chuyển tri thức của mình thành ngữ liệu là mời họ cung cấp nguyên liệu miễn phí cho chính hệ thống đang gạt họ ra.[1] Vế thứ nhất củng cố mệnh đề 1 chứ không bác bỏ nó. Vế thứ hai thì chương này không có lời bác bỏ hoàn chỉnh và cần nói rõ như vậy. Điều duy nhất có thể nói là lựa chọn không tham gia hiện không tồn tại, hồ sơ về một làng nghề vẫn được viết dù cộng đồng có tham gia hay không, chỉ khác ở chỗ ai viết. Câu hỏi đúng vì thế không phải tham gia hay không mà là tham gia với điều kiện gì.

1. Emily M. Bender, Timnit Gebru, Angelina McMillan-Major và Shmargaret Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, ACM, New York, 2021, trang 610 tới 623.

### 3.2.4. Hệ quả: dữ liệu không phải ngữ liệu

Giữa một tệp dữ liệu và một hệ thống suy luận có bốn tầng chuyển hóa và mỗi tầng giải một bài toán khác.

Dữ liệu ghi lại rằng thông tin tồn tại. Một bức ảnh chùa Thiên Mụ là một tập hợp điểm ảnh. Ngữ liệu trả lời câu hỏi đó là gì, cùng bức ảnh ấy trở thành ngữ liệu khi hệ thống biết đó là chùa Thiên Mụ ở Huế, khởi dựng năm 1601, thuộc loại hình di sản tôn giáo, có tọa độ xác định. Tri thức xuất hiện khi các mảnh ngữ liệu nối thành quan hệ có nghĩa. Chùa Thiên Mụ và Đại Nội cùng thuộc quần thể di tích cố đô, cùng gắn với triều Nguyễn, thường nằm trong cùng một hành trình.[1] Ký ức là phần tri thức lấy ra được đúng lúc. Một thư viện triệu cuốn không tạo nên ký ức tốt nếu không ai tìm được cuốn cần tìm.

Nhiều tỉnh tự hào về kho dữ liệu đồ sộ. Chưa chuẩn hóa, chưa định danh, chưa liên kết thì với máy đó là kho chứa, không phải kho ký ức.

Con số cho thấy khoảng cách này phổ biến đến mức nào. Trong đợt trích xuất tháng Mười năm 2024 của dự án Web Data Commons, dữ liệu có cấu trúc xuất hiện trên 16,5 triệu trong tổng số 37,5 triệu tên miền được thu thập, tức 44 phần trăm.[2] Sau hơn một thập niên phổ biến các chuẩn mở, hơn một nửa số tên miền vẫn chưa tự mô tả mình bằng ngôn ngữ máy đọc được. Khoảng trống ấy tập trung ở đúng những chủ thể ít nguồn lực nhất.

1. Bộ Văn hóa, Thể thao và Du lịch, công bố cuối năm 2024. Số kiểm kê tích lũy nhiều năm, mức độ cập nhật không đồng đều giữa các địa phương.

2. Web Data Commons, Đại học Mannheim, báo cáo trích xuất từ bản Common Crawl tháng 10 năm 2024, công bố ngày 10 tháng 1 năm 2025, https://webdatacommons.org/structureddata/2024-12/stats/stats.html. Dữ liệu có cấu trúc xuất hiện trên 16,5 triệu trong 37,5 triệu tên miền cấp thanh toán, tức 44,12%, và trên 51,25% số trang. Mẫu là kho Common Crawl chứ không phải toàn bộ web.

### 3.2.5. Điều không bao giờ vào được bộ nhớ

Có một phần của W không chỉ chưa được ánh xạ mà về nguyên tắc rất khó ánh xạ. Friedrich Hayek (1945) đã mô tả nó: phần lớn tri thức có ích không tồn tại ở dạng khái quát và thống kê mà tồn tại rải rác dưới dạng hiểu biết về hoàn cảnh cụ thể của một thời điểm và một địa điểm, thứ chỉ người tại chỗ nắm được.[1] Ông kết luận rằng không cơ quan kế hoạch trung ương nào thay thế được thị trường, không phải vì thiếu năng lực tính toán mà vì loại tri thức quyết định bị hủy hoại trên đường khái quát hóa.

Lập luận viết ra để chống kế hoạch hóa tập trung, tám mươi năm sau mô tả chính xác giới hạn của khả đọc thuật toán, lần này đối tượng không phải một ủy ban kế hoạch mà là một lớp hạ tầng suy luận toàn cầu. Người bán bún bò biết giờ nào khách đông, tháng nào mưa làm vắng khách, con hẻm nào ngập trước. Người trông coi một ngôi đình biết dòng họ nào giữ chìa khóa gian nào. Tri thức ấy không mất đi khi không được số hóa. Nó chỉ vô hình với hệ thống đang thay mặt hàng triệu người quyết định.

Quy mô của phần vô hình ở các địa phương của Việt Nam rất lớn. Cả nước đã kiểm kê hơn 40 nghìn di tích và gần 70 nghìn di sản văn hóa phi vật thể.[2] Nghị quyết 80-NQ/TW đặt mục tiêu số hóa toàn bộ di sản đã xếp hạng cấp quốc gia và cấp quốc gia đặc biệt.[3] Cam kết mạnh, nhưng số hóa không phải khả đọc. Một hồ sơ quét thành tệp ảnh vẫn nằm ngoài mọi chuỗi truy hồi

1. Friedrich A. Hayek, "The Use of Knowledge in Society", American Economic Review, tập 35, số 4, 1945, trang 519 tới 530.

2. Bộ Văn hóa, Thể thao và Du lịch, số liệu kiểm kê công bố cuối năm 2024. Số kiểm kê tích lũy nhiều năm, mức độ cập nhật giữa các địa phương không đồng đều và việc sáp nhập đơn vị hành chính cấp tỉnh năm 2025 làm thay đổi cách quy chiếu địa bàn.

3. Nghị quyết 80-NQ/TW của Bộ Chính trị, năm 2026. Đây là chỉ tiêu chính sách, không phải kết quả đã đạt.

nếu không có định danh bền, không mô tả cấu trúc, không liên kết. Chuyển từ giấy sang tệp là điều kiện cần. Chuyển từ tệp sang thực thể mới là điều kiện đủ.

### 3.2.6. Thí nghiệm với Huế, khâu thứ nhất

Trở lại gia đình ở Lyon. Ràng buộc lớn nhất họ đưa ra là tránh đi bộ nhiều dưới nắng, cùng một đứa trẻ chín tuổi. Cỗ máy cần bốn trường dữ liệu để xác nhận Huế đáp ứng được: quãng đường đi bộ trong khuôn viên mỗi điểm, tỷ lệ có bóng che, thời lượng tham quan trung bình, khả năng tiếp cận cho trẻ nhỏ. Không phải hồ sơ di tích nào của Việt Nam cũng chứa bốn trường ấy.

Huế không bị loại vì kém. Huế bị loại vì không có căn cứ để khẳng định mình không kém. Trong một hệ thống phải chịu trách nhiệm cho khuyến nghị của mình, thiếu bằng chứng được xử lý gần giống bằng chứng phủ định.

Cổng thông tin năm ngôn ngữ không giải quyết việc này, vì nó được thiết kế cho người đọc. Câu hỏi của khâu truy hồi là khác, cùng khối nội dung ấy có tồn tại ở dạng máy lấy được mà không cần đọc giao diện hay không, mỗi di tích có định danh ổn định hay không, lịch biểu diễn nhã nhạc tuần tới nằm trong một tệp có lược đồ hay nằm trong một tấm ảnh áp phích. Khoảng cách giữa hai câu hỏi ấy là khoảng cách giữa số hóa và khả đọc và nó không xuất hiện trong bất kỳ chỉ tiêu nào. Tổn thất loại này không để lại dấu vết. Không báo cáo nào ghi rằng một trăm nghìn du khách đã không đến vì tác nhân thiếu bốn trường dữ liệu. Không ai nhận được thông báo rằng mình vừa bị đưa ra khỏi một chuỗi suy luận.

Điều không thể truy hồi thì không thể suy luận. Với máy, thứ không tìm lại được tương đương với thứ chưa từng tồn tại. Chi phí của vô hình không phải là không được nhìn thấy mà là không bao giờ được cân nhắc. Tổn thất xảy ra trước khi có ai quan sát được nó và không để lại dấu vết trong bất kỳ báo cáo nào. Vì thế, nguyên tắc của quản trị là không quản trị thứ không có định danh. Mỗi thực thể trọng điểm cần một mã duy nhất, bất biến qua mọi lần đổi tên hay sáp nhập, kèm khai báo quan hệ kế thừa. Không công bố dữ liệu chỉ cho người đọc, mọi dữ liệu công khai phải có bản dành cho máy đọc kèm lược đồ và điều kiện sử dụng. Một quyết định chỉ tồn tại dưới dạng ảnh quét là một quyết định không tồn tại. Không coi im lặng là quyền mờ, phần tri thức không đưa vào lược đồ phải là một quyết định được ghi lại, có hội đồng và có lý do, không phải một thiếu sót được biện minh về sau.

Nếu là Giám đốc Sở Du lịch. Bỏ chỉ tiêu số tin bài và lượt truy cập cổng thông tin. Thay bằng ba con số: bao nhiêu % trong hai trăm thực thể trọng điểm có định danh bền, bao nhiêu % có dữ liệu cập nhật trong chín mươi ngày và bao nhiêu % có đủ các trường mô tả ràng buộc thực tế của khách như quãng đường đi bộ, bóng che, thời lượng, khả năng tiếp cận. Ba con số này rẻ để đo và giải thích được phần lớn chênh lệch giữa các điểm đến trong tập ứng viên của tác nhân.

## 3.3. RÀNG BUỘC THỨ HAI: YÊU CẦU ĐỒNG ĐO

Hai đối tượng chỉ so sánh được khi cùng nằm trên một hệ đo. Việc đưa các đối tượng khác loại về một hệ đo chung là biểu diễn. Không có biểu diễn thì không có so sánh và không có so sánh thì không có lựa chọn.

### 3.3.1. Chứng minh

So sánh là một quan hệ thứ tự trên một tập. Một quan hệ thứ tự đòi hỏi mọi phần tử được chiếu lên cùng một thang giá trị. Một bảng giá, một bức ảnh, một tọa độ và một đoạn đánh giá thuộc bốn loại khác nhau, không có thang chung sẵn có. Vậy trước khi so sánh phải tồn tại một phép chiếu đưa cả bốn về một không gian duy nhất. Phép chiếu ấy là biểu diễn.

Cơ quan thống kê thực hiện đúng phép chiếu này khi quy mọi hoạt động kinh tế về một đơn vị tiền tệ. Nhà sinh học thực hiện nó khi quy hàng vạn loài về một hệ danh pháp. Máy thực hiện nó khi quy thế giới về các véc-tơ. Ba hệ đo, một nguyên lý, một cái giá, thứ không chiếu được lên hệ đo thì không tham gia được phép so sánh.

Điều không thể biểu diễn thì không thể so sánh và điều không thể so sánh thì không thể được lựa chọn. Mọi phép chiếu lên hệ đo đều giữ lại một phần và bỏ đi phần còn lại.

### 3.3.2. Phản ví dụ

• ***Phản ví dụ 2A. So sánh trong cùng một loại.*** Một hệ thống xếp hai khách sạn theo giá không cần hệ đo nào ngoài đơn vị tiền tệ vốn đã có sẵn. Không có phép chiếu nào phải dựng. Phản ví dụ thành công và khoanh đúng phạm vi, đồng đo chỉ bắt buộc khi phải so các đối tượng khác loại hoặc so theo nhiều chiều không cùng đơn vị. Nhưng đó chính xác là việc mà một tác nhân lập kế hoạch luôn phải làm, vì nó phải cân một đêm ở Huế với một buổi lặn ở Nha Trang và một ngày ở Hà Nội trong cùng một ngân sách. Hệ quả cho địa phương ngược với trực giác thông thường. Địa phương an toàn khi bị so trong loại, dễ thua khi bị so liên loại và nền kinh tế tác nhân đẩy mọi thứ về phía so liên loại.

• ***Phản ví dụ 2B. Hệ thống học thẳng từ tín hiệu tới hành động.*** Một xe tự hành ánh xạ trực tiếp từ điểm ảnh sang góc lái, không có khái niệm nào được đặt tên, không có lược đồ nào do người thiết kế viết ra. Phản ví dụ thất bại, nhưng theo cách đáng lo hơn là nếu nó thành công. Biểu diễn vẫn tồn tại, chỉ là nó nằm ẩn trong các lớp trung gian và không ai đặt tên cho nó, kể cả người xây hệ thống. Khi biểu diễn tường minh, một địa phương bị gán sai danh mục còn biết mình bị gán sai. Khi biểu diễn hoàn toàn ẩn, sai lệch vẫn xảy ra mà không còn đối tượng nào để kiểm tra hay khiếu nại. Biểu diễn ẩn không xóa vấn đề, nó xóa khả năng phát hiện vấn đề.

• ***Phản ví dụ 2C. Con người chọn bằng trực giác.*** Một du khách chọn Huế thay vì Hội An mà không hề quy hai nơi về một thang điểm nào. Phản ví dụ thất bại ở chỗ nó không phản bác được mệnh đề, vì mệnh đề nói về hệ thống phải so sánh tường minh và cũng vì các nghiên cứu về quyết định cho thấy con người vẫn dựng một thang chung ngầm rồi rút gọn nó bằng kinh nghiệm.

Điều đáng chú ý nằm ở chỗ khác. Trực giác của người là thứ duy nhất có thể so sánh mà không cần đồng đo, nên khi quyền chọn chuyển từ người sang máy, đúng thứ vốn cứu được những nơi khó mô tả cũng biến mất.

### 3.3.3. Phản biện

• ***Phản bác thứ nhất.*** Yann LeCun lập luận rằng các mô hình ngôn ngữ tự hồi quy là ngõ cụt và trí tuệ thực sự đòi hỏi những mô hình thế giới dự đoán ở tầng biểu diễn trừu tượng.[1] Nếu ông đúng, kiến trúc thống trị hôm nay sẽ bị thay. Bác bỏ một phần, phản biện này đúng ở tầng sản phẩm và không chạm tới mệnh đề 2. Một mô hình thế giới vẫn phải mã hóa môi trường thành một không gian trừu tượng và chính LeCun nhấn mạnh bước ấy hơn cả các kiến trúc hiện hành. Phản biện của ông bác bỏ một họ kiến trúc, không bác bỏ ràng buộc sinh ra chúng.

• ***Phản bác thứ hai.*** Nếu mô hình đọc trọn được một hồ sơ dài thì nó đâu cần nén thành véc-tơ. Bác bỏ: việc chọn hồ sơ nào đáng đọc trọn vẹn lại dựa trên một biểu diễn nén thực hiện trước đó. Ngữ cảnh dài dời bước nén lùi một nấc, không xóa nó. Đây là cùng một cấu trúc phản bác đã gặp ở mệnh đề 1 và sự lặp lại ấy không ngẫu nhiên, mọi nỗ lực gỡ bỏ một ràng buộc bằng cách tăng dung lượng đều chỉ dời ràng buộc sang chỗ khác.

1. Yann LeCun, "A Path Towards Autonomous Machine Intelligence", bản 0.9.2, ngày 27 tháng 6 năm 2022, công bố trên OpenReview.

### 3.3.4. Hệ quả thứ nhất: mọi thực thể đều bị nén

Một khách sạn trong thực tế có hàng nghìn thuộc tính. Máy phải tổng hợp tất cả thành một biểu diễn duy nhất để so sánh trong thời gian tính bằng phần nghìn giây. Quá trình nén giữ lại cả cấu trúc quan hệ, nên máy đo sự gần gũi bằng cấu trúc thông tin chứ không bằng khoảng cách địa lý. Một khu nghỉ dưỡng ở Phú Quốc và một khu nghỉ dưỡng ở Bali nằm gần nhau trong không gian biểu diễn nếu phục vụ cùng nhóm khách, cùng mức giá, cùng xuất hiện trong những hành trình giống nhau. Hai địa điểm cách nhau một con phố có thể ở rất xa nhau nếu phục vụ hai mục đích khác hẳn.

Nén càng mạnh thì hệ thống càng nhanh và càng mất chi tiết. Thứ được giữ lại chỉ là phần đủ để hệ thống hoàn thành nhiệm vụ của nó, mà nhiệm vụ ấy do người thiết kế định nghĩa chứ không do địa phương.

### 3.3.5. Hệ quả thứ hai: phần dễ nén là phần ít phân biệt

Điều được giữ lại trước hết là những gì gọi tên được, rồi tới thuộc tính đo đếm được, rồi tới quan hệ, rồi tới tín hiệu hành vi. Điều bị bỏ lại là cảm giác bước vào một ngôi đình lúc sáng sớm khi hương trầm còn phảng phất, là bối cảnh khiến một chiếc nón trên cánh đồng khác hẳn chiếc nón trong cửa hàng lưu niệm, là tính sống của một lễ hội mỗi năm mỗi khác. Trớ trêu là những yếu tố bị bỏ lại thường chính là nguồn gốc của lợi thế cạnh tranh. Khách du lịch không bay hàng nghìn km để nhìn một điểm trên bản đồ. Nhà đầu tư không chỉ đọc số liệu dân số.

Từ đó suy ra một quy luật, phần bản sắc dễ chuyển thành dữ liệu nhất cũng là phần mà mọi địa phương khác đều có: địa chỉ, giá vé, giờ mở cửa, số sao. Phần khó chuyển nhất lại là phần duy nhất không sao chép được. Một địa phương càng chuẩn hóa tốt thì càng dễ so sánh và càng dễ so sánh thì càng dễ bị thay thế bằng một phương án tương đương rẻ hơn.

Phần bản sắc dễ nén nhất là phần ít phân biệt nhất. Chuẩn hóa mua được khả năng được so sánh và trả giá bằng khả năng không thể thay thế. Đây không phải lý do để từ chối chuẩn hóa, vì từ chối dẫn thẳng tới vô hình. Đây là lý do để chuẩn hóa có chiến lược, đưa vào lược đồ đủ để lọt tập ứng viên, giữ lại ngoài lược đồ phần chỉ trải nghiệm trực tiếp mới chạm tới được.

### 3.3.6. Hệ quả thứ ba: địa phương là một tập hợp thực thể

Máy không đọc Huế như một khái niệm. Máy đọc các thực thể cấu thành nên Huế cùng các cạnh nối giữa chúng, phân bố trên bảy lớp: con người, địa điểm, tổ chức, sự kiện, sản phẩm, trải nghiệm, tài sản.

Số lượng thực thể đã định danh và số quan hệ đã khai báo, chứ không phải tài sản thực có, quyết định độ phân giải của địa phương trong mắt máy. Chương trình OCOP minh họa khoảng cách này. Đến hết tháng Sáu năm 2026, cả nước có 21.028 sản phẩm đạt từ ba sao trở lên với 10.657 chủ thể và 126 sản phẩm năm sao cấp quốc gia.[1] Một hệ thống phân hạng công phu, có tiêu chí, có hội

1. Bộ Nông nghiệp và Môi trường, số liệu Chương trình Mỗi xã một sản phẩm tới hết tháng 6 năm 2026. Đây là số liệu quản lý nhà nước tổng hợp từ quyết định phân hạng, nên phản ánh số quyết định đã ban hành chứ không phản ánh số sản phẩm đang lưu thông.

đồng, có quyết định hành chính. Câu hỏi bỏ ngỏ là bao nhiêu trong số hơn hai mươi nghìn sản phẩm ấy tồn tại như thực thể có định danh bền, nối được với chủ thể sản xuất, vùng nguyên liệu, làng nghề và lễ hội nơi chúng ra đời. Một thứ hạng trên giấy và một thực thể máy phân giải được là hai loại tài sản khác nhau. Địa phương tồn tại với máy đúng bằng số thực thể được định danh và số quan hệ được khai báo, không bằng số tài sản nó thực có.

### 3.3.7. Biểu diễn là một vấn đề triết học

Câu hỏi mà biểu diễn đặt ra cũ hơn máy tính rất nhiều. Plato mô tả những người bị xích chỉ nhìn thấy cái bóng in trên vách hang.[1] Khác biệt duy nhất của hôm nay và cũng là điều khiến tình huống nghiêm trọng hơn, là những cái bóng bây giờ có quyền ra quyết định.

Kant cho công cụ chính xác hơn. Ông phân biệt sự vật như nó tự thân tồn tại với sự vật như nó xuất hiện với chủ thể, rồi chỉ ra rằng mọi kinh nghiệm đều đã bị nhào nặn bởi các phạm trù của giác tính.[2] Đó gần như là định nghĩa của bản thể luận trong khoa học máy tính, chỉ khác một điểm quyết định, phạm trù của Kant là điều kiện phổ quát của nhận thức, còn phạm trù của một hệ thống dữ liệu do một nhóm kỹ sư cụ thể chọn trong một hoàn cảnh cụ thể. Máy không tiếp cận Huế. Máy tiếp cận Huế đã đi qua lược đồ của nó.

1. Plato, Cộng hòa (Politeia), quyển VII, đoạn 514a tới 520a, dụ ngôn cái hang, khoảng năm 375 trước Công nguyên. Bản tiếng Anh thường dẫn là Republic, bản dịch của G. M. A. Grube, hiệu đính bởi C. D. C. Reeve, Hackett, Indianapolis, 1992.

2. Immanuel Kant, Kritik der reinen Vernunft, bản A năm 1781 và bản B năm 1787, Johann Friedrich Hartknoch, Riga. Phân biệt giữa vật tự thân và hiện tượng, cùng học thuyết về các phạm trù của giác tính. Bản tiếng Anh thường dẫn là Critique of Pure Reason, bản dịch của Paul Guyer và Allen W. Wood, Cambridge University Press, Cambridge, 1998.

Heidegger đẩy nhận xét ấy tới điểm căng thẳng nhất. Kỹ thuật hiện đại, theo ông, không phải công cụ mà là một cách phơi mở thế giới, biến mọi thứ thành nguồn dự trữ sẵn sàng được huy động.[1] Dòng sông thôi là dòng sông, nó thành trữ lượng thủy điện. Một lễ hội thành mục trong lịch sự kiện có nút đặt vé. Một làng nghề thành điểm dừng bốn mươi lăm phút. Không gì trong đó là dối trá. Vấn đề là sau đủ nhiều lần, cách mô tả bắt đầu định hình cách tồn tại. Máy càng hiểu thế giới hơn, thế giới càng phải tự biến mình thành thứ mà máy hiểu được.

### 3.3.8. Thí nghiệm Huế, khâu thứ hai

Giả sử Huế đã vượt khâu một và có mặt trong tập ứng viên. Cỗ máy bây giờ phải dựng biểu diễn của nó.

Ở lớp thực thể cấp một, Huế giàu liên kết: cố đô, triều Nguyễn, di sản thế giới, sông Hương, Đại Nội. Đi xuống một bậc thì mạng lưới thưa dần. Bún bò Huế có phải một thực thể riêng nối với vùng nguyên liệu, cơ sở đạt chuẩn, công thức được ghi nhận hay không. Nghệ nhân nhã nhạc còn hành nghề có hồ sơ thực thể nào không. Chiếc nón bài thơ có nối được với làng nghề, nghệ nhân, cửa hàng và lễ hội nơi nó xuất hiện hay không. Xuống bậc thứ ba, tức lớp trải nghiệm mà du khách thực sự trả tiền, gần như trống.

Bảy lớp thực thể của Huế phân bố rất lệch. Địa điểm gần đầy đủ vì bản đồ số nào cũng thu thập. Sự kiện rời rạc, chủ yếu quanh các kỳ festival lớn. Sản phẩm được chuẩn hóa một phần nhờ hệ

1. Martin Heidegger, "Die Frage nach der Technik", bài giảng năm 1953, in trong Vorträge und Aufsätze, Günther Neske, Pfullingen, 1954. Bản tiếng Anh, "The Question Concerning Technology", Harper and Row, New York, 1977.

thống phân hạng nông sản. Bốn lớp còn lại gần như trống: nghệ nhân, phường hội, trải nghiệm, tri thức chế biến. Phần Huế đã khả đọc là phần bất động sản. Phần chưa khả đọc là phần làm nên Huế.

Với gia đình ở Lyon, hệ quả rất cụ thể. Cỗ máy dựng được một Huế gồm di tích và khách sạn, không dựng được một Huế gồm người và nghề. Nó sẽ đề xuất Đại Nội. Nó sẽ không đề xuất một buổi học gói bánh với một gia đình ở Kim Long, dù đó chính là thứ mà một đứa trẻ chín tuổi nhớ lâu nhất.

Từ đây kéo theo nguyên tắc quản trị là không số hóa trước khi định nghĩa bản thể luận, số hóa mà chưa có lược đồ thực thể chỉ tạo ra tệp, không tạo ra tri thức. Không đo bằng số hồ sơ, chỉ đo bằng mật độ quan hệ. Mỗi thực thể trọng điểm cần tối thiểu năm tới mười cạnh nối tới các thực thể đã định danh. Một hồ sơ đứng một mình không tham gia được suy luận. Không giao khả đọc cho phòng tin học, quyết định giữ gì và bỏ gì trong lược đồ là quyết định về bản sắc, phải do hội đồng có đại diện văn hóa và cộng đồng thông qua.

Nếu là Trung tâm Bảo tồn Di tích. Ba lớp thực thể đang trống có chi phí tạo lập thấp nhất trong toàn bộ chuỗi. Một hồ sơ nghệ nhân gồm tên, nghề, cơ sở, sản phẩm, lễ hội liên quan và một định danh bền tốn chưa tới một giờ làm việc. Một trăm hồ sơ như vậy làm thay đổi độ phân giải của cả thành phố trong mắt máy và rẻ hơn một phóng sự truyền hình.

## 3.4. RÀNG BUỘC THỨ BA: CHÚ Ý KHAN HIẾM VÀ THÔNG TIN BẤT CÂN XỨNG

Nếu kênh đầu ra hẹp hơn tập ứng viên, mọi hệ thống buộc phải áp một trật tự lên tập ấy. Không tồn tại đầu ra không được xếp hạng. Một phép chọn ngẫu nhiên cũng là một hàm xếp hạng với trọng số suy biến.

### 3.4.1. Chứng minh

Gọi tập ứng viên là C, số mục được trả về là k, với k nhỏ hơn số phần tử của C. Việc chọn k phần tử từ C tương đương với việc xác định một quan hệ ưu tiên phân tách C thành phần được chọn và phần bị loại. Quan hệ ấy là xếp hạng. Nếu k bằng số phần tử của C thì không cần xếp hạng, nhưng khi ấy hệ thống đã đẩy toàn bộ việc chọn lọc sang con người và không còn thay con người quyết định.

Vì sao k luôn nhỏ? Herbert Simon trả lời từ năm 1971, sự dồi dào của thông tin tạo ra sự khan hiếm của chú ý, nên thứ cần được phân bổ hiệu quả là chú ý chứ không phải thông tin.[1] Xếp hạng là cơ chế phân bổ một nguồn lực khan hiếm.

Còn một nửa nữa của ràng buộc. Chất lượng thật không quan sát trực tiếp được. George Akerlof chỉ ra rằng khi người mua không phân biệt được hàng tốt với hàng xấu, thị trường đẩy hàng tốt ra ngoài.[2] Michael Spence bổ sung lời giải, bên có chất lượng cao phải phát ra một tín hiệu đủ tốn kém để bên kém không bắt

1. Herbert A. Simon, "Designing Organizations for an Information-Rich World", 1971. Đã dẫn đầy đủ ở chú thích 3.

2. George A. Akerlof, "The Market for Lemons: Quality Uncertainty and the Market Mechanism", Quarterly Journal of Economics, tập 84, số 3, 1970, trang 488 tới 500.

chước được.[1] Xếp hạng của máy chính là một cơ chế đọc tín hiệu và câu hỏi trung tâm của mọi địa phương là tín hiệu nào còn tốn kém. Một bộ ảnh đẹp không còn tốn kém vì ai cũng tạo được. Dữ liệu thời gian thực và cam kết thực thi thì có, vì chỉ nơi vận hành tốt mới dám công bố.

### 3.4.2. Phản ví dụ

- ***Phản ví dụ 3A. Xếp theo một tiêu chí khách quan duy nhất.*** Một ứng dụng bản đồ liệt kê nhà hàng theo khoảng cách tăng dần. Không có trọng số nào do ai chọn, chỉ có một đại lượng vật lý. Phản ví dụ thất bại. Việc chọn khoảng cách làm tiêu chí duy nhất tự nó đã là một phán quyết về giá trị và là phán quyết có hệ quả phân phối rõ rệt, nó chuyển toàn bộ lợi thế cho những nơi nằm gần điểm xuất phát của du khách, tức các trục trung tâm và loại các làng nghề ngoại vi bất kể chất lượng. Một tiêu chí đo được không đồng nghĩa với một tiêu chí trung lập.

- ***Phản ví dụ 3B. Xếp hạng có chèn ngẫu nhiên.*** Nhiều hệ thống gợi ý cố ý dành một tỷ lệ nhỏ lượt hiển thị cho các phương án chưa đủ dữ liệu, để tự học thêm. Ở phần ngẫu nhiên ấy, thứ tự không phản ánh điểm số nào. Phản ví dụ thành công một phần và đây là phản ví dụ quan trọng nhất của cả chương. Về hình thức, chèn ngẫu nhiên vẫn là một hàm xếp hạng với phân phối xác suất, nên mệnh đề đứng vững. Về chính sách, nó chứng minh một điều mà các nền tảng ít khi nói ra. Việc dành phần khám phá cho những nơi chưa có dữ liệu là khả thi về kỹ thuật và đã được triển khai ở quy mô lớn. Vậy việc các hệ thống hiện nay không dành phần ấy cho các điểm đến mới nổi là một lựa

1. Michael Spence, "Job Market Signaling", Quarterly Journal of Economics, tập 87, số 3, 1973, trang 355 tới 374.

chọn chứ không phải một tất yếu. Đó là chỗ duy nhất trong toàn chương mà chính sách công có đòn bẩy trực tiếp lên thuật toán với các cơ chế bảo đảm tỷ lệ hiển thị.[1]

• ***Phản ví dụ 3C. Trả về toàn bộ kết quả.*** Một cơ sở dữ liệu trả về mọi bản ghi khớp điều kiện, không sắp thứ tự. Phản ví dụ thành công trong phạm vi hẹp và tự bộc lộ giới hạn của nó. Hệ thống ấy không xếp hạng vì nó đẩy toàn bộ việc chọn lọc sang con người, tức nó chưa thay ai quyết định. Ngay khi một tác nhân phải trả lời thay người dùng, kênh đầu ra thu hẹp và trật tự xuất hiện trở lại. Vì thế được phát biểu chặt hơn, nó đúng khi và chỉ khi hệ thống thay con người rút gọn tập lựa chọn.

### 3.4.3. Năm nhóm tín hiệu

Trọng số không trả lời câu hỏi điều gì đúng. Nó trả lời câu hỏi điều gì nên được tin hơn trong hoàn cảnh này. Năm nhóm lặp lại trong hầu hết hệ thống. (1) Thẩm quyền của nguồn: dữ liệu từ cơ quan quản lý, tổ chức quốc tế hay trường đại học được coi là ít khả năng sai hơn. (2) Bằng chứng độc lập: một khẳng định được nhiều nguồn không liên quan xác nhận thì nặng hơn. (3) Tính cập nhật: dữ liệu phải còn đúng ở đúng thời điểm quyết định, vì hậu quả của thông tin lỗi thời không còn là một cú nhấp chuột phí phạm mà là một giao dịch thất bại. (4) Mức độ phổ biến: nơi đã nổi tiếng tiếp tục nổi tiếng và vòng phản hồi này không phân biệt chất lượng với quán tính. (5) Sự phù hợp: máy không xếp hạng thế giới, máy xếp hạng thế giới cho từng người. Bốn nhóm đầu cải thiện được bằng hạ tầng. Nhóm thứ tư thì không mua được bằng ngân sách kỹ thuật. Đó là toàn bộ chiến lược khả thi của một địa phương chưa nổi tiếng.

1. Cai-Nicolas Ziegler và cộng sự, "Improving Recommendation Lists Through Topic Diversification", 2005, đã dẫn ở phần trước. Kết luận về khả năng triển khai ở quy mô lớn là suy luận của tác giả từ tài liệu kỹ thuật công bố của các nền tảng.

### 3.4.4. Phản biện

• ***Phản bác thứ nhất.*** Cá nhân hóa giải quyết bất bình đẳng, vì khi mỗi người nhận một bảng xếp hạng riêng thì địa phương nhỏ luôn có cơ hội đứng đầu ở một phân khúc. Bác bỏ: cá nhân hóa cần chính những tín hiệu mà địa phương nhỏ thiếu nhất. Muốn biết một nơi phù hợp với gia đình có trẻ nhỏ, hệ thống cần dữ liệu về những gia đình có trẻ nhỏ đã tới và đã hài lòng. Không lưu lượng thì không tín hiệu phân khúc, không tín hiệu thì không được đề xuất cho phân khúc ấy. Cá nhân hóa không xóa vòng lặp tự củng cố, nó chia vòng lặp thành nhiều vòng nhỏ hơn.

• ***Phản bác thứ hai.*** Yêu cầu các nền tảng công khai trọng số và cho kiểm toán sẽ giải quyết vấn đề trung lập. Bác bỏ một phần: minh bạch đáng theo đuổi nhưng đặt sai chỗ. Trọng số trong hệ thống hiện đại là hệ quả của quá trình học, phân bố qua hàng tỷ tham số, chứ không phải những con số gõ vào một bảng cấu hình. Công bố chúng hiếm khi tạo ra hiểu biết dùng được. Chưa kể, trọng số là các tọa độ trong không gian tham số, chúng liên tục thay đổi theo thời gian khi mô hình học từ dữ liệu mới. Đòn bẩy của địa phương nằm ở phía đầu vào, bảo đảm rằng giá trị mình theo đuổi tồn tại dưới dạng tín hiệu đo được.

• ***Phản bác thứ ba, mạnh nhất.*** Judea Pearl chỉ ra rằng phần lớn học máy vận hành ở bậc thấp nhất của thang nhân quả, bậc tương quan và không trả lời được câu hỏi can thiệp hay phản thực.[1] Áp vào xếp hạng, điều này làm suy yếu chính khái niệm mức độ phù hợp vừa dựng ở trên. Một hệ thống quan sát thấy các gia đình có trẻ nhỏ hài lòng với Đà Nẵng hơn Huế không biết

1. Judea Pearl và Dana Mackenzie, The Book of Why: The New Science of Cause and Effect, Basic Books, New York, 2018.

vì sao. Có thể vì Huế kém phù hợp thật. Có thể vì các gia đình từng chọn Huế được đưa tới đó theo một tiêu chí khác. Có thể vì một tháng thời tiết xấu. Có thể vì chính việc bị xếp hạng thấp đã khiến các nhà cung cấp tốt không xuất hiện trong tập ứng viên. Xếp hạng dựa trên tương quan không phân biệt được bốn khả năng ấy và vì kết quả xếp hạng tác động ngược lên hành vi quan sát được ở vòng sau, hệ thống đang học từ chính hậu quả của nó. Đây không phải lỗi sửa được bằng thêm dữ liệu. Đây là giới hạn của cả một bậc suy luận và nó buộc mọi phát biểu dạng thuật toán đã chọn đúng nơi phù hợp nhất phải được đọc với sự dè dặt lớn hơn nhiều so với vẻ tự tin của giao diện.

### 3.4.5. Hệ quả thứ nhất: ngưỡng cắt

Nếu hệ thống trả ba gợi ý, phương án thứ tư có thể gần như tương đương phương án thứ ba, nhưng một bên được nhìn thấy còn bên kia biến mất. Trước ngưỡng, địa phương được đề xuất, được đặt dịch vụ, tiếp tục sinh dữ liệu. Sau ngưỡng, nó không nhận lượt tiếp cận, không tạo phản hồi, không tích lũy thêm tín hiệu nào. Quanh ngưỡng cắt, một chênh lệch rất nhỏ về điểm số tạo ra một chênh lệch rất lớn về cơ hội. Ranh giới giữa hiện diện và vô hình thường không phải khoảng cách giữa tốt và xấu mà là khoảng cách giữa thứ ba và thứ tư.

### 3.4.6. Hệ quả thứ hai: ưu thế tự tích lũy

Thẩm quyền số hành xử như một dạng vốn. Tích lũy chậm, mất nhanh, sinh lợi kép. Một điểm đến được nhiều nguồn nhắc tới thì mỗi lần xuất hiện lại làm tăng thẩm quyền cho lần sau. Một địa phương ít được biết đến có thể sở hữu giá trị đặc biệt nhưng

thiếu dấu vết xác nhận, ít được đề xuất nên ít được trải nghiệm, ít được trải nghiệm nên ít được đánh giá, ít được đánh giá nên hệ thống càng thiếu căn cứ để đề xuất.

Một chi tiết nói lên nhiều điều về cấu trúc quyền lực này. Wikidata, lớp neo thực thể mà nhiều hệ thống dựa vào, chứa hơn 122 triệu mục và được duy trì bởi khoảng bốn mươi nghìn người dùng tích cực.[1] Với phần lớn thực thể nằm ngoài rìa sự chú ý toàn cầu, mô tả đầu tiên được viết ra thường là mô tả duy nhất và nó ở lại rất lâu.

Mỗi vòng lựa chọn sinh ra dữ liệu làm tăng xác suất được lựa chọn ở vòng sau. Ưu thế và bất lợi đều tự tích lũy, nên chênh lệch giữa các địa phương giãn ra ngay cả khi chất lượng thực không đổi.

### 3.4.7. Thí nghiệm Huế, khâu thứ ba

Gia đình ở Lyon quan tâm lịch sử. Trong tập ứng viên toàn cầu cho một chuyến đi mười ngày ưu tiên di sản, Huế không cạnh tranh với Đà Nẵng. Huế cạnh tranh với Kyoto.

Hai thành phố giống nhau đến mức đáng chú ý. Cùng là cố đô. Cùng có quần thể di sản được UNESCO ghi danh. Cùng có nghệ thuật cung đình. Cùng bán một lời hứa nhìn thấy một triều đại đã mất. Rồi mọi thứ khác hẳn.

Kyoto đón khoảng 56 triệu lượt khách năm 2024, với 16,3

1. Wikidata, trang thống kê chính thức. Cao với số mục vì đếm trực tiếp. Số người dùng tích cực phụ thuộc định nghĩa của dự án. Về thiết kế của dự án xem Denny Vrandečić và Markus Krötzsch, "Wikidata: A Free Collaborative Knowledgebase", Communications of the ACM, 2014.

triệu lượt lưu trú, trong đó 8,21 triệu lượt quốc tế, một kỷ lục.[1] Huế đón hơn 4,3 triệu lượt trong sáu tháng đầu năm 2026 và đặt mục tiêu 7,5 triệu lượt cho cả năm. Chênh lệch quy mô là rõ. Nó không phải điều đáng chú ý nhất.

Có một chi tiết dễ bị bỏ qua. Một khảo sát của Đại học Kyoto cho thấy hơn 80% cư dân báo rằng đời sống hằng ngày bị ảnh hưởng vì chen chúc trên giao thông công cộng. Chính quyền thành phố đã phải triển khai hàng loạt biện pháp phân tán dòng khách. Huế thì còn dư địa sức chứa trong phần lớn thời gian năm.[2]

Hệ thống xếp hạng đang gửi thêm khách tới nơi muốn ít khách hơn và gửi ít khách hơn tới nơi cần thêm khách. Nó không sai một phép tính nào. Nó chỉ chưa từng được hỏi về sức chứa. Nguyên nhân đọc được ngay từ năm nhóm tín hiệu. Kyoto vượt trội ở mức độ phổ biến, ở khối lượng bằng chứng độc lập, ở mật độ dữ liệu thời gian thực, ở khả năng hoàn tất giao dịch bằng máy. Không nhóm nào đo sức chứa còn lại, mức độ hài lòng của cư dân, hay lợi ích biên của một du khách tăng thêm đối với cộng đồng sở tại. Những đại lượng ấy tồn tại, đo được và trong trường hợp Kyoto đã được đo. Chúng chỉ không có mặt trong bất kỳ trường dữ liệu nào mà tác nhân đọc.

Hai kết luận đi ngược nhau và đều đúng. Với Huế, bài học không phải là cần nổi tiếng hơn Kyoto, điều vừa bất khả thi vừa không đáng mong muốn. Bài học là bốn trong năm nhóm tín

1. Thống kê du lịch thành phố Kyoto, số liệu năm 2024. Con số 56 triệu là lượt tham quan, 16,3 triệu là lượt lưu trú qua đêm, còn số của Huế theo phương pháp thống kê Việt Nam. Ba con số theo ba định nghĩa khác nhau nên chỉ so sánh được về bậc độ lớn.

2. Khảo sát của Đại học Kyoto về thái độ cư dân với du lịch, dẫn lại theo tường thuật báo chí. Kết quả loại này nhạy cảm với cách đặt câu hỏi và địa bàn lấy mẫu.

hiệu cải thiện được bằng hạ tầng và nhóm thứ năm là nhóm duy nhất mà một điểm đến ít đông đúc có lợi thế cấu trúc, với điều kiện lợi thế ấy tồn tại dưới dạng dữ liệu. Với Kyoto, bài học là một điểm đến quá tải không có công cụ nào để tự hạ thứ hạng của mình, bởi hệ thống không có trường dữ liệu nào tiếp nhận câu nói xin đừng gửi thêm người đến đây vào tháng Tư.

Không tồn tại kết quả nào không bị xếp hạng và không tồn tại bảng xếp hạng nào trung lập. Mỗi bảng xếp hạng là một quan điểm về giá trị được viết dưới dạng thuật toán. Vì thế nguyên tắc quản trị là không đo bằng lượt xem khi cạnh tranh diễn ra ở tập ứng viên. Chỉ số đúng là tỷ lệ lọt vào tập ứng viên trên một bộ kịch bản yêu cầu cố định, đo định kỳ. Không tuyên bố một giá trị mà không tạo trường dữ liệu cho nó. Sức chứa còn lại, tỷ lệ doanh thu ở lại cộng đồng, cam kết môi trường phải tồn tại dưới dạng dữ liệu đo được, nếu không thuật toán sẽ không thấy và mọi tuyên bố dừng ở văn bản. Không kết luận về chất lượng từ thứ hạng. Thứ hạng thấp có thể phản ánh chất lượng, có thể phản ánh thiếu dữ liệu, có thể phản ánh hậu quả của lần xếp hạng trước. Kiểm tra sự tồn tại của tín hiệu trước khi kết luận về sản phẩm.

Nếu là Trung tâm Xúc tiến đầu tư. Bộ hồ sơ mời gọi đầu tư hiện được viết cho người đọc và gửi kèm thư mời. Cùng nội dung ấy phải tồn tại ở dạng máy đọc được, với các trường mà một tác nhân sàng lọc địa điểm sẽ hỏi trước tiên: diện tích đất sạch còn lại theo từng khu, công suất điện đấu nối được, thời gian thông quan trung bình, mức lương kỹ thuật viên, thời gian cấp phép thực tế của mười hai tháng gần nhất. Đây là những tín hiệu tốn kém theo nghĩa của Spence, vì chỉ nơi thực sự làm được mới dám công bố kèm ngày cập nhật.

## 3.5. RÀNG BUỘC THỨ TƯ: CHI PHÍ GIAO DỊCH

Ủy nhiệm không phải hệ quả logic của sự hữu hạn. Nó xảy ra khi và chỉ khi chi phí giao dịch của việc giao việc cho máy thấp hơn chi phí tự làm. Vì vậy, mệnh đề này có điều kiện và điều kiện ấy đang được thỏa mãn theo một thứ tự dự báo được.

### 3.5.1. Chứng minh

Ronald Coase lập luận từ năm 1937 rằng ranh giới giữa tự làm và giao cho bên khác được quyết định bởi chi phí giao dịch, gồm chi phí tìm đối tác, mô tả yêu cầu, thương lượng, giám sát và khắc phục sai sót.[1] Suốt ba thập niên, chi phí giao dịch của việc ủy quyền cho máy quá cao, phải mô tả nhiệm vụ bằng ngôn ngữ hình thức, phải tin rằng máy không hiểu sai, phải kiểm tra lại toàn bộ kết quả. Ngôn ngữ tự nhiên kéo chi phí mô tả xuống gần bằng không. Các giao thức thanh toán và xác thực đang kéo chi phí giám sát xuống theo. Khi tổng chi phí giao dịch giảm đủ sâu, biên giới dịch chuyển. Ủy nhiệm vì thế bùng nổ không phải vì máy đột nhiên thông minh hơn mà vì việc thuê máy đột nhiên rẻ hơn.

### 3.5.2. Phản ví dụ

Đã có điều kiện, nên phản ví dụ hữu ích là tìm nơi chi phí giao dịch đã thấp mà ủy nhiệm vẫn không xảy ra, chứ không phải tìm nơi ủy nhiệm không xảy ra, hoặc ngược lại.

1. Ronald H. Coase, "The Nature of the Firm", Economica, tập 4, số 16, 1937, trang 386 tới 405.

• ***Phản ví dụ 4A. Y tế.*** Chi phí mô tả và giám sát nhiều tác vụ y khoa đã giảm mạnh, hệ thống đọc được hồ sơ, nhắc lịch, sàng lọc triệu chứng. Vậy mà quyết định lâm sàng vẫn không được ủy nhiệm. Phản ví dụ thành công và buộc phát biểu lại mệnh đề. Biến còn thiếu là khả năng hoàn tác. Một lần đặt nhầm bàn ăn sửa được trong ba mươi giây. Một lần chẩn đoán nhầm thì không. Chi phí giao dịch thấp là điều kiện cần, không phải điều kiện đủ.

• ***Phản ví dụ 4B. Giao dịch tài chính tự động.*** Ở chiều ngược lại, việc ủy nhiệm cho máy trong giao dịch chứng khoán đã diễn ra từ nhiều thập niên, dù rủi ro rất lớn và hậu quả khó hoàn tác. Phản ví dụ thành công và bổ sung biến thứ ba là quy mô lợi ích biên. Khi lợi ích của tốc độ đủ lớn, người ta chấp nhận rủi ro không hoàn tác được và bù bằng hạn mức, ngắt tự động và bảo hiểm. Cấu trúc bù rủi ro ấy chính là thứ mà hạ tầng công của một địa phương phải xây nếu muốn ủy nhiệm mở rộng ra ngoài nhóm giao dịch nhỏ.

• ***Phát biểu lại:*** Ủy nhiệm xảy ra khi chi phí giao dịch giảm đủ sâu và đồng thời hậu quả sai sót hoặc hoàn tác được, hoặc bù đắp được bằng một cơ chế chịu rủi ro. Thiếu vế thứ hai, chi phí thấp một mình không đủ. Phiên bản chặt hơn này giữ nguyên khuyến nghị về thứ tự đầu tư và làm rõ lý do. Vé tham quan được ủy nhiệm trước không chỉ vì rẻ mà vì hủy được. Hợp đồng đầu tư đến sau không chỉ vì lớn mà vì không có nút hoàn tác.

### 3.5.3. Một dự báo kiểm chứng được

Cách đọc theo chi phí cho một dự báo mà cách đọc theo năng lực không cho được. Nếu ủy nhiệm lan theo chi phí giao dịch, nó sẽ lan nhanh nhất ở những giao dịch giá trị nhỏ, lặp lại nhiều, hậu

quả sai thấp và hoàn tác được, chứ không phải ở những giao dịch quan trọng nhất. Đặt bàn ăn tối trước khi mua bảo hiểm. Gia hạn dịch vụ trước khi chọn trường cho con. Vé tham quan trước khi hợp đồng đầu tư.

Dữ liệu ngành ủng hộ dự báo này và đồng thời cho thấy khoảng cách hạ tầng. Theo Phocuswright năm 2026, khoảng 61% doanh nghiệp du lịch đang thử nghiệm hoặc mở rộng ứng dụng tác nhân, nhưng chỉ 6% triển khai ở quy mô thực sự.[1] Một nghiên cứu công bố tháng Sáu năm 2026 cho thấy chỉ 11% tổ chức khách sạn hoàn tất được một giao dịch đặt phòng qua tác nhân theo thời gian thực. Gần chín phần mười ngành vẫn tối ưu cho các kênh phân phối của thời kỳ trước.

### 3.5.4. Phản biện

• ***Phản bác thứ nhất, mạnh nhất.*** Gary Marcus cảnh báo nhiều năm rằng học sâu đang chạm trần, rằng các hệ thống hiện nay giòn, không đáng tin ngoài phân bố huấn luyện và bịa ra thông tin theo cách khó dự đoán.[2] Nếu ông đúng, ủy nhiệm giao dịch tài chính cho chúng là liều lĩnh. Vế thứ nhất phải được thừa nhận rằng bằng chứng thực nghiệm đứng về phía Marcus nhiều hơn về phía các lời quảng bá. Nhưng kết luận không phải là ủy nhiệm sẽ không xảy ra, mà là nó xảy ra đúng theo thứ tự mà logic chi

1. Phocuswright, Budgets, Barriers and the Race to Agentic AI, công bố tháng 1 năm 2026, https://www.phocuswright.com/Travel-Research/Consumer-Trends/Budgets-Barriers-and-the-Race-to-Agentic-AI. Khảo sát toàn cầu với lãnh đạo cấp cao ngành du lịch: 61% doanh nghiệp đang thử nghiệm hoặc mở rộng tác nhân AI, trong đó 6% đã triển khai ở quy mô nhiều lĩnh vực. Mẫu nghiêng về doanh nghiệp lớn và phương Tây.

2. Gary Marcus, "The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence", arXiv:2002.06177, 2020. Xem thêm Gary Marcus, "Deep Learning: A Critical Appraisal", arXiv:1801.00631, 2018.

phí giao dịch dự báo, từ những giao dịch nhỏ và hoàn tác được, nơi chi phí một lần sai thấp hơn lợi ích của hàng trăm lần đúng. Phản biện của Marcus không phải lý do để chờ mà là lý do để xác định thứ tự đầu tư. Nó cũng đặt ra một câu hỏi chương này không trả lời được: Ai chịu trách nhiệm khi máy đặt nhầm hoặc mô tả sai dịch vụ của một hộ kinh doanh.

• ***Phản bác thứ hai.*** Yêu cầu hạ tầng khả hành động là một tiêu chuẩn loại trừ viết bằng ngôn ngữ trung lập, vì phần lớn hộ kinh doanh, hợp tác xã và ban quản lý di tích không có năng lực kỹ thuật lẫn vốn để đáp ứng. Phản bác này đúng và chỉ ra một rủi ro phân phối thực. Kết luận rút ra là chuyển tiêu chuẩn lên cấp phù hợp, không phải từ bỏ tiêu chuẩn. Không hộ kinh doanh nào tự xây lưới điện và cũng không hộ nào nên tự xây cổng thanh toán cho tác nhân. Hạ tầng khả hành động thuộc loại hàng hóa công có tính chia sẻ, giống chợ đầu mối hay đường liên xã. Để từng chủ thể tự lo thì kết quả gần như chắc chắn là chỉ các chuỗi lớn còn hiện diện.

### 3.5.5. Hệ quả: khả đọc không còn đủ

Một tác nhân không hoàn thành nhiệm vụ chỉ bằng việc biết một khách sạn tồn tại. Nó cần biết còn phòng hay không, giá lúc này bao nhiêu, điều kiện hủy thế nào, thanh toán bằng cách nào và quan trọng nhất là có giao diện để hoàn tất hay không.

Năm điều kiện tách ra từ thực tế vận hành. Giao diện kết nối để máy truy cập và thao tác. Khả năng đặt chỗ để kiểm tra tình trạng rồi hoàn tất. Khả năng thanh toán, vì nếu bước này buộc quay lại con người thì toàn bộ lợi thế tốc độ biến mất. Cơ chế xác nhận trả về mã đặt chỗ, vé, điều kiện đổi hủy, bởi một xác nhận mơ hồ nguy hiểm hơn một lời từ chối dứt khoát. Dữ liệu thời gian thực, vì dữ liệu tĩnh mất giá rất nhanh trong môi trường này.

### 3.5.6. Thí nghiệm Huế, khâu thứ tư

Gia đình ở Lyon yêu cầu đặt trọn phòng và vé trước khi bay. Tác nhân phải hoàn tất một chuỗi: đặt phòng, mua vé Đại Nội và các lăng, giữ chỗ một buổi nhã nhạc, đặt bàn quán bún bò buổi sáng, thuê xe đi lăng vua Gia Long, mua bảo hiểm, rồi xử lý tình huống mưa lớn khiến một điểm đóng cửa.

Ở mỗi mắt xích, câu hỏi không phải dịch vụ ấy có tồn tại hay không. Tất cả đều tồn tại và phần lớn đều tốt. Câu hỏi là mắt xích ấy có giao diện để máy hoàn tất giao dịch, có trả về mã xác nhận, có báo được tình trạng còn chỗ theo thời gian thực hay không. Chỉ cần một mắt xích buộc quay lại con người, chẳng hạn, buổi nhã nhạc chỉ nhận đặt qua điện thoại, thì tác nhân có động cơ chuyển cả cụm hoạt động ấy sang một điểm đến khác nơi chuỗi chạy liền mạch.

Mắt xích yếu không phải mắt xích lớn. Khách sạn ở Huế phần lớn đã nằm trong các hệ thống phân phối quốc tế, nên khâu lưu trú gần như đã khả hành động. Những mắt xích đứt lại là vé di tích, suất diễn nghệ thuật truyền thống, lớp học nấu ăn, buổi trải nghiệm làm nón, tức đúng những sản phẩm mang bản sắc và đúng những chủ thể nhỏ nhất.

Phần khả hành động của địa phương lại là phần ít bản sắc nhất. Nếu không can thiệp, tác nhân sẽ dựng nên một Huế gồm khách sạn quốc tế và vận chuyển, còn phần văn hóa để lại cho du khách tự xoay xở, mà phần lớn du khách sẽ không xoay xở.

Điều không thể thực thi thì không được giao. Trong nền kinh tế tác nhân, được đề xuất là điều kiện cần, được giao việc mới là điều kiện đủ. Vì thế, nguyên tắc của quản trị là không xây

giao diện khi chưa chuẩn hóa thực thể. Một giao diện trỏ vào dữ liệu chưa định danh chỉ nhân bản sự lộn xộn với tốc độ cao hơn. Không để từng chủ thể tự lo hạ tầng khả hành động: cổng đặt chỗ, thanh toán và xác thực là hàng hóa công cấp tỉnh, giao cho từng hộ tương đương với quyết định loại họ khỏi chuỗi. Không bắt đầu từ sản phẩm cao cấp nhất: Theo chi phí giao dịch, vé tham quan và vé sự kiện sẽ được ủy nhiệm trước và đó cũng là nhóm mang bản sắc nhiều nhất đang đứng ngoài chuỗi.

Nếu là Chủ tịch tỉnh: Phép đo duy nhất cần yêu cầu là tỷ lệ chuỗi hoàn tất. Chọn năm hành trình mẫu tương ứng năm nhóm khách chính, để một tác nhân thực hiện trọn vẹn, đếm số lần buộc phải quay lại con người. Con số ấy chính là khoảng cách giữa được biết đến và được giao việc và nó sẽ chỉ thẳng vào những đơn vị đang giữ vé giấy.

## 3.6. ĐIỀU KIỆN TIỀN ĐỀ: BẢN THỂ LUẬN

Ba mệnh đề đầu đều giả định rằng đã tồn tại một tập các loại đối tượng mà hệ thống công nhận. Không có hạng mục thì không có thực thể, không có thực thể thì không có gì để truy hồi, biểu diễn hay xếp hạng. Bản thể luận vì thế không phải thao tác thứ năm mà là điều kiện tiền đề của cả bốn.

### 3.6.1. Chứng minh

Truy hồi là phép chọn tập con của một tập đã được xác định. Muốn tập ấy tồn tại, các phần tử của nó phải phân biệt được với nhau, tức phải thuộc về những loại đã được định nghĩa. Biểu diễn là phép chiếu lên một hệ đo và hệ đo chỉ dựng được khi biết cần đo những chiều nào, tức khi biết đối tượng thuộc loại gì. Xếp hạng so sánh các phần tử trong cùng một tập và một tập không đồng loại thì không so sánh được. Vậy cả ba thao tác đều lấy bản thể luận làm đầu vào. Bản thể luận dựng trên bốn thành phần. Khái niệm xác định những lớp đối tượng được công nhận là tồn tại. Danh mục xác định thực thể thuộc nhóm nào. Thuộc tính cung cấp các chiều để so sánh. Quan hệ nối các thực thể thành mạng lưới.

Hệ quả rất cụ thể. Nếu một làng nghề chỉ được gắn nhãn cửa hàng, khả năng nó được đề xuất khi người dùng hỏi về trải nghiệm văn hóa gần bằng không, dù thực tế nó là nơi giàu trải nghiệm nhất vùng. Nếu một hệ thống có khái niệm bảo tàng nhưng không có khái niệm không gian văn hóa cộng đồng, hai loại hình ấy được đối xử rất khác. Điều máy không gọi tên được thì máy khó đề xuất.

### 3.6.2. Phản ví dụ

• ***Phản ví dụ 5A. Tìm kiếm ngữ nghĩa thuần véc-tơ.*** Một hệ thống tìm theo độ tương đồng, không danh mục, không lược đồ, không nhãn. Người dùng hỏi về nơi yên tĩnh có kiến trúc cổ và hệ thống trả về đúng thứ cần, chưa từng phải phân loại gì. Phản ví dụ thất bại. Các cụm trong không gian véc-tơ chính là một bản thể luận, chỉ khác ở chỗ nó hình thành từ dữ liệu huấn luyện thay vì được ai đó viết ra. Bản thể luận ngầm không dễ chịu hơn bản thể luận tường minh: không ai đọc được nó, không ai sửa được nó và nó thừa hưởng nguyên vẹn mọi thiên lệch của kho văn bản đã sinh ra nó. Một loại hình chỉ được nhắc tới bằng vài chục văn bản tiếng Việt sẽ không có cụm riêng, nên sẽ bị hút vào cụm gần nhất.

• ***Phản ví dụ 5B. Gắn thẻ tự do.*** Cộng đồng gắn thẻ tự do cho ảnh và địa điểm, không theo bảng phân loại nào. Trật tự nổi lên từ dưới thay vì được áp từ trên. Phản ví dụ thành công một phần và mở ra một chiến lược thật. Thẻ tự do quả thực không cần lược đồ ban đầu, nhưng qua thời gian chúng hội tụ về một tập nhỏ do tần suất sử dụng, tức bản thể luận tái xuất hiện, chỉ theo hướng ngược lại. Ai gắn thẻ nhiều nhất thì người đó định nghĩa. Với một địa phương, đây là con đường rẻ nhất để đưa những loại hình chưa có ô vào ngôn ngữ chung, không phải xin thêm một lớp khái niệm mà là tạo đủ khối lượng sử dụng nhất quán để một thuật ngữ tự trở thành chuẩn.

• ***Phản ví dụ 5C. Mô hình đa phương thức mô tả tự do.*** Đưa cho hệ thống một bức ảnh ngôi đình, nó mô tả trôi chảy cả chức năng thờ tự lẫn sinh hoạt cộng đồng mà không cần bất kỳ nhãn nào. Phản ví dụ thất bại ở tầng hành động. Mô tả trôi chảy không giúp ngôi đình lọt qua một bộ lọc truy vấn theo loại hình, không giúp

nó xuất hiện trong danh sách địa điểm mở cửa sau mười bảy giờ, không giúp nó bán được một vé. Khoảng cách giữa hiểu và xử lý được chính là khoảng cách mà bản thể luận tồn tại.

### 3.6.3. Ai viết bản thể luận

Không cơ quan nào định nghĩa toàn bộ thế giới số. Các nền tảng bản đồ thương mại gán hàng tỷ địa điểm vào những loại hình có sẵn. Các dự án bản đồ mở đóng vai trò không kém, cơ sở dữ liệu của OpenStreetMap đã vượt mười tỷ nút với khoảng 2,25 triệu người từng đóng góp và dữ liệu ấy thành lớp nền cho rất nhiều hệ thống.[1] Các bách khoa toàn thư mở cung cấp lớp thực thể và quan hệ. Các chuẩn dữ liệu có cấu trúc cho phép một trang web tự khai báo mình là sự kiện hay tuyến du lịch.

Về lý thuyết các bộ từ vựng chuẩn rất phong phú. Trên thực tế, phần được dùng lại rất hẹp. Bộ dữ liệu về mức độ sử dụng schema.org công bố năm 2026 cho thấy, chỉ 43 thuật ngữ, gồm 12 lớp đối tượng và 31 vị từ, xuất hiện trên hơn mười triệu tên miền, tương đương chưa tới 0,8% toàn bộ từ vựng.[2] Ngôn ngữ chung thực sự của web có cấu trúc chỉ gồm vài chục từ.

Hai hàm ý trái chiều. Mô tả mình bằng thuật ngữ nằm ngoài nhóm phổ dụng gần như không tạo hiệu lực. Nhưng đúng vì nhóm phổ dụng quá hẹp, mọi bản sắc đều đang bị buộc tự diễn

1. OpenStreetMap, trang thống kê cộng đồng, https://wiki.openstreetmap.org/wiki/Stats. Đây là số liệu chạy liên tục. Độ tin cậy cao với số nút vì là phép đếm trực tiếp. Số người đóng góp là ước lượng và các cách đếm khác nhau cho kết quả chênh lệch tới vài bậc.

2. Web Data Commons, bộ dữ liệu trích xuất dữ liệu có cấu trúc từ kho Common Crawl, Đại học Mannheim, https://webdatacommons.org/structureddata/. Ghi chú biên tập: bản mới nhất xác minh được là bản trích xuất Common Crawl tháng 10 năm 2024, công bố tháng 1 năm 2025. Tỷ lệ trên tổng từ vựng schema.org là ước lượng của tác giả.

đạt bằng một vốn từ nghèo hơn rất nhiều so với thực tại mà nó muốn mô tả.

Khu vực công là nguồn thẩm quyền chưa được khai thác đúng mức. Danh mục di tích quốc gia, di sản phi vật thể, khu du lịch quốc gia, công viên địa chất đều là hệ thống phân loại chính thức, đã có tiêu chí, đã qua thẩm định, đã ban hành bằng văn bản. Số hóa theo chuẩn mở, có định danh bền, liên kết được với các cơ sở tri thức toàn cầu thì chúng thành nguồn xác nhận rất mạnh. Nằm trong tệp văn bản trên cổng thông tin thì chúng gần như không tồn tại.

### 3.6.4. Mỗi bản thể luận là một thế giới quan

Michel Foucault mở cuốn sách nổi tiếng nhất của mình về trật tự tri thức bằng một tràng cười, khi đọc bảng phân loại động vật trong một bách khoa toàn thư Trung Hoa hư cấu do Borges dựng ra, nơi động vật được chia thành những nhóm như thuộc về hoàng đế, đã ướp xác, thuần dưỡng, vừa mới đánh vỡ một chiếc bình và nhìn từ xa trông như ruồi.[1] Tiếng cười đến từ chỗ không có mặt phẳng chung nào để các nhóm ấy cùng tồn tại. Nhưng nó cũng làm lộ ra một điều là mọi bảng phân loại đều mang một trật tự tùy thuộc thời đại của người phân loại và bảng phân loại của chính chúng ta chỉ có vẻ tự nhiên vì ta đang đứng bên trong nó.

Đó không phải chuyện đùa với một địa phương. Danh mục loại hình địa điểm của các nền tảng bản đồ cũng tùy thuộc thời đại đúng như bách khoa toàn thư của Borges, chỉ khác là nó vận hành ở quy mô hàng tỷ đối tượng và không ai cười. Trong bảng

1. Jorge Luis Borges, "El idioma analítico de John Wilkins", đăng lần đầu năm 1942, in trong Otras inquisiciones, Sur, Buenos Aires, 1952. Michel Foucault dẫn lại ở phần mở đầu Les mots et les choses, Gallimard, Paris, 1966.

ấy, một ngôi đình làng phải chọn giữa nơi thờ tự, điểm tham quan và địa danh, trong khi bản chất của nó là cả ba cộng một, chức năng thứ tư không có ô nào tương ứng, đó là nơi cộng đồng tự quản. Quan hệ giữa quyền lực và tri thức, trong nền kinh tế tác nhân, có một hình thức rất tầm thường, nó là danh sách các giá trị được phép điền vào một trường dữ liệu.

### 3.6.5. Phản biện

- ***Phản bác thứ nhất.*** Các mô hình hiện đại không cần bản thể luận cứng để hiểu rằng một ngôi đình vừa là nơi thờ tự vừa là không gian sinh hoạt cộng đồng. Bác bỏ: nhận xét đúng ở tầng hội thoại, ít đúng ở tầng hành động. Khi một tác nhân phải lọc, so sánh, đặt chỗ hoặc thanh toán, nó làm việc với các trường dữ liệu, bộ lọc và điều kiện truy vấn chứ không với sự thấu hiểu tinh tế. Ở đó một nhãn sai vẫn loại thực thể khỏi tập ứng viên bất kể mô hình hiểu tới đâu.

- ***Phản bác thứ hai.*** Một tỉnh không thể tác động vào các bản thể luận toàn cầu, nên bàn về chủ quyền bản thể luận là khát vọng suông. Bác bỏ một phần: thêm một lớp khái niệm mới vào chuẩn quốc tế quả thực nằm ngoài tầm với của hầu hết địa phương và cũng không nên là mục tiêu. Nhưng phần lớn quyền định nghĩa trên thực tế nằm ở tầng thấp hơn nhiều, như dùng tên gọi nào một cách nhất quán trên mọi văn bản chính thức, gán thực thể của mình vào danh mục nào, công bố dữ liệu theo chuẩn nào, cấp định danh bền cho tài sản nào, ai viết mô tả đầu tiên trong các cơ sở tri thức mở. Đó là một chuỗi quyết định hành chính rất cụ thể, phần lớn không tốn kém và phần lớn đang bị bỏ trống. Bởi các hệ thống ngôn ngữ lớn, tự nó cũng không ngừng cào dữ liệu cập nhật, học máy và huấn luyện lại mô hình liên tục.

### 3.6.6. Thí nghiệm với Huế, điều kiện tiền đề

Huế có ba loại tài sản không có ô nào tương ứng trong các bản thể luận đang vận hành. Nhà rường là một loại hình kiến trúc ở, không phải bảo tàng, không phải khách sạn, không phải di tích, dù có thể đồng thời là cả ba. Phủ đệ là dinh thự của hoàng thân, một loại hình gần như chỉ Huế có. Ca Huế thính phòng là một hình thức diễn xướng gắn với không gian nhỏ và với người nghe cụ thể, khác hẳn một buổi biểu diễn bán vé. Ba loại hình ấy hoặc bị ép vào ô gần giống nhất hoặc biến mất. Khi bị ép, chúng mất đúng phần khiến chúng khác biệt. Gia đình ở Lyon sẽ không bao giờ được đề xuất một đêm ca Huế thính phòng trong một phủ đệ, không phải vì trải nghiệm ấy không tồn tại mà vì không có hạng mục nào để mô tả nó.

Ai định nghĩa hạng mục thì người đó quyết định điều gì được phép tồn tại. Địa phương nào chỉ nhận các hạng mục do người khác tạo ra sẽ vĩnh viễn được diễn giải bằng ngôn ngữ của người khác. Vì thế, nguyên tắc quản trị là không đổi tên mà không khai báo kế thừa. Một tên mới không kèm quan hệ kế thừa khiến hệ thống coi đó là thực thể khác và xóa toàn bộ thẩm quyền số đã tích lũy. Không để cộng đồng đứng ngoài việc định nghĩa chính mình, trao vai trò biên tập chính thức cho một nhóm tại chỗ ở mỗi địa bàn. Chi phí gần bằng không so với một chiến dịch truyền thông, tác động dài hạn lớn hơn nhiều. Không coi phân loại là việc kỹ thuật. Quyết định về tên gọi, danh mục và định danh phải qua hội đồng có đại diện văn hóa và cộng đồng, không nằm trọn trong tay đơn vị vận hành hệ thống thông tin.

Nếu là sở quản lý, việc cần làm trong một quý là lập bảng ánh xạ. Với mỗi loại hình bản địa như nhà rường, phủ đệ, ca Huế thính phòng, phường hội nghề, chỉ rõ nó sẽ được khai báo bằng lớp phổ dụng nào, mất mát gì khi ánh xạ và bổ sung thuộc tính nào để bù phần mất. Đây là công việc xử lý một lần nhưng dùng được trong nhiều năm và không đơn vị nào ngoài sở quản lý của địa phương có chức năng làm điều đó.

## 3.7. NGUYÊN LÝ KHẢ ĐỌC THUẬT TOÁN

Bốn ràng buộc sinh ra bốn thao tác. Bốn thao tác nối nhau bằng phép nhân. Từ đó suy ra mệnh đề tổng hợp của chương.

**NGUYÊN LÝ KHẢ ĐỌC THUẬT TOÁN**

*Một địa phương chỉ có thể được máy lựa chọn khi đồng thời thỏa mãn bốn điều kiện: có thể được truy hồi, có thể được biểu diễn đúng, có thể vượt qua ngưỡng xếp hạng và có thể được thực thi. Thiếu bất kỳ điều kiện nào, xác suất được lựa chọn tiến về không, bất kể giá trị thực của địa phương lớn đến đâu.*

Mệnh đề này sẽ bị bác bỏ nếu quan sát được một trong ba điều sau. Nếu một hệ thống liên tục đưa đúng thực thể vào tập lựa chọn dù thực thể ấy không có định danh ổn định, không có nguồn đối chiếu và có nhiều bản ghi xung đột. Nếu chất lượng khả đọc đo trước không có tương quan đáng kể với xác suất xuất hiện trong kết quả do tác nhân dựng. Hoặc nếu việc cải thiện dữ liệu nền không làm giảm sai lệch thực thể trong các vòng kiểm toán lặp lại. Ba phép thử này đo được và đó là điều phân biệt một mệnh đề với một cách nói. Ở đây dùng chữ nguyên lý chứ không dùng chữ định lý, vì mệnh đề trên không được suy ra từ một hệ tiên đề hình thức mà được rút ra từ bốn ràng buộc quan sát được cùng các điều kiện bác bỏ vừa nêu.

Nguyên lý này không nói về trí tuệ nhân tạo. Nó nói về mọi hệ thống suy luận hữu hạn phải hành động trong một thế giới lớn hơn

bộ nhớ của nó. Các mô hình hôm nay chỉ là hiện thân đương thời của bốn ràng buộc ấy. Khi chúng bị thay thế, nguyên lý vẫn còn.

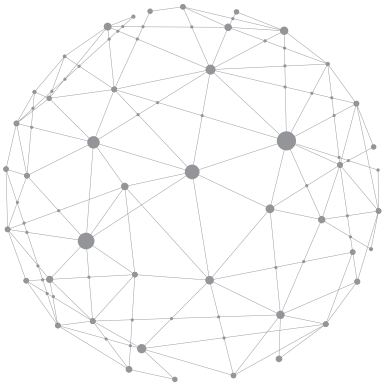
### 3.7.1. Hành trình mà cỗ máy trả về

Gia đình ở Lyon nhận được bản kế hoạch mười ngày. Bốn ngày ở Hà Nội và vịnh Hạ Long, ba ngày ở Hội An và Đà Nẵng, hai ngày ở Thành phố Hồ Chí Minh, một ngày dự phòng. Huế xuất hiện, dưới dạng một chuyến đi trong ngày từ Đà Nẵng, gồm Đại Nội và một lăng, ăn trưa tại một nhà hàng có mặt trên hệ thống đặt bàn quốc tế. Không có đêm nào ở Huế. Không có ca Huế trên sông. Không có buổi học nấu ăn. Không có nghệ nhân nào. Không có phủ đệ nào. Doanh thu mà Huế nhận được từ gia đình này bằng hai vé tham quan, một bữa trưa và một phần phí vận chuyển. Huế không thua vì kém. Huế thua bốn lần liên tiếp ở bốn khâu khác nhau, mỗi lần vì một lý do có thể sửa được và không lần nào ai được báo tin.[1]

Từng bước dẫn tới kết quả ấy đều đúng theo logic của cỗ máy. Bốn trường dữ liệu về điều kiện tham quan không tồn tại nên ràng buộc của khách không xác nhận được. Lớp thực thể cấp hai và cấp ba trống nên không có gì để đề xuất ngoài di tích. Bốn trong năm nhóm tín hiệu nghiêng về những nơi đã dẫn trước. Chuỗi đặt chỗ đứt ở đúng các sản phẩm mang bản sắc. Và ba loại hình đặc thù nhất không có hạng mục để tồn tại.

Nguyên lý cho biết điều kiện cần và đủ. Nó không cho biết cách xây. Mười một mệnh đề nền và mười bảy nguyên tắc trong chương này mô tả những ràng buộc mà một thiết kế phải tôn trọng, không mô tả bản thiết kế.

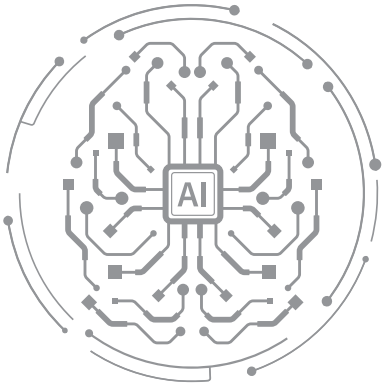
1. Giả định của tác giả. Kết quả bốn khâu trong đoạn này là hệ quả của kịch bản hư cấu đã dựng ở đầu chương, không phải một lần chạy thử có thật trên hệ thống cụ thể.



## CHƯƠNG 4

# NIỀM TIN VÀ KÝ ỨC CỦA MÁY

*Hai bộ lọc đầu và cái vòng. Máy phân giải thực thể, đối chiếu xuất xứ, tự viết câu mô tả, rồi ghi nhớ. Chương giải thích vì sao sửa sai muộn tốn kém phi tuyến và vì sao quản trị thương hiệu trở thành quản trị ký ức.*


AI

Chương này ở lại trong hai bộ lọc đầu của mô hình dựng ở Chương 2. Phân giải trả lời câu hỏi máy có nhận ra địa phương là một thực thể duy nhất hay không. Chứng thực trả lời câu hỏi máy có tin những gì được nói về thực thể ấy hay không. Nhưng chương này còn thêm một chi tiết mà Chương 2 chỉ nêu tên về "cái vòng". Bộ lọc là một cửa, đi qua một lần là xong. Cái vòng thì không có lần nào là lần cuối, vì mỗi lần được đọc lại làm cho lần đọc sau dễ hơn và mỗi lần bị đọc sai lại làm cho cái sai bền hơn. Bộ lọc quyết định anh có được vào. Cái vòng quyết định anh có ra được nữa hay không.

Máy không biết thế giới bằng quan sát trực tiếp mà bằng những thực thể mà nó phân giải, đối chiếu và chứng thực được. Khi một mô tả được chấp nhận đủ lâu, nó không còn chỉ là tri thức mà trở thành ký ức thuật toán. Ký ức ấy tạo ra phụ thuộc đường dẫn, khiến thẩm quyền tích lũy chậm nhưng mất rất nhanh, còn chi phí sửa sai tăng lên sau mỗi vòng lặp.

Ở chương trước, cuốn sách đã trả lời câu hỏi máy đọc địa phương bằng kiến trúc gì. Một hệ thống suy luận hiện đại, dù mang tên công cụ hay mô hình nào, đều phải đi qua bốn chức năng tối thiểu: truy hồi, biểu diễn, xếp hạng và ủy nhiệm. Chương này đi sâu hơn vào ba câu hỏi còn bỏ ngỏ. Máy xác định đối tượng bằng cách nào? Vì sao nó tin một nguồn hơn một nguồn khác? Và vì sao một mô tả được lặp lại lâu ngày lại trở nên khó thay đổi đến vậy?

Con người thường hình dung trí tuệ nhân tạo giống một người đọc cực nhanh, tiếp nhận hàng triệu tài liệu rồi tổng hợp thành câu trả lời. Cách hình dung ấy đúng nhưng chưa đủ. Trước khi có thể suy luận, hệ thống phải xác định điều đang được nhắc tới là thực thể nào. Trước khi có thể tin, nó phải biết thông tin ấy đến từ đâu. Trước khi có thể ghi nhớ, nó phải trải qua nhiều vòng xác nhận để biến tri thức thành ký ức. Điều máy biết vì thế không được xây dựng từ một lần đọc mà từ một chuỗi tích lũy kéo dài của các bằng chứng.

Chương được tổ chức quanh bốn chuyển động nối tiếp: phân giải, chứng thực, tổng hợp và củng cố. Phần A khảo sát quá trình hình thành tri thức, từ việc xác định thực thể đến việc thiết lập xuất xứ và thẩm quyền, rồi tới khâu bản lề là lúc máy tự viết ra một mô tả mới từ những bằng chứng đã chọn. Phần B khảo sát động lực của tri thức sau khi đã hình thành: vòng lặp củng cố, ký ức thuật toán, phụ thuộc đường dẫn, hiện tượng khóa chặt và chi phí sửa sai. Khâu tổng hợp ở cuối Phần A chính là chỗ nối giữa hai nửa, bởi những mô tả do máy sinh ra sẽ quay trở lại hệ thống dưới hình thức dữ liệu mới.

Đối với quản trị thương hiệu địa phương, ý nghĩa của chương vượt ra ngoài phạm vi kỹ thuật. Một địa phương không chỉ cạnh

tranh để được máy nhìn thấy, mà còn phải cạnh tranh để được nhận diện đúng, tin tưởng đúng và ghi nhớ đúng. Trong nền kinh tế tác nhân, một mô tả sai không chỉ là một lỗi dữ liệu. Nếu không được sửa đủ sớm, nó có thể trở thành một phần của ký ức thuật toán, lan qua nhiều hệ thống suy luận và làm tăng chi phí sửa chữa sau mỗi vòng lặp. Quản trị dữ liệu vì thế đã trở thành quản trị niềm tin và ký ức của máy.

**PHẠM VI CỦA MÔ HÌNH**

*Chương này không mô tả một thuật toán cụ thể. Nó dựng một mô hình khái niệm chung từ những cơ chế thường gặp trong các hệ thống truy hồi, đồ thị tri thức, công cụ tìm kiếm, nền tảng bản đồ, mô hình tạo sinh và tác nhân. Mỗi hệ thống thực tế có thể thực hiện các cơ chế này theo cách khác nhau, thực hiện một phần, hoặc không thực hiện. Khi viết rằng máy phân giải, máy đánh giá nguồn hay máy ghi nhớ, cần hiểu đó là mô tả một xu hướng kiến trúc chứ không phải một thuộc tính phổ quát của mọi hệ thống trí tuệ nhân tạo.*

PHẦN A

# MÁY HÌNH THÀNH TRI THỨC BẰNG CÁCH NÀO

## 4.1. PHÂN GIẢI THỰC THỂ: BÀI TOÁN CỦA MỘT CÁI TÊN

Con người có thể đọc một cái tên và gần như ngay lập tức hiểu đối tượng đang được nhắc tới. Khi nghe "Hội An", chúng ta không chỉ nghĩ đến một địa danh hành chính mà còn liên tưởng đến phố cổ, đèn lồng, sông Hoài và hàng loạt ký ức văn hóa tích lũy từ trước. Máy không có những liên tưởng ấy. Đối với một hệ thống suy luận, một cái tên trước hết chỉ là một chuỗi ký tự. Câu hỏi phải trả lời trước tiên rất cơ bản: đối tượng này là ai. "Thanh Hà" là một người, một xã, một phường, một làng vải thiều hay một doanh nghiệp? "Phú Quốc" là thành phố, hòn đảo, khu dự trữ sinh quyển hay một điểm đến du lịch?

Đó là lý do bài toán đầu tiên của một hệ thống tri thức không phải là đọc văn bản mà là phân giải thực thể, tức quá trình nhận diện rằng nhiều cách gọi khác nhau có thể chỉ cùng một đối tượng, hoặc ngược lại, cùng một tên gọi có thể đại diện, được gắn nhãn cho nhiều đối tượng hoàn toàn khác nhau. Đây là bước chuyển từ ký hiệu sang đối tượng tồn tại trong thế giới thực.

### 4.1.1. Không có thực thể thì không có tri thức

Trong ngôn ngữ hằng ngày, chúng ta thường nghĩ rằng tri thức bắt đầu từ việc đọc. Một người mở sách, đọc một bài báo hay nghe một câu chuyện rồi mới hiểu thế giới. Vì vậy, khi trí tuệ nhân tạo có khả năng xử lý hàng triệu tài liệu, nhiều người cũng hình dung rằng nó đơn giản là một "người đọc khổng lồ". Thực tế không phải như vậy.

Đối với máy, đọc văn bản chưa bao giờ là bước đầu tiên. Trước khi có thể đọc, máy phải biết mình đang đọc về đối tượng nào. Nếu một câu viết rằng "Huế đang phát triển mạnh du lịch đêm", hệ thống phải xác định "Huế" ở đây là thành phố trực thuộc Trung ương của Việt Nam, chứ không phải một người có tên Huế, một doanh nghiệp hay một địa danh khác. Nếu không xác định được đối tượng, toàn bộ phần còn lại của câu gần như mất ý nghĩa.

Đó là lý do mọi hệ thống tri thức đều phải giải quyết trước bài toán phân giải thực thể (Entity resolution). Đây là quá trình xác định rằng một tên gọi trong dữ liệu thực sự đang đề cập đến thực thể nào trong thế giới. Sau khi thực thể được xác định, hệ thống mới có thể gắn cho nó các thuộc tính, kết nối với các thực thể khác, đối chiếu thông tin từ nhiều nguồn và cuối cùng mới hình thành tri thức.

Điều này cũng lý giải vì sao nhiều hệ thống tri thức hiện đại không chỉ tổ chức thế giới như một tập hợp văn bản, mà còn như một tập hợp thực thể và quan hệ. Văn bản chỉ là phương tiện biểu đạt. Tri thức chỉ xuất hiện khi máy biết chính xác văn bản ấy đang nói về ai, về cái gì và về mối quan hệ nào.

Với địa phương cũng vậy. Trước khi một hệ thống có thể kết luận rằng một điểm đến đáng ghé thăm, một di sản có giá trị hay một doanh nghiệp đáng tin cậy, nó phải nhận diện đúng từng thực thể liên quan. Nếu bước đầu tiên này thất bại, mọi phép suy luận phía sau đều có nguy cơ sai lệch. Vì thế, có thể nói rằng không có thực thể thì không có tri thức, bởi tri thức của máy luôn bắt đầu từ việc xác định đúng đối tượng chứ không phải từ việc đọc đúng câu chữ.

• ***Mệnh đề phân giải trước (hệ thống):*** Khi bước xác định thực thể sai, các tầng suy luận phía sau vẫn tiếp tục vận hành trên đối tượng sai mà thường không có đủ tín hiệu để nhận ra nguồn gốc của sai lệch. Hệ quả là một hồ sơ mạch lạc về một thực thể không tồn tại vẫn có thể được tạo ra và được dùng như thể nó đúng.

### 4.1.2. Một cái tên không đủ để định danh

Con người thường đánh giá thấp sự nhập nhằng của tên gọi vì chúng ta luôn có ngữ cảnh để suy luận. Khi nghe đến "Long Biên", người Hà Nội có thể nghĩ ngay đến quận Long Biên, trong khi một người khác lại liên tưởng đến cầu Long Biên. Nếu cuộc trò chuyện nói về lịch sử, "Long Biên" có thể là tên một đơn vị hành chính tồn tại từ nhiều thế kỷ trước. Chỉ trong một từ ngắn, đã có nhiều thực thể hoàn toàn khác nhau cùng tồn tại.

Điều tương tự xảy ra với rất nhiều địa danh ở Việt Nam. "Huế" có thể là thành phố, tỉnh cũ trước khi điều chỉnh địa giới, quần thể di tích cố đô, thương hiệu ẩm thực hay vùng văn hóa Trung Bộ. "Thanh Hà" có thể là một phường, một xã, một huyện, một làng gốm nổi tiếng ở Hội An, một vùng trồng vải của tỉnh Hải Dương cũ hoặc tên của rất nhiều người. "Hội An" có thể chỉ thành phố, khu phố cổ được UNESCO công nhận, thương cảng lịch sử hay thương hiệu du lịch. Cùng một tên gọi nhưng mỗi thực thể lại có tọa độ, thuộc tính, lịch sử và quan hệ hoàn toàn khác nhau.

Máy không thể dựa vào trực giác để phân biệt những trường hợp ấy. Nó cần nhiều tín hiệu hơn một cái tên, như vị trí địa lý, loại hình thực thể, mã định danh, quan hệ với các thực thể khác, nguồn dữ liệu và nhiều thuộc tính bổ sung. Chỉ khi những yếu

tố đó được kết hợp, hệ thống mới có thể xác định chính xác thực thể đang được đề cập.

Đây chính là khác biệt giữa tên gọi và danh tính. Tên chỉ là một nhãn ngôn ngữ mà con người dùng để chỉ sự vật. Danh tính là tập hợp những đặc điểm giúp một thực thể trở nên duy nhất và không thể nhầm lẫn với bất kỳ thực thể nào khác. Trong thế giới của máy, tri thức không được xây dựng trên tên gọi mà trên danh tính. Một tên có thể được dùng cho hàng trăm thực thể. Ngược lại, một thực thể có thể mang nhiều tên và nhiều mã trong các hệ thống khác nhau, nhưng những dấu hiệu ấy cần cùng quy chiếu về một danh tính thống nhất.

Đối với quản trị địa phương, sự khác biệt này có ý nghĩa đặc biệt quan trọng. Một địa phương không thể giả định rằng chỉ cần có tên trên bản đồ là đủ để máy nhận diện chính xác. Trong nền kinh tế tác nhân, khả năng được phân giải đúng thành một thực thể duy nhất là điều kiện đầu tiên để được truy hồi, được chứng thực và cuối cùng được đề xuất. Trước khi cạnh tranh về chất lượng hay thương hiệu, địa phương phải giải quyết được bài toán tưởng như đơn giản nhưng lại mang tính nền tảng: một cái tên không bao giờ đủ để định danh.

### 4.1.3. Thực thể được ghép từ những gì

Nếu tên gọi chưa đủ để định danh, vậy điều gì tạo nên một thực thể? Các hệ thống tri thức hiện đại trả lời bằng một tập hợp bằng chứng, không bằng một thuộc tính duy nhất. Máy không kết luận hai thông tin đang nói về cùng một đối tượng chỉ vì chúng có cùng tên. Nó tổng hợp nhiều tín hiệu khác nhau, mỗi tín hiệu đóng góp một phần xác suất và chỉ khi tổng thể các bằng chứng đủ mạnh, hệ thống mới coi chúng là một thực thể duy nhất.

• ***Thành phần đầu tiên là tên gọi:*** Đây là dấu hiệu dễ nhận biết nhất nhưng cũng dễ gây nhầm lẫn nhất. Một thực thể có thể có nhiều tên chính thức, tên viết tắt, tên tiếng Anh, tên lịch sử, tên địa phương hay tên thương mại. Thành phố Hồ Chí Minh vẫn thường được gọi là Sài Gòn. Thành phố Huế từng được ghi là Thừa Thiên Huế trong nhiều cơ sở dữ liệu. Hội An có thể xuất hiện dưới dạng Hoi An, Hội An Ancient Town hoặc Ancient Town of Hoi An. Máy phải biết rằng các cách gọi ấy có thể cùng chỉ một thực thể.

• ***Thành phần thứ hai là địa điểm:*** Tọa độ địa lý giúp loại bỏ phần lớn sự nhập nhằng của tên gọi. Một "Thanh Hà" ở Hội An hoàn toàn khác với "Thanh Hà" ở Hải Dương. Một "Long Biên" nằm phía đông sông Hồng khác với cây cầu Long Biên bắc qua con sông ấy. Vị trí không chỉ là kinh độ và vĩ độ, mà còn bao gồm quốc gia, tỉnh, thành phố, quận, phường và các đơn vị hành chính liên quan. Địa điểm đặt một thực thể vào đúng bối cảnh không gian của nó.

• ***Thành phần thứ ba là danh mục:*** Máy cần biết đối tượng đang được nói tới thuộc loại nào: một địa phương, một khách sạn, một bảo tàng, một nhà hàng, một di tích, một lễ hội hay một doanh nghiệp. Cùng một tên nhưng thuộc các danh mục khác nhau sẽ tạo nên các thực thể khác nhau. Việc phân loại này chính là bước kết nối giữa thực thể và bản thể luận (ontology) đã được trình bày ở chương trước.

• ***Thành phần thứ tư là thuộc tính:*** Một khách sạn có địa chỉ, số phòng, hạng sao và giờ nhận phòng. Một di tích có niên đại, cấp xếp hạng và cơ quan quản lý. Một lễ hội có thời gian tổ chức, địa điểm diễn ra và chu kỳ lặp lại. Những thuộc tính này không chỉ giúp mô tả thực thể mà còn giúp phân biệt nó với các thực thể khác có tên gần giống.

• ***Thành phần thứ năm là quan hệ:*** Trong thế giới của máy, rất ít thực thể tồn tại một mình. Một làng nghề thuộc một xã, xã thuộc một huyện, huyện thuộc một tỉnh. Một khách sạn nằm gần một điểm tham quan. Một lễ hội được tổ chức tại một quảng trường. Một doanh nghiệp quản lý nhiều cơ sở lưu trú. Chính mạng lưới quan hệ này giúp máy hiểu rằng các thực thể là những nút trong một mạng lưới tri thức rộng lớn chứ không phải các hòn đảo rời rạc.

• ***Cuối cùng là định danh:*** Đây thường là lớp bằng chứng có khả năng neo thực thể mạnh nhất, với điều kiện mã được cấp đúng và duy trì ổn định. Một mã sai lại làm cho lỗi hợp nhất trở nên khó phát hiện hơn. Mỗi hệ thống thường gán cho thực thể một mã định danh riêng: mã địa giới hành chính, mã doanh nghiệp, mã của OpenStreetMap, Wikidata, Google hay các cơ sở dữ liệu chuyên ngành. Tên có thể thay đổi, thuộc tính có thể được cập nhật, nhưng mã định danh là sợi chỉ giúp các hệ thống nhận ra rằng mọi thay đổi ấy vẫn đang nói về cùng một đối tượng.

Nhìn dưới góc độ này, một thực thể không phải là một cái tên hay một bản ghi dữ liệu. Nó là một tập hợp bằng chứng hội tụ về cùng một đối tượng trong thế giới thực. Càng nhiều bằng chứng độc lập cùng chỉ về một thực thể, khả năng nhận diện của máy càng chính xác. Đây cũng là lý do các hệ thống suy luận hiện đại không ngừng đối chiếu và bổ sung dữ liệu từ nhiều nguồn khác nhau. Chúng không tìm kiếm thêm thông tin chỉ để biết nhiều hơn, mà để giảm sự bất định trong việc xác định danh tính của thực thể.

Đối với địa phương, điều này mang một ý nghĩa chiến lược. Xây dựng hiện diện số không chỉ là tạo thêm nội dung, mà còn là làm giàu tập bằng chứng về từng thực thể của địa phương. Mỗi

địa danh, di tích, doanh nghiệp, lễ hội hay sản phẩm đặc trưng đều cần được mô tả đầy đủ, liên kết chặt chẽ và định danh nhất quán. Khi tập bằng chứng đủ mạnh, máy mới có thể nhận diện đúng và sử dụng chúng làm nền tảng cho mọi bước suy luận tiếp theo. Bảng 4.1 tổng hợp sáu lớp bằng chứng ấy.

**Bảng 4.1.** Sáu lớp bằng chứng cấu thành một thực thể

| Lớp bằng chứng | Câu hỏi nó trả lời | Rủi ro khi thiếu |
|---|---|---|
| **Tên gọi** | Đối tượng được gọi là gì | Trùng tên, nhập nhằng, tách một thực thể thành nhiều |
| **Địa điểm** | Đối tượng nằm ở đâu | Không loại được các trùng tên khác tỉnh, gắn sai tọa độ |
| **Danh mục** | Đối tượng thuộc loại nào | Bị xếp sai nhóm nên không xuất hiện trong tập ứng viên |
| **Thuộc tính** | Đối tượng có đặc điểm gì | Không đủ dữ kiện để máy so sánh và xếp hạng |
| **Quan hệ** | Đối tượng gắn với những gì | Bị tách khỏi mạng lưới, mất toàn bộ suy luận bắc cầu |
| **Định danh** | Đâu là mã duy nhất của đối tượng | Mỗi hệ thống hiểu một kiểu, không hợp nhất được |

### MỆNH ĐỀ DANH TÍNH HỘI TỤ

*Một thực thể được nhận diện không nhờ tên gọi mà nhờ sự hội tụ của sáu lớp bằng chứng: tên gọi, địa điểm, danh mục, thuộc tính, quan hệ và định danh. Khi các lớp ấy cùng quy chiếu về một đối tượng, danh tính số trở nên bền. Khi chúng phân kỳ, thực thể có nguy cơ bị tách đôi hoặc bị hợp nhất nhầm.*

### 4.1.4. Sai ở bước phân giải sẽ sai toàn bộ chuỗi

Trong khoa học máy tính có một nguyên tắc nổi tiếng: "Garbage in, garbage out", nghĩa là dữ liệu đầu vào sai sẽ tạo ra kết quả đầu ra sai. Đối với các hệ thống trí tuệ nhân tạo hiện đại, nguyên tắc ấy cần được mở rộng thêm một bước, nếu phân giải thực thể sai, toàn bộ chuỗi suy luận phía sau cũng sẽ sai. Vấn đề không còn chỉ là chất lượng dữ liệu, mà là chất lượng của việc xác định đối tượng ngay từ điểm xuất phát.

Hãy hình dung một tác nhân AI được giao nhiệm vụ tìm khách sạn gần bến thuyền Tràng An ở Ninh Bình. Nếu hệ thống nhầm lẫn giữa hai khách sạn có tên gần giống nhau, mọi bước tiếp theo, từ đọc đánh giá, so sánh giá phòng, kiểm tra tình trạng còn phòng cho đến đặt chỗ, đều được thực hiện trên sai đối tượng. Khả năng suy luận của AI có thể rất mạnh, nhưng toàn bộ lập luận ấy chỉ là một lời giải hoàn hảo cho một bài toán đã bị đặt sai ngay từ đầu.

Điều tương tự xảy ra với doanh nghiệp. Một nhà hàng có nhiều chi nhánh, một khu nghỉ dưỡng đổi chủ sở hữu hoặc một thương hiệu sử dụng nhiều tên thương mại khác nhau đều có thể khiến hệ thống hợp nhất hoặc tách nhầm thực thể. Khi đó, đánh giá của khách hàng, dữ liệu đặt chỗ, hình ảnh và thông tin liên hệ có thể bị gán sai, dẫn đến những kết luận không còn phản ánh đúng thực tế.

Ở quy mô địa phương, hậu quả còn lớn hơn. Một địa danh có thể bị nhầm với địa danh cùng tên ở tỉnh khác. Một di tích được gắn nhầm tọa độ. Một lễ hội bị đồng nhất với một sự kiện hoàn toàn khác. Một làng nghề bị xem như một doanh nghiệp riêng lẻ. Sai sót ban đầu ấy sẽ lan sang các hệ thống bản đồ, công cụ tìm kiếm, mô hình ngôn ngữ và các tác nhân AI sử dụng chúng làm nguồn

tham chiếu. Mỗi hệ thống tiếp tục tạo ra mô tả mới dựa trên dữ liệu đã sai, khiến sai lệch được khuếch đại qua từng vòng truy hồi.

Điều đáng chú ý là năng lực suy luận không thể tự sửa chữa một thực thể bị nhận diện sai. Một hệ thống có thể lập luận rất tinh vi về giá trị của một di sản, nhưng nếu ngay từ đầu nó nhầm di sản này với một địa điểm khác thì mọi lập luận sau đó đều mất giá trị. Sức mạnh của trí tuệ nhân tạo không bù đắp được sai lầm ở tầng định danh. Ngược lại, nó có thể khiến sai lầm ấy lan truyền nhanh hơn và thuyết phục hơn.

Vì vậy, trong nền kinh tế tác nhân, bài toán quan trọng nhất không phải là làm cho AI suy luận tốt hơn, mà là bảo đảm AI bắt đầu từ đúng thực thể. Phân giải thực thể không phải một bước tiền xử lý kỹ thuật có thể xem nhẹ. Nó là nền móng của toàn bộ kiến trúc tri thức. Nếu nền móng ấy lệch đi một vài centimet, toàn bộ công trình suy luận xây phía trên sẽ nghiêng theo. Chính vì vậy, có thể diễn đạt lại nguyên tắc cũ bằng một mệnh đề phù hợp hơn với thời đại trí tuệ nhân tạo: phân giải sai, suy luận sai (Wrong entity in, wrong reasoning out).

Nếu suy luận đủ mạnh, máy sẽ tự sửa được lỗi phân giải. Các mô hình ngày càng có khả năng đọc ngữ cảnh dài và tự phát hiện mâu thuẫn. Một hệ thống đủ tinh vi lẽ ra phải nhận ra rằng nó đang nhầm hai đối tượng khác nhau. Suy luận chỉ sửa được những mâu thuẫn mà nó nhìn thấy. Khi hai thực thể bị hợp nhất từ tầng dữ liệu, các thuộc tính mâu thuẫn của chúng biến mất trước khi mô hình kịp so sánh, vì hệ thống đã coi chúng là thuộc tính của cùng một đối tượng. Lỗi phân giải không tạo ra tín hiệu bất thường mà tạo ra một hồ sơ trông rất mạch lạc về một thực thể không tồn tại. Năng lực suy luận trong trường hợp này làm cho kết quả sai trở nên thuyết phục hơn chứ không đúng hơn.

### 4.1.5. Địa phương phải tự giải bài toán phân giải

Nếu phân giải thực thể là điều kiện đầu tiên để máy hình thành tri thức, thì một câu hỏi tự nhiên xuất hiện: ai sẽ làm công việc ấy? Liệu các nền tảng số, các mô hình trí tuệ nhân tạo hay các công cụ tìm kiếm sẽ tự động nhận diện chính xác mọi địa phương? Câu trả lời là không.

Các hệ thống suy luận hiện đại có thể hợp nhất dữ liệu từ hàng triệu nguồn, nhưng chúng chỉ có thể làm việc với những gì đã tồn tại. Máy không thể tự phát minh ra một định danh chuẩn nếu không có định danh. Máy không thể tự xây dựng một bản thể luận chính xác nếu không có ai định nghĩa các khái niệm. Máy cũng không thể tự tạo ra những mối liên kết giữa các thực thể nếu dữ liệu ban đầu chưa từng được thiết kế theo cách đó. Càng thiếu cấu trúc ở đầu vào, hệ thống càng phải dựa vào suy đoán thống kê và xác suất sai lệch càng lớn.

Điều đó có nghĩa là bài toán phân giải không thể được giao hoàn toàn cho các nền tảng công nghệ. Chính địa phương phải chủ động xây dựng hạ tầng nhận diện của mình. Nếu muốn được máy đọc đúng, địa phương phải tạo ra đủ bằng chứng để các hệ thống khác nhau có xu hướng đi đến cùng một kết luận về từng thực thể.

Nền tảng đầu tiên là định danh. Mỗi thực thể quan trọng của địa phương, từ đơn vị hành chính, di tích, điểm tham quan, làng nghề, lễ hội cho tới doanh nghiệp và sản phẩm đặc trưng, cần có một danh tính ổn định và duy nhất. Danh tính ấy không chỉ phục vụ quản lý nội bộ mà còn giúp các hệ thống bên ngoài nhận ra rằng những dữ liệu xuất hiện ở nhiều nguồn khác nhau thực chất đang nói về cùng một đối tượng.

Trên nền tảng đó là bản thể luận, tức tập các loại thực thể tồn tại trong hệ sinh thái của địa phương và quan hệ giữa chúng. Kế đến là lược đồ dữ liệu, quy định mỗi loại thực thể cần có những trường thông tin nào. Chương 3 đã trình bày hai lớp này, nên ở đây chỉ cần nhắc lại một điểm: cả hai đều là quyết định quản trị chứ không phải lựa chọn kỹ thuật, bởi chúng xác định địa phương được xếp vào loại nào, mà loại đó quyết định địa phương xuất hiện trong tập ứng viên nào.

Hai lớp còn lại là đồ thị tri thức, nơi các thực thể được nối với nhau thành mạng lưới thay vì nằm rời rạc. Lớp cuối là mã định danh chuẩn, điểm neo giúp mọi hệ thống quy chiếu về cùng một đối tượng bất kể tên gọi hay ngôn ngữ.

Điều Chương 4 bổ sung nằm ở phần vận hành các lớp ấy theo thời gian. Một hạ tầng nhận diện chỉ giữ được giá trị nếu địa phương duy trì được bốn việc: giữ ánh xạ về mã cũ mỗi khi đổi tên hoặc điều chỉnh địa giới, quản lý phiên bản để biết mỗi thuộc tính đã thay đổi khi nào, chỉ định rõ nguồn nào có thẩm quyền với loại thông tin nào, xử lý trùng lặp trước khi các hệ thống bên ngoài kịp hợp nhất sai. Không có bốn việc này, một bộ dữ liệu được xây dựng công phu vẫn có thể mất toàn bộ lịch sử tích lũy chỉ sau một lần nâng cấp hệ thống.

Nhìn từ góc độ quản trị, những thành phần này không còn là các khái niệm thuần túy của khoa học dữ liệu. Chúng tạo thành hạ tầng nhận diện của địa phương trong nền kinh tế tác nhân. Nếu đường sá, điện và nước là hạ tầng của nền kinh tế công nghiệp, thì định danh, bản thể luận, lược đồ dữ liệu, đồ thị tri thức và mã định danh chuẩn chính là hạ tầng của nền kinh tế nơi máy trở thành chủ thể đọc và lựa chọn. Vì thế, nhiệm vụ của địa phương không phải là chờ các mô hình trí tuệ nhân tạo hiểu

mình tốt hơn. Nhiệm vụ của địa phương là xây dựng một danh tính số đủ rõ ràng để các hệ thống khác nhau có xu hướng hội tụ về cùng một kết luận. Chỉ khi bài toán phân giải được giải quyết ngay từ đầu, các bước tiếp theo như chứng thực, xếp hạng và ủy nhiệm mới có thể vận hành trên một nền tảng đáng tin cậy.

*Nếu bạn là người phụ trách chuyển đổi số của tỉnh:*

- Lập danh mục các thực thể lõi cần định danh trước, thường không quá hai trăm mục, bao gồm: đơn vị hành chính, di tích xếp hạng, lễ hội định kỳ, làng nghề, sản phẩm chỉ dẫn địa lý, cơ sở lưu trú đạt chuẩn.
- Chọn một mã định danh chuẩn cho mỗi thực thể và công bố công khai bảng ánh xạ sang các mã đang lưu hành ở những hệ thống khác.
- Yêu cầu mọi cổng thông tin của tỉnh nhúng dữ liệu có cấu trúc kèm mã định danh, thay vì chỉ đăng bài viết dạng văn xuôi.
- Đặt một đầu mối duy nhất chịu trách nhiệm khi hai hệ thống của tỉnh mô tả cùng một thực thể theo hai cách khác nhau.

Lấy một ví dụ dựng lại để theo suốt chương. Đây là một trường hợp giả định, ghép từ những kiểu lỗi thường gặp, không phải một địa phương có thật. Giả sử có một làng nghề gốm mang hai tên song song: tên hành chính của thôn và tên nghề mà người dân quen dùng. Cổng thông tin huyện đăng theo tên hành chính, các bài báo du lịch dùng tên nghề, còn trên nền tảng bản đồ thì tọa độ được người dùng gắn nhầm sang một khu dân cư cách đó bốn cây số. Không bản ghi nào có mã định danh. Kết quả là các hệ thống bên ngoài không nhìn thấy một làng nghề mà nhìn thấy hai thực thể mờ nhạt, mỗi thực thể chỉ có một nửa bằng chứng.

## 4.2. XUẤT XỨ, ĐỐI CHIẾU VÀ THẨM QUYỀN

Nếu mục 4.1 trả lời câu hỏi "máy biết một thực thể là ai", thì tiết này trả lời câu hỏi tiếp theo: "vì sao máy tin rằng điều đó là đúng". Trong thế giới của con người, niềm tin có thể hình thành từ trực giác, kinh nghiệm hay uy tín cá nhân. Chúng ta có thể tin lời một người dân địa phương lâu năm, một hướng dẫn viên nhiều kinh nghiệm hay một chuyên gia đầu ngành mà không cần kiểm chứng từng chi tiết. Nhưng máy không có trực giác. Một hệ thống AI không biết ai đáng tin nếu không có bằng chứng để đánh giá. Vì vậy, sau khi một thực thể được phân giải thành công, bài toán tiếp theo là xác lập thẩm quyền của thông tin. Một khách sạn có thật không. Một lễ hội còn diễn ra hay đã dừng tổ chức. Một nhà hàng đã đóng cửa hay vẫn hoạt động. Một di tích được xếp hạng cấp quốc gia hay chỉ là một cách gọi trong cộng đồng. Mọi phát biểu đều phải đi kèm với nguồn gốc và khả năng kiểm chứng.

Điều này khiến khái niệm nguồn dữ liệu trở nên quan trọng hơn bao giờ hết. Trong những hệ thống có cơ chế truy hồi, xếp hạng và truy vết nguồn, máy không chỉ xử lý nội dung mà còn có thể dùng thông tin về nguồn để đánh giá bằng chứng. Hai câu giống hệt nhau nhưng xuất hiện trên một cổng thông tin chính thức, trong một cơ sở dữ liệu được chuẩn hóa hay trên một bài đăng vô danh sẽ có trọng lượng hoàn toàn khác nhau. Đối với máy, thông tin không chỉ mang nội dung mà còn mang theo lịch sử hình thành, mức độ đáng tin cậy và dấu vết xác minh của nó. Bởi vậy, phần lớn hệ thống suy luận hiện đại vận hành dựa trên ba câu hỏi liên tiếp. Thông tin đến từ đâu? Có được đối chiếu với các nguồn khác hay không? Và nguồn đó có đủ thẩm quyền để được sử dụng trong quá trình suy luận không? Chỉ khi trả lời được cả ba câu hỏi này, dữ liệu mới có thể trở thành bằng chứng.

Đối với quản trị địa phương, đây là một thay đổi có ý nghĩa chiến lược. Trong nền kinh tế tác nhân, cạnh tranh không chỉ là có nhiều dữ liệu hơn, mà còn là có dữ liệu đáng tin hơn. Một địa phương có thể sở hữu hàng nghìn bài viết quảng bá, nhưng nếu thiếu nguồn chính thức, thiếu cơ chế xác thực và thiếu khả năng đối chiếu với các hệ thống tri thức khác, AI vẫn sẽ coi đó là thông tin có độ tin cậy thấp. Ngược lại, một tập dữ liệu nhỏ nhưng có nguồn gốc rõ ràng, được nhiều hệ thống độc lập xác nhận và được cập nhật thường xuyên lại có giá trị lớn hơn rất nhiều.

Mục này vì thế không bàn về kỹ thuật kiểm chứng dữ liệu đơn thuần, mà đi sâu vào kiến trúc của niềm tin số. Bởi trong thời đại AI, điều quyết định một địa phương có được đọc hay không không chỉ là khả năng được nhận diện, mà còn là khả năng được tin tưởng. Nếu phân giải tạo ra danh tính số, thì nguồn gốc, đối chiếu và thẩm quyền chính là những yếu tố tạo ra uy tín số của địa phương trong trí nhớ của máy.

### 4.2.1. Muốn biết điều gì đúng, trước hết phải biết nó đến từ đâu

Khi con người tiếp nhận một thông tin, câu hỏi đầu tiên thường là: "Điều này có đúng không?" Chúng ta đối chiếu với hiểu biết, kinh nghiệm hoặc trực giác của mình để đánh giá tính xác thực.

Máy không làm như vậy. Một hệ thống AI không có trải nghiệm sống để tự phán đoán đúng sai. Nó không biết Sa Pa có đẹp hay không, một lễ hội có nổi tiếng hay không, hay một món ăn có thật sự đặc sắc hay chỉ là lời quảng cáo. Trước khi đánh giá nội dung, máy phải trả lời một câu hỏi cơ bản hơn nhiều: *"Ai nói điều này?"*. Nếu thông tin đến từ cổng dữ liệu của cơ quan quản lý, trọng số sẽ

khác. Nếu đến từ một bài báo chính thống, trọng số sẽ khác. Nếu được nhiều nguồn độc lập cùng xác nhận, mức độ tin cậy lại tăng lên. Ngược lại, một phát biểu chỉ xuất hiện trên một bài đăng vô danh hoặc một trang web ít được biết đến sẽ có giá trị chứng cứ thấp hơn, dù nội dung có thể hoàn toàn chính xác.

Đó là lý do trong khoa học dữ liệu và trí tuệ nhân tạo xuất hiện khái niệm xuất xứ dữ liệu (provenance). Xuất xứ mô tả toàn bộ lịch sử của một dữ liệu, nó được tạo ra bởi ai, vào thời điểm nào, từ nguồn nào, đã được chỉnh sửa ra sao và đã đi qua những hệ thống nào trước khi đến tay người sử dụng. Nói cách khác, nếu dữ liệu là một vật chứng thì xuất xứ dữ liệu chính là chuỗi bảo quản chứng cứ của dữ liệu đó.

Một tài liệu không chỉ có nội dung mà còn có lý lịch. Một bức ảnh không chỉ có hình ảnh mà còn có nơi chụp, thời gian chụp và người công bố. Một địa danh không chỉ có tên gọi mà còn có cơ quan định danh, hồ sơ pháp lý và các nguồn xác nhận độc lập. Đối với AI, những thông tin này đôi khi quan trọng không kém chính nội dung mà chúng mô tả.

Trong những hệ thống có khả năng truy vết nguồn, xuất xứ có thể trở thành một lớp thông tin nền của quá trình đánh giá bằng chứng. Khi tổng hợp câu trả lời từ nhiều nguồn, hệ thống không chỉ đọc văn bản mà còn có thể sử dụng dấu vết về đường đi của tri thức. Thông tin xuất phát từ đâu, được ai xác nhận, có còn mới hay đã lỗi thời, có mâu thuẫn với các nguồn khác hay không, tất cả đều ảnh hưởng đến việc thông tin đó có được sử dụng làm bằng chứng hay bị loại bỏ.

Đối với địa phương, điều này tạo ra một yêu cầu hoàn toàn mới. Không còn đủ để công bố thông tin. Cần phải công bố theo

cách mà nguồn gốc của thông tin có thể được máy nhận diện và kiểm chứng. Một trang web chính thức, dữ liệu mở có định danh rõ ràng, tài liệu có ngày cập nhật, cơ quan chịu trách nhiệm và lịch sử thay đổi minh bạch đều góp phần xây dựng xuất xứ cho địa phương.

Có thể xem đây là sự chuyển dịch từ nội dung sang nguồn gốc của nội dung. Trong nền kinh tế tác nhân, điều AI tin tưởng không chỉ là một mệnh đề, mà là toàn bộ chuỗi bằng chứng đứng phía sau mệnh đề ấy. Vì vậy, trước khi hỏi một thông tin có đúng hay không, máy luôn hỏi trước một câu tưởng như đơn giản nhưng mang tính quyết định: *"Ai là người đã nói điều này và vì sao tôi nên tin họ?".*

*Mệnh đề xuất xứ:* Trong những hệ thống có khả năng truy vết nguồn, hai mệnh đề giống hệt nhau về câu chữ vẫn nhận trọng số khác nhau nếu xuất xứ khác nhau. Hệ quả là giá trị của một thông tin không nằm trọn trong nội dung của nó mà một phần nằm ở nơi nó được phát ra.

### 4.2.2. Một thông tin mạnh vì nhiều nguồn cùng xác nhận

Trong đời sống hằng ngày, khi nhiều người độc lập cùng kể một câu chuyện giống nhau, chúng ta có xu hướng tin câu chuyện ấy hơn. Không phải vì số đông luôn đúng, mà vì nhiều bằng chứng độc lập làm giảm khả năng sai lệch. Máy cũng vận hành theo một logic tương tự. Một hệ thống có nhiều nguồn để đối chiếu hiếm khi đặt niềm tin tuyệt đối vào một nguồn duy nhất. Thay vào đó, nó liên tục đối chiếu thông tin giữa nhiều nguồn khác nhau để tìm kiếm sự nhất quán. Một địa danh xuất hiện trên trang web chính thức của địa phương là một tín hiệu. Nếu cùng địa danh ấy cũng

có trên OpenStreetMap, Wikipedia, Google Maps, báo chí chính thống và cơ sở dữ liệu của cơ quan quản lý, mức độ tin cậy sẽ tăng lên đáng kể. Mỗi nguồn không chỉ bổ sung thông tin mà còn đóng vai trò như một người làm chứng cho những nguồn còn lại.

Đó chính là nguyên lý của xác minh chéo giữa các nguồn, trong tài liệu tiếng Anh thường gọi là "source corroboration". Cần phân biệt nó với "cross validation" trong học máy, vốn là kỹ thuật chia dữ liệu để đánh giá mô hình và không liên quan đến bài toán ở đây. Xác minh chéo trong quản trị tri thức là quá trình so sánh cùng một mệnh đề trên nhiều nguồn khác nhau để xem mức độ đồng thuận đến đâu.

Hãy hình dung một lễ hội truyền thống của địa phương. Nếu lễ hội chỉ được nhắc đến trên một trang web quảng bá du lịch, AI sẽ xem đây là một thông tin cần thận trọng. Nhưng nếu lễ hội đồng thời xuất hiện trên cổng thông tin của Ủy ban nhân dân tỉnh, trong cơ sở dữ liệu của Bộ Văn hóa, trên Wikipedia, được định vị trên Google Maps, có bài viết của báo chí chính thống và được cộng đồng chia sẻ bằng hình ảnh, đánh giá và lịch sử nhiều năm liên tiếp, AI sẽ có nhiều cơ sở hơn để kết luận rằng đây là một sự kiện có thật, đang tồn tại và có giá trị.

Điều quan trọng là các nguồn này không hoàn toàn giống nhau về chức năng. Trang web chính thức cung cấp thông tin gốc. Cơ quan quản lý xác nhận tính pháp lý. OpenStreetMap bổ sung dữ liệu không gian. Google tổng hợp và đối chiếu từ nhiều nguồn khác nhau. Wikipedia đóng vai trò như một bách khoa toàn thư cộng đồng có cơ chế biên tập. Báo chí cung cấp bối cảnh và sự kiện. Cộng đồng mang lại bằng chứng thực tế thông qua hình ảnh, đánh giá và trải nghiệm. Chính sự đa dạng ấy tạo nên một mạng lưới chứng cứ phong phú hơn bất kỳ nguồn đơn lẻ nào.

Điều này cũng giải thích vì sao một trang web đẹp nhưng đứng một mình vẫn là một trang web yếu. Một địa phương có thể đầu tư rất lớn cho cổng thông tin điện tử, nhưng nếu thông tin trên đó không được các hệ sinh thái tri thức khác tham chiếu, liên kết và xác nhận, AI vẫn khó có đủ cơ sở để xem đó là bằng chứng mạnh. Trong nền kinh tế tác nhân, sức mạnh của thông tin không chỉ đến từ chất lượng của từng nguồn mà còn đến từ mức độ kết nối giữa các nguồn.

Vì vậy, quản trị dữ liệu địa phương không còn là quản trị từng cơ sở dữ liệu riêng lẻ. Đó là quá trình xây dựng một hệ sinh thái chứng thực, nơi cùng một thực thể được phản ánh nhất quán trên nhiều nền tảng, nhiều tổ chức và nhiều cộng đồng khác nhau. Khi đó, mỗi nguồn không còn hoạt động độc lập mà trở thành một mắt xích trong mạng lưới xác nhận lẫn nhau. Có thể tóm tắt nguyên lý này bằng một mệnh đề đơn giản: Trong thời đại AI, một thông tin không mạnh vì được lặp lại nhiều lần, mà mạnh vì được nhiều nguồn độc lập cùng xác nhận. Đó cũng là sự chuyển dịch từ tư duy công bố thông tin sang tư duy xây dựng mạng lưới bằng chứng. Một địa phương càng có nhiều nguồn đáng tin cùng mô tả nhất quán về mình, địa phương đó càng có khả năng được AI tin tưởng, truy hồi và lựa chọn trong quá trình suy luận.

**MỆNH ĐỀ ĐỐI CHIẾU ĐỘC LẬP**

*Khi nhiều nguồn độc lập cùng xác nhận một mệnh đề, độ tin cậy của mệnh đề thường tăng. Việc lặp lại trong các nguồn phụ thuộc lẫn nhau không tạo ra mức tăng tương đương, dù số lần xuất hiện có thể lớn hơn nhiều.*

Vấn đề là tính độc lập của nguồn không phải lúc nào cũng quan sát được. Mười trang web có thể trông như mười nhân chứng, trong khi thực chất chỉ là mười bản sao của cùng một mệnh đề ban đầu. Máy nhìn thấy số lần xuất hiện dễ hơn nhiều so với nhìn thấy quan hệ phụ thuộc giữa các nguồn. Chính khoảng mù ấy là điều khiến vòng lặp được bàn ở Phần B trở nên nguy hiểm.

### 4.2.3. Mọi nguồn đều có trọng số

Trong thế giới của máy, không tồn tại khái niệm "mọi nguồn đều đáng tin như nhau". Hai câu văn có thể giống hệt nhau. Hai bộ dữ liệu có thể chứa cùng một thông tin. Nhưng nếu chúng đến từ hai nguồn khác nhau, giá trị của chúng trong quá trình suy luận có thể hoàn toàn khác nhau. Điều này phản ánh một nguyên lý cơ bản của mọi hệ thống tri thức, không phải mọi dữ liệu đều bình đẳng.

Trong nghiên cứu khoa học, người ta phân biệt giữa nguồn sơ cấp (primary source) và nguồn thứ cấp (secondary source). Nguồn sơ cấp là nơi thông tin được tạo ra lần đầu tiên: một văn bản pháp luật, một báo cáo điều tra, một cơ sở dữ liệu gốc, một quyết định hành chính, một bản đồ chính thức hay kết quả đo đạc trực tiếp. Nguồn thứ cấp là những tài liệu diễn giải, tổng hợp hoặc trích dẫn lại các nguồn sơ cấp như sách, bài báo phân tích hay các bài viết chuyên khảo.

Trong môi trường số, còn xuất hiện một nhóm dữ liệu ngày càng quan trọng là nội dung do người dùng tạo ra (User Generated Content - UGC): đánh giá trên Google Maps, bài viết trên mạng xã hội, hình ảnh, video, bình luận hay trải nghiệm cá nhân. Đây là nguồn dữ liệu có quy mô khổng lồ và phản ánh thực tế rất nhanh, nhưng cũng chứa nhiều sai lệch, thiên kiến và thông tin

chưa được kiểm chứng. Thay vì loại bỏ hẳn một loại nguồn, hệ thống thường gán cho mỗi nguồn một trọng số khác nhau.

Một quyết định công nhận di tích của cơ quan nhà nước thường có giá trị xác nhận về mặt pháp lý cao hơn một bài viết trên blog. Một bài báo khoa học phản biện thường có trọng số chuyên môn cao hơn một bài quảng cáo. Một nghìn đánh giá của người dùng về chất lượng phục vụ có thể đáng tin hơn một lời giới thiệu từ chính doanh nghiệp. Ngược lại, khi đánh giá cảm nhận của du khách, trải nghiệm thực tế của cộng đồng lại có giá trị mà các văn bản hành chính không thể thay thế.

Nói cách khác, giá trị của một nguồn phụ thuộc vào câu hỏi đang được đặt ra. Nếu AI cần biết một khách sạn có được cấp phép hay không, cơ quan quản lý là nguồn có thẩm quyền cao nhất. Nếu AI cần biết khách sạn có sạch sẽ và phục vụ tốt hay không, hàng nghìn đánh giá của khách hàng lại trở thành bằng chứng quan trọng hơn. Thẩm quyền là một quan hệ giữa nguồn thông tin và mục đích sử dụng thông tin, chứ không phải một thuộc tính tuyệt đối.

Đó là lý do nhiều hệ thống hiện đại xây dựng một mô hình về thẩm quyền (authority) và mức độ tin cậy (trust). Thẩm quyền phản ánh vị thế và khả năng xác nhận của nguồn phát thông tin. Mức độ tin cậy phản ánh mức độ mà hệ thống có thể dựa vào nguồn đó trong từng ngữ cảnh cụ thể. Một nguồn có thẩm quyền cao nhưng đã lỗi thời có thể bị giảm trọng số. Một nguồn từ cộng đồng nhưng liên tục được xác nhận là chính xác có thể dần tích lũy được mức độ tin cậy lớn hơn.

Đối với địa phương, điều này mang ý nghĩa chiến lược. Không chỉ cần tạo ra dữ liệu, mà còn phải xây dựng hệ sinh thái các nguồn có thẩm quyền. Cổng thông tin chính thức, cơ sở dữ liệu

mở, hồ sơ pháp lý, bản đồ chuẩn hóa, báo chí uy tín, các tổ chức chuyên môn và cộng đồng người dùng đều có vai trò khác nhau trong việc hình thành niềm tin của AI. Mỗi nguồn giống như một nhân chứng với mức độ uy tín riêng. Khi các nhân chứng đáng tin cùng kể một câu chuyện nhất quán, xác suất được AI tin tưởng sẽ tăng lên đáng kể.

Có thể xem đây là một quy luật của nền kinh tế tác nhân: Dữ liệu không chỉ có nội dung mà còn có địa vị. Nguồn càng có thẩm quyền, thông tin càng có trọng lượng trong quá trình suy luận của máy. Vì vậy, quản trị dữ liệu địa phương không chỉ là quản lý dữ liệu, mà còn là quản trị thẩm quyền của dữ liệu. Trong một thế giới nơi AI ngày càng trở thành người ra quyết định đầu tiên, ai sở hữu những nguồn thông tin đáng tin cậy hơn sẽ có lợi thế lớn hơn trong việc định hình cách máy nhìn nhận và lựa chọn địa phương. Bảng 4.2 cho thấy trọng số của một nguồn thay đổi ra sao theo loại câu hỏi.

**Bảng 4.2.** Trọng số nguồn thay đổi theo loại câu hỏi

| Câu hỏi đang được đặt ra | Nguồn có trọng số cao nhất | Nguồn có trọng số thấp |
|---|---|---|
| **Cơ sở này có được cấp phép không?** | Cơ quan quản lý nhà nước | Đánh giá của người dùng |
| **Di tích này được xếp hạng cấp nào?** | Quyết định công nhận chính thức | Bài viết du lịch |
| **Giờ mở cửa hôm nay?** | Kênh do chính chủ thể cập nhật | Bài viết đã đăng từ nhiều năm trước |
| **Trải nghiệm thực tế ra sao?** | Tập đánh giá của người dùng | Tài liệu quảng bá |
| **Số liệu thống kê năm gần nhất?** | Cơ quan thống kê | Báo chí dẫn lại không nêu nguồn |
| **Sự kiện này có diễn ra không** | Thông báo chính thức mới nhất | Lịch sự kiện tổng hợp tự động |

Tuy nhiên, xếp trọng số theo thẩm quyền chỉ là hợp thức hóa tiếng nói của bên mạnh. Nếu nguồn nhà nước và tổ chức lớn luôn được ưu tiên thì tri thức cộng đồng, tiếng nói của người dân và của các chủ thể nhỏ sẽ bị đẩy ra rìa. Cơ chế trọng số vì thế tái sản xuất bất bình đẳng sẵn có.

Phản bác này đúng ở phần mô tả và sai ở phần suy ra. Trọng số không phải là một thứ hạng chung mà là một hàm của câu hỏi, như Bảng 4.2 cho thấy. Với câu hỏi pháp lý, cơ quan quản lý mạnh hơn. Với câu hỏi về trải nghiệm, tập đánh giá của người dùng mạnh hơn bất kỳ văn bản hành chính nào. Rủi ro thực sự không nằm ở việc có trọng số mà nằm ở việc một cộng đồng không có kênh nào để bằng chứng của mình được ghi nhận dưới dạng máy đọc được. Đó là lý do việc mở kênh chứng thực cho các chủ thể nhỏ là một biện pháp phân phối chứ không chỉ là một biện pháp kỹ thuật.

### 4.2.4. Thẩm quyền là tài sản tích lũy

Trong thế giới vật chất, uy tín của một con người không thể mua ngay trong một ngày. Một nhà khoa học cần nhiều năm nghiên cứu. Một trường đại học cần nhiều thập kỷ đào tạo. Một tờ báo cần hàng nghìn bài viết chính xác để xây dựng niềm tin. Một thương hiệu cần nhiều năm giữ lời hứa với khách hàng.

Thẩm quyền của dữ liệu cũng vận hành theo logic ấy. Một trang web vừa được tạo hôm nay hiếm khi ngay lập tức trở thành nguồn tham chiếu có trọng lượng. Một cơ sở dữ liệu mới công bố cũng hiếm khi lập tức được các hệ thống khác đối chiếu tới. Thẩm quyền là một tài sản được tích lũy, không phải một thuộc tính được gán.

- ***Sự tích lũy đó diễn ra qua ba yếu tố quan trọng:***

• ***Yếu tố thứ nhất là thời gian.*** *Một nguồn dữ liệu tồn tại lâu dài, liên tục được duy trì và không ngừng được sử dụng sẽ dần hình thành lịch sử của chính nó. Lịch sử ấy là bằng chứng rằng nguồn thông tin không chỉ xuất hiện nhất thời rồi biến mất.*
• ***Yếu tố thứ hai là độ ổn định.*** *Một trang web thường xuyên ngừng hoạt động, thay đổi cấu trúc hoặc mất dữ liệu sẽ khó tạo dựng niềm tin. Ngược lại, một hệ thống duy trì được khả năng truy cập, định danh ổn định và cấu trúc nhất quán trong nhiều năm sẽ được các hệ thống khác dễ dàng tham chiếu và tin cậy hơn.*
• ***Yếu tố thứ ba là sự nhất quán.*** *Nếu cùng một địa danh nhưng mỗi lần công bố lại có một tên gọi, một địa chỉ hay một danh mục khác nhau, AI sẽ coi đây là tín hiệu bất ổn. Ngược lại, khi một thực thể được mô tả thống nhất qua thời gian và nhất quán với các nguồn khác, mức độ tin cậy sẽ dần được tích lũy.*

Điều này giải thích vì sao nhiều cơ sở dữ liệu của các tổ chức lớn có giá trị không chỉ vì nội dung của chúng, mà còn vì lịch sử vận hành. Hàng chục năm cập nhật liên tục, hàng triệu lần được tham chiếu và vô số lần được kiểm chứng đã tạo nên một loại vốn vô hình: vốn thẩm quyền. Đối với địa phương, đây là một góc nhìn rất khác về chuyển đổi số. Một cổng dữ liệu không chỉ là nơi công bố thông tin mà còn là nơi tích lũy niềm tin của máy. Mỗi lần cập nhật đúng hạn, mỗi lần sửa sai minh bạch, mỗi lần duy trì định danh ổn định đều góp thêm một phần vào tài sản ấy. Thẩm quyền số không được xây dựng bằng những chiến dịch truyền thông ngắn hạn, mà bằng hàng nghìn hành động nhỏ được lặp lại một cách bền bỉ.

Có thể xem đây là một dạng vốn danh tiếng của dữ liệu. Giống như thương hiệu của một doanh nghiệp hay uy tín của một quốc gia, thẩm quyền của một nguồn thông tin cần nhiều năm để hình thành nhưng có thể bị suy giảm rất nhanh nếu đánh mất sự nhất quán và độ tin cậy.

**MỆNH ĐỀ THẨM QUYỀN TÍCH LŨY**

*Thẩm quyền số được tích lũy qua thời gian nhờ độ ổn định, tính nhất quán, khả năng kiểm chứng và lịch sử được các nguồn khác tham chiếu. Vì mọi điều kiện ấy đều cần thời gian, thẩm quyền hình thành chậm, trong khi một đợt gián đoạn hoặc mâu thuẫn có thể làm suy giảm đáng kể phần tích lũy của nhiều năm.*

### 4.2.5. Mâu thuẫn giữa các nguồn được giải quyết như thế nào

Nếu mọi nguồn đều thống nhất với nhau, bài toán của AI sẽ rất đơn giản. Nhưng thế giới thực không vận hành như vậy. Một trang web ghi nhà hàng mở cửa đến 22 giờ. Google Maps hiển thị 23 giờ. Trang Facebook thông báo đóng cửa để sửa chữa. Báo chí gọi một di tích bằng tên cũ, trong khi cơ quan quản lý đã đổi sang tên mới. Wikipedia ghi dân số theo số liệu năm 2019, còn cơ quan thống kê vừa công bố kết quả điều tra năm 2025.[1]

1. Giả định của tác giả. Bốn nguồn mâu thuẫn trong đoạn này là ví dụ dựng để minh hoạ cơ chế đối chiếu, không phải một trường hợp có thật đã được kiểm chứng.

Máy phải làm gì khi cùng một câu hỏi lại có nhiều câu trả lời khác nhau? Đây là bài toán giải quyết xung đột giữa các nguồn (conflict resolution). Bước đầu tiên thường là đếm phiếu, tức xem có bao nhiêu nguồn độc lập cùng xác nhận một thông tin. Nếu sáu nguồn đều ghi cùng một địa chỉ và chỉ một nguồn ghi khác, khả năng cao nguồn duy nhất đó chứa lỗi. Tuy nhiên, bỏ phiếu chỉ là một tín hiệu, không phải là lời giải cuối cùng. Tiếp theo là thẩm quyền của từng nguồn. Nếu thông tin từ cơ quan quản lý mâu thuẫn với một bài đăng cá nhân, AI thường ưu tiên nguồn có thẩm quyền cao hơn. Nhưng ngay cả điều này cũng chưa đủ.

Một yếu tố quan trọng khác là độ mới của thông tin. Một dữ liệu rất đáng tin nhưng đã ba năm không cập nhật có thể kém giá trị hơn một nguồn chính thức vừa được cập nhật sáng nay. Trong thế giới số, tính đúng không chỉ phụ thuộc vào nguồn mà còn phụ thuộc vào thời điểm.

Cuối cùng là suy luận. Một hệ thống có thể kết hợp nhiều cơ chế cùng lúc, từ đếm phiếu và trọng số nguồn đến kiểm tra độ mới và xem xét tính nhất quán giữa các dữ kiện. Nếu một khách sạn được thông báo đã đóng cửa nhưng vẫn liên tục nhận đặt phòng, có đánh giá mới của khách và xuất hiện trong hệ thống đặt chỗ thời gian thực, AI có thể kết luận rằng thông tin "đã đóng cửa" nhiều khả năng đã lỗi thời. Ngược lại, nếu nhiều nguồn chính thức cùng xác nhận việc ngừng hoạt động, AI sẽ điều chỉnh kết luận của mình.

Điều quan trọng là hệ thống không tìm kiếm mọi sự thật. Trong mỗi lần suy luận, nó buộc phải chọn một phiên bản đủ đáng tin để sử dụng tại thời điểm ra quyết định. Đó là kết quả của việc cân nhắc đồng thời nhiều yếu tố: số lượng nguồn xác nhận,

mức độ thẩm quyền, độ mới của dữ liệu và sự nhất quán với các dữ kiện khác. Phiên bản được chọn không nhất thiết là phiên bản đúng, mà là phiên bản có bằng chứng mạnh nhất trong số những gì hệ thống truy hồi được.

Đối với địa phương, điều này đặt ra một yêu cầu rất thực tế. Nếu thông tin về cùng một thực thể khác nhau trên trang web chính thức, Google Maps, OpenStreetMap, báo chí và các nền tảng cộng đồng, AI không thể biết ngay nguồn nào đúng. Địa phương càng để tồn tại nhiều phiên bản khác nhau của cùng một sự thật, chi phí suy luận của máy càng lớn và khả năng được tin tưởng càng giảm. Vì vậy, quản trị dữ liệu không chỉ là tạo ra thông tin mới mà còn là giảm xung đột giữa các nguồn. Một địa phương được mô tả nhất quán trên toàn bộ hệ sinh thái số sẽ giúp AI không phải giải quyết quá nhiều mâu thuẫn, từ đó tăng xác suất được truy hồi, được tin tưởng và được lựa chọn.

Có thể tóm tắt nguyên lý này bằng một mệnh đề: Máy không tìm kiếm sự thật tuyệt đối. Máy lựa chọn phiên bản có bằng chứng mạnh nhất tại thời điểm ra quyết định. Đó cũng là lý do trong nền kinh tế tác nhân, quản trị tri thức không chỉ là quản trị dữ liệu, mà còn là quản trị sự đồng thuận của dữ liệu. Khi các nguồn cùng kể một câu chuyện nhất quán, AI sẽ có đủ cơ sở để biến câu chuyện ấy thành sự thật mà nó sử dụng để hành động.

Trong quá trình lựa chọn, hệ thống không dừng lại để tìm ra sự thật tuyệt đối mà chọn một phiên bản đủ đáng tin để hành động tại thời điểm ra quyết định. Vì vậy, một thông tin dù đúng nhưng không có bằng chứng, không được các nguồn độc lập xác nhận hoặc không đáp ứng các tiêu chí đánh giá của hệ thống, vẫn có thể bị bỏ qua. Trên thực tế, nó gần như vận hành như thể không tồn tại.

Để tăng khả năng được lựa chọn, địa phương cần tác động vào bốn yếu tố. Thứ nhất là bằng chứng, với câu hỏi kiểm tra: *Có bao nhiêu nguồn độc lập cùng xác nhận điều này?* Thay vì lặp lại cùng một thông điệp trên một kênh, mục tiêu là mở rộng số lượng nguồn độc lập cùng xác thực một thông tin. Thứ hai là thẩm quyền, với câu hỏi: *Nguồn nào có vị thế cao nhất đối với vấn đề này?* Nguồn chính thức và có thẩm quyền cần luôn là nơi cập nhật thông tin sớm và đầy đủ nhất. Thứ ba là tính mới, được kiểm tra bằng câu hỏi: *Thông tin này được cập nhật lần cuối khi nào?* Điều đó đòi hỏi mọi bản ghi đều có dấu thời gian rõ ràng và được cập nhật theo chu kỳ ổn định. Cuối cùng là tính nhất quán trong suy luận, với câu hỏi: *Các dữ kiện có logic và khớp với nhau không?* Địa phương cần loại bỏ các mâu thuẫn giữa các cơ sở dữ liệu và các hệ thống của chính mình để hình thành một bức tranh thống nhất mà máy có thể tin cậy.

- *Nếu bạn là người phụ trách cổng dữ liệu của địa phương:*

- *Coi mọi bản ghi thiếu nhãn thời gian là bản ghi hỏng, vì nó sẽ thua trong mọi lần máy so độ mới.*
- *Đo mức độ mâu thuẫn giữa các hệ thống của chính địa phương trước khi đi đo mức độ sai lệch của bên ngoài. Phần lớn xung đột mà máy nhìn thấy có nguồn gốc nội bộ.*
- *Duy trì cổng dữ liệu liên tục, kể cả trong giai đoạn không có nội dung mới. Gián đoạn phục vụ làm hỏng chỉ số ổn định nhiều hơn là thiếu bài viết.*
- *Không đổi cấu trúc địa chỉ liên kết khi nâng cấp hệ thống, hoặc nếu buộc phải đổi thì giữ chuyển hướng vĩnh viễn.*

Trở lại làng nghề. Khi một hệ thống cố xác định làng còn hoạt động hay không, nó gặp ba nguồn mâu thuẫn: một quyết định công nhận làng nghề truyền thống từ nhiều năm trước, một bài báo gần đây nói phần lớn lò đã ngừng đỏ lửa. Nguồn thứ ba là một trang du lịch quảng bá tour tham quan làng. Không nguồn nào có nhãn thời gian rõ ràng ngoài bài báo. Theo cơ chế, trong điều kiện thiếu những nguồn mới có thẩm quyền để đối trọng, nguồn tiêu cực nhưng mới nhất có nhiều khả năng được ưu tiên.

## 4.3. KHI MÁY TỰ VIẾT CÂU MÔ TẢ

Trong kỷ nguyên của công cụ tìm kiếm truyền thống, phần lớn nội dung hiển thị trước người dùng là những gì con người trực tiếp biên soạn với tiêu đề, đoạn giới thiệu, mô tả điểm đến hay bài viết trên trang web. AI tạo sinh đã thay đổi điều đó. Ngày càng nhiều câu trả lời không còn là sự sao chép nguyên văn từ một nguồn nào, mà là những đoạn văn mới được máy tự tổng hợp từ nhiều bằng chứng khác nhau. Nói cách khác, AI không chỉ đọc dữ liệu mà còn viết lại dữ liệu theo cách của chính nó.

Sự thay đổi này làm xuất hiện một lớp cạnh tranh hoàn toàn mới. Địa phương không còn chỉ cạnh tranh để có một bài viết hay, một trang web đẹp hay một vị trí cao trên công cụ tìm kiếm. Quan trọng hơn, địa phương phải cạnh tranh để khi AI tổng hợp thông tin từ hàng chục nguồn khác nhau, câu mô tả mà AI tự sinh ra vẫn phản ánh đúng bản sắc, đúng giá trị và đúng điều mà địa phương muốn truyền tải. Một khi AI đã viết ra một "phiên bản" của địa phương, chính phiên bản ấy có thể trở thành điều hàng triệu người đọc đầu tiên.

Điều đáng chú ý là AI không tự sáng tạo từ khoảng trống. Với những hệ thống kết hợp truy hồi, câu mô tả được hình thành từ các mảnh dữ liệu đã truy hồi được, những thực thể đã phân giải, những nguồn đã được đánh giá về mức độ tin cậy và tri thức tích lũy sẵn trong mô hình. Vì vậy, chất lượng của câu trả lời cuối cùng phụ thuộc trực tiếp vào chất lượng của hạ tầng tri thức mà địa phương đã xây dựng trước đó. Nếu dữ liệu rời rạc, mâu thuẫn hoặc thiếu ngữ cảnh, AI sẽ tạo ra một mô tả nghèo nàn hoặc thậm chí sai lệch. Ngược lại, khi tri thức được chuẩn hóa và nhất quán, AI có nhiều khả năng tái tạo đúng bản sắc của địa phương.

Mục này sẽ phân tích cách các mô hình ngôn ngữ tạo sinh xây dựng một câu mô tả từ tập hợp bằng chứng, vì sao hai AI khác nhau có thể viết hai đoạn giới thiệu khác nhau về cùng một địa phương và quan trọng hơn, địa phương cần làm gì để câu chuyện mà AI kể không chỉ chính xác mà còn phản ánh được những giá trị cốt lõi mà mình muốn gìn giữ và lan tỏa. Đây là bước chuyển từ quản trị dữ liệu sang quản trị cách AI kể câu chuyện về địa phương.

Cần nói rõ phạm vi trước khi đi tiếp. Trong mục này, cụm máy tự viết mô tả chủ yếu chỉ những hệ thống tạo sinh có khả năng kết hợp tri thức đã học trong tham số với dữ liệu được truy hồi từ nguồn bên ngoài. Một mô hình chỉ dựa vào tri thức đã học, một hệ thống chỉ trả về danh sách tài liệu, một tác nhân biết gọi công cụ và cơ sở dữ liệu là ba kiến trúc khác nhau. Không phải câu trả lời nào cũng có đầy đủ dấu vết nguồn. Chính sự thiếu vắng ấy là một phần của rủi ro được bàn dưới đây.

### 4.3.1. Từ truy hồi sang tổng hợp

Trong giai đoạn đầu của internet, hệ thống tìm kiếm chủ yếu thực hiện một nhiệm vụ tương đối đơn giản là truy hồi tài liệu. Khi người dùng nhập một truy vấn, máy tìm những trang phù hợp nhất rồi hiển thị danh sách liên kết. Người đọc phải tự mở từng trang, đối chiếu thông tin và tự hình thành kết luận. Giá trị của công cụ tìm kiếm nằm ở khả năng tìm đúng tài liệu.

Các mô hình ngôn ngữ tạo sinh đã thay đổi logic đó. Sau khi truy hồi các nguồn liên quan, AI không dừng lại ở việc đưa ra danh sách tài liệu mà tiếp tục thực hiện một bước mới là "tổng hợp". Thông tin từ nhiều nguồn được kết hợp, đối chiếu và diễn đạt lại thành một đoạn văn hoàn chỉnh bằng ngôn ngữ tự nhiên.

Thứ người dùng nhận được không còn là tập hợp các liên kết mà là một câu trả lời mới.

Vì phải đi qua nhiều tầng xử lý như vậy, câu trả lời cuối cùng thường không còn giữ được dấu vết của từng nguồn ban đầu. Quá trình ấy đánh dấu sự chuyển đổi từ đọc tài liệu sang tái cấu trúc tri thức. AI không còn đóng vai trò người chuyển phát thông tin mà trở thành người biên tập. Nó lựa chọn dữ kiện nào đáng giữ lại, bỏ qua dữ kiện nào, quyết định thứ tự trình bày và kết nối các ý thành một câu chuyện có vẻ mạch lạc. Mỗi câu trả lời vì thế là một sản phẩm mới được tạo ra từ kho bằng chứng đã truy hồi.

Hệ quả với địa phương nằm ở một chỗ khác. Không còn đủ để có một trang web thật tốt hay một bài báo nổi bật. Điều quan trọng hơn là phải tạo ra một hệ sinh thái tri thức nhất quán, bởi AI sẽ đọc tất cả rồi tự viết lại bằng ngôn ngữ của nó. Nếu các nguồn đều phản ánh cùng một bản sắc, AI có nhiều khả năng tái tạo đúng hình ảnh mà địa phương mong muốn. Ngược lại, nếu thông tin phân mảnh và mâu thuẫn, chính AI sẽ tổng hợp sự mâu thuẫn ấy thành câu trả lời dành cho hàng triệu người dùng. Nói cách khác, trong nền kinh tế tác nhân, địa phương không còn cạnh tranh để được trích dẫn, mà cạnh tranh để được AI tổng hợp thành phiên bản tốt nhất của chính mình.

Các hệ thống ngày càng dẫn nguồn nên vấn đề mất dấu vết sẽ tự biến mất. Nhiều công cụ hiện nay đã hiển thị đường dẫn tới tài liệu gốc bên cạnh câu trả lời. Khi việc dẫn nguồn trở thành chuẩn mực, người đọc sẽ truy ngược được và địa phương lấy lại quyền kiểm soát câu chuyện của mình. Tuy nhiên, dẫn nguồn giải quyết vấn đề minh bạch chứ không giải quyết vấn đề tổng hợp. Một câu trả lời được ghép từ nhiều mảnh dữ liệu vẫn là một mệnh đề mới mà không nguồn nào trong danh sách trích dẫn

từng phát biểu nguyên văn. Danh sách nguồn cho biết máy đã đọc gì, không cho biết máy đã suy ra điều đó bằng cách nào và đã bỏ qua điều gì. Ngoài ra, phần lớn người dùng không mở các đường dẫn ấy, nên trong thực tế câu trả lời tổng hợp vẫn là phiên bản duy nhất mà đa số tiếp nhận.

### 4.3.2. Mô tả là sản phẩm của nhiều nguồn

Một đoạn mô tả do AI tạo ra hiếm khi bắt nguồn từ một tài liệu duy nhất. Phía sau vài câu văn ngắn gọn có thể là hàng chục hoặc hàng trăm mảnh thông tin được thu thập từ nhiều nguồn khác nhau. Mỗi nguồn chỉ đóng góp một phần của bức tranh tổng thể, nơi cung cấp tên gọi chính thức, nơi bổ sung lịch sử, nơi mô tả địa lý, nơi phản ánh đánh giá của cộng đồng, nơi cập nhật tình trạng mới nhất. AI có nhiệm vụ kết hợp tất cả thành một câu trả lời thống nhất. Quá trình này thường được gọi là hợp nhất tri thức (knowledge fusion). Thay vì coi mỗi tài liệu là một câu chuyện độc lập, AI xem chúng như những bằng chứng về cùng một thực thể. Các thông tin trùng lặp được củng cố lẫn nhau, các chi tiết bổ sung được ghép lại, còn những điểm mâu thuẫn phải được đánh giá để quyết định nên tin nguồn nào hơn. Mục tiêu là xây dựng một mô hình tri thức thống nhất về đối tượng đang được hỏi, không phải sao chép từng tài liệu.

Sau khi hợp nhất tri thức, AI tiếp tục thực hiện bước tóm lược (summarization). Một điểm đến có thể có hàng nghìn trang tài liệu, nhưng người dùng thường chỉ nhận được vài trăm từ. Máy phải quyết định điều gì đủ quan trọng để giữ lại và điều gì sẽ bị loại bỏ. Mỗi lần tóm lược là một lần giảm độ phong phú của thế giới thực xuống thành một phiên bản ngắn gọn hơn, dễ đọc hơn và dễ sử dụng hơn.

Cuối cùng là bước suy luận (reasoning). AI không chỉ ghép các câu lại với nhau mà còn tạo ra những kết nối logic giữa các dữ kiện. Nếu nhiều nguồn cùng mô tả một địa phương có mật độ di sản cao, hoạt động văn hóa phong phú và lượng du khách quốc tế lớn, AI có thể kết luận rằng đây là một trung tâm du lịch văn hóa, ngay cả khi không có tài liệu nào viết nguyên văn câu đó. Nhận định cuối cùng là kết quả của quá trình suy luận dựa trên nhiều bằng chứng.

Điều này tạo ra một thay đổi căn bản trong quản trị thông tin. Trước đây, địa phương chỉ cần xuất hiện trên một vài kênh quan trọng. Ngày nay, mọi nguồn đều có thể trở thành nguyên liệu cho cùng một câu mô tả. Một sai sót nhỏ trên một cơ sở dữ liệu, một thông tin cũ trên một trang web ít người biết hay một định nghĩa không thống nhất trong một kho dữ liệu mở đều có thể được AI hấp thụ và đưa vào quá trình tổng hợp. Trong thời đại AI tạo sinh, điều cần quản trị không còn là từng tài liệu riêng lẻ mà là toàn bộ hệ sinh thái tri thức mà máy sẽ sử dụng để viết nên câu chuyện về địa phương.

### 4.3.3. Vì sao AI khác Wikipedia

Thoạt nhìn, một câu trả lời của AI và một bài viết trên Wikipedia có vẻ giống nhau. Cả hai đều là những đoạn văn ngắn gọn, trình bày kiến thức theo lối bách khoa, giúp người đọc nhanh chóng nắm được thông tin về một địa điểm, một tổ chức hay một sự kiện. Tuy nhiên, phía sau hai dạng văn bản ấy là hai cơ chế hoàn toàn khác nhau.

Wikipedia là một kho lưu trữ tri thức. Mỗi bài viết là một tài liệu cụ thể được cộng đồng biên soạn, chỉnh sửa và lưu lại. Khi

người đọc mở trang về Hội An hay Vịnh Hạ Long, họ đang đọc đúng bài viết đã được các biên tập viên xây dựng trước đó. Nội dung có thể được cập nhật theo thời gian, nhưng giữa hai lần cập nhật, mọi người đều nhìn thấy cùng một phiên bản.

AI tạo sinh không hoạt động như vậy. Nó không có một "bài viết về Hội An" được lưu sẵn để lấy ra khi cần. Thay vào đó, mỗi lần nhận câu hỏi, AI lại truy hồi thông tin, lựa chọn bằng chứng, sắp xếp ý tưởng và tạo ra một đoạn văn mới ngay tại thời điểm trả lời. Hai người dùng đặt cùng một câu hỏi ở hai thời điểm khác nhau, hoặc trên hai mô hình AI khác nhau, hoàn toàn có thể nhận được hai cách diễn đạt khác nhau dù cùng nói về một thực thể.

Điều này khiến AI gần giống một người biên tập thời gian thực hơn là một thư viện. Mỗi câu trả lời là kết quả của quá trình tổng hợp tức thời giữa dữ liệu đã truy hồi, tri thức tích lũy trong mô hình, các bước suy luận và ngữ cảnh của cuộc hội thoại. Nội dung không tồn tại trước khi câu hỏi được đặt ra. Nó được sinh ra trong chính quá trình trả lời.

Chính vì vậy, không tồn tại một "phiên bản chính thức" duy nhất của đoạn mô tả mà AI tạo ra. Thứ tương đối ổn định không phải từng câu chữ mà là hệ thống bằng chứng phía sau. Nếu hệ sinh thái tri thức về một địa phương phong phú, nhất quán và đáng tin cậy, nhiều mô hình AI khác nhau sẽ có xu hướng tạo ra những mô tả tương đồng. Ngược lại, nếu dữ liệu phân tán hoặc mâu thuẫn, mỗi mô hình có thể kể một câu chuyện khác nhau về cùng một địa phương.

Đây là sự khác biệt căn bản giữa tri thức được lưu trữ và tri thức được sinh ra. Wikipedia quản trị từng bài viết. AI quản trị quá trình tạo ra bài viết. Vì thế, trong nền kinh tế tác nhân, mục

tiêu của địa phương không còn là viết một bài thật hay, mà là xây dựng một hệ tri thức đủ mạnh để bất kỳ AI nào cũng có thể tự tạo ra một câu chuyện chính xác và nhất quán.

### 4.3.4. Khi sai lệch nhỏ trở thành mô tả chính thức

Khả năng tự tổng hợp của AI mang lại nhiều lợi ích, nhưng cũng tạo ra một rủi ro mới, một sai lệch rất nhỏ có thể trở thành nền tảng cho hàng loạt câu trả lời sau này. Không phải mọi thông tin mà AI tạo ra đều chính xác. Trong nghiên cứu về mô hình ngôn ngữ, hiện tượng AI tạo ra những thông tin nghe có vẻ hợp lý nhưng không có cơ sở được gọi là ảo giác (hallucination). AI có thể nhầm ngày thành lập một di tích, gán sai một lễ hội cho một địa phương, hoặc kết nối hai sự kiện vốn không liên quan. Những lỗi này thường không đến từ ác ý mà từ quá trình mô hình cố gắng tạo ra một câu trả lời hoàn chỉnh khi bằng chứng chưa đủ hoặc mâu thuẫn.

Nguy hiểm hơn là hiện tượng lan truyền sai lệch (propagation). Một thông tin không chính xác, nếu được đăng lại trên nhiều trang web, xuất hiện trong nhiều cơ sở dữ liệu hoặc được nhiều AI khác nhau lặp lại, sẽ dần có vẻ như là sự thật. Các hệ thống AI sau đó lại truy hồi chính những nguồn đã bị ô nhiễm ấy, khiến sai lệch tiếp tục được khuếch đại. Một lỗi nhỏ ban đầu có thể lan rộng thành một "sự thật số" được nhiều người tin.

Khi sai lệch được lặp đi lặp lại, nó còn tạo ra hiệu ứng xác nhận (confirmation). Con người có xu hướng tin một thông tin hơn khi nhìn thấy nó xuất hiện ở nhiều nơi. AI cũng gặp một vấn đề tương tự, nếu nhiều nguồn cùng nói giống nhau, hệ thống thường coi đó là bằng chứng mạnh hơn. Khi nguồn gốc của các

bản sao không còn được phân biệt rõ, việc lặp lại nhiều lần có thể bị nhầm là sự xác nhận độc lập, dù tất cả đều bắt nguồn từ cùng một lỗi ban đầu.

Đối với địa phương, đây là một thách thức quản trị mới. Một mô tả sai về lịch sử, một tọa độ nhầm trên bản đồ, một tên gọi không thống nhất hay một thông tin lỗi thời tưởng như rất nhỏ có thể được AI hấp thụ, tổng hợp rồi phổ biến tới hàng triệu người dùng. Khi nhiều hệ thống tiếp tục sử dụng chính những câu trả lời đó làm nguồn tham khảo, sai lệch càng trở nên khó sửa hơn. Vì vậy, trong thời đại AI tạo sinh, việc sửa lỗi không còn là chỉnh lại một bài viết hay một trang web. Điều cần làm là làm sạch toàn bộ hệ sinh thái tri thức, đồng bộ thông tin giữa các nguồn, thiết lập các nguồn gốc có thẩm quyền và phát hiện sai lệch trước khi chúng được khuếch đại. Trong một môi trường mà AI liên tục đọc, tổng hợp và tái tạo tri thức, một sai sót nhỏ hôm nay có thể trở thành mô tả mặc định của ngày mai.

### 4.3.5. Địa phương mất quyền kể câu chuyện của mình như thế nào

Trong nhiều thế kỷ, câu chuyện về một địa phương chủ yếu được kể bởi chính địa phương đó qua sử sách, qua người dân, qua hướng dẫn viên, qua báo chí địa phương hay các cơ quan quản lý. Dù cách kể có thể khác nhau, chủ thể kể chuyện vẫn tương đối rõ ràng. Kỷ nguyên AI tạo sinh đang làm thay đổi trật tự ấy. Người kể chuyện đầu tiên về một địa phương ngày càng không còn là con người, mà là một cỗ máy tổng hợp tri thức.

Sự chuyển đổi này không diễn ra đột ngột mà âm thầm qua nhiều bước. Ban đầu, một thông tin cũ không được cập nhật trên

trang web chính thức. Một tên gọi không thống nhất xuất hiện trên bản đồ số. Một bài báo chứa sai sót vẫn tồn tại nhiều năm mà không ai sửa. Mỗi sai lệch riêng lẻ có vẻ không đáng kể. Tuy nhiên, AI không nhìn từng nguồn riêng biệt mà đọc toàn bộ hệ sinh thái thông tin. Những dữ liệu ấy trở thành nguyên liệu để máy tạo nên câu mô tả mới.

Khi AI bắt đầu kể câu chuyện ấy cho hàng triệu người dùng, quá trình lại bước sang một giai đoạn khác. Các blogger, doanh nghiệp, nền tảng du lịch và thậm chí các cơ quan truyền thông có thể sử dụng chính câu trả lời của AI làm tài liệu tham khảo. Những nội dung mới lại được xuất bản lên internet, trở thành dữ liệu để các AI khác tiếp tục truy hồi trong tương lai. Một vòng lặp tự củng cố bắt đầu hình thành. Chuỗi chuyển hóa ấy có thể được mô tả rất đơn giản nhưng không ai sửa. Chuỗi chuyển giao quyền kể chuyện, máy tự kể, người khác trích dẫn, sai trở thành mặc định. Điều đáng lo ngại là ở giai đoạn cuối, rất ít người còn truy ngược được nguồn gốc của thông tin. Người đọc chỉ thấy cùng một nội dung xuất hiện trên nhiều trang web, nhiều chatbot và nhiều nền tảng khác nhau. Sự lặp lại tạo ra cảm giác về tính xác thực, mặc dù tất cả có thể bắt nguồn từ cùng một sai lệch ban đầu. Quyền lực không còn nằm ở người tạo ra lỗi, mà ở cơ chế khuếch đại của hệ sinh thái số.

Đối với địa phương, đây là một sự dịch chuyển về chủ quyền diễn ngôn. Nếu không chủ động quản trị dữ liệu, cập nhật nguồn chính thức và duy trì sự nhất quán giữa các hệ thống tri thức, địa phương sẽ dần đánh mất quyền định nghĩa chính mình. Câu chuyện về lịch sử, văn hóa, bản sắc hay lợi thế cạnh tranh sẽ không còn do địa phương trực tiếp kể, mà do các mô hình AI liên tục tái tạo từ những dữ liệu đang tồn tại trên internet. Vì vậy, trong nền kinh tế tác nhân, bảo vệ bản sắc không còn chỉ là bảo

tồn di sản hay quảng bá hình ảnh. Đó còn là năng lực quản trị câu chuyện mà AI sẽ kể. Một địa phương thực sự có chủ quyền số không phải là địa phương có nhiều dữ liệu nhất, mà là địa phương tạo ra đủ bằng chứng có thẩm quyền để các hệ thống khác nhau, khi tự viết về mình, có xu hướng hội tụ về cùng một câu chuyện cốt lõi: chính xác, nhất quán và phản ánh đúng bản sắc mà địa phương muốn truyền lại cho thế giới.

Ở khâu tổng hợp, hệ thống không trích lại nguyên văn bài báo. Nó viết một câu mới: "làng nghề từng nổi tiếng, nay chỉ còn vài hộ duy trì sản xuất". Câu ấy không có trong bất kỳ nguồn nào, nhưng nghe hợp lý và được rút ra từ tập bằng chứng sẵn có. Từ thời điểm này, phiên bản do máy viết bắt đầu tồn tại độc lập với mọi tài liệu gốc.

PHẦN B
# TRI THỨC ẤY BIẾN ĐỔI NHƯ THẾ NÀO

## 4.4. VÒNG LẶP THAY CHO ĐƯỜNG THẲNG

Trong cách hình dung cũ, thông tin vận động theo một đường thẳng. Một địa phương tạo ra dữ liệu. Trang web đăng tải dữ liệu ấy. Công cụ tìm kiếm thu thập. Người dùng tìm thấy, đọc và đưa ra quyết định. Khi một thông tin sai được phát hiện, người ta có thể quay về nơi xuất bản ban đầu, sửa lại nội dung và chờ phiên bản mới được cập nhật. Chuỗi này tương đối dễ quan sát bởi có người viết, có nơi đăng, có người đọc và có một điểm được xem là nguồn gốc.

Trong hệ thống trí tuệ nhân tạo, thông tin không còn đi theo con đường ấy. Nó vận động trong một vòng. Một dữ kiện được xuất bản sẽ được máy truy hồi, đưa vào câu trả lời, rồi được người dùng sao chép, diễn đạt lại và xuất bản tiếp dưới hình thức mới. Đầu ra của vòng trước trở thành đầu vào của vòng sau. Mỗi lần quay lại, dữ kiện ấy mang thêm dấu vết của lần xử lý trước đó. Bởi vậy, quyền kể câu chuyện của địa phương không chỉ được quyết định tại thời điểm câu chuyện được viết ra. Nó được quyết định qua toàn bộ vòng đời của thông tin khi được thu thập, phân giải, tổng hợp, trích dẫn, xếp hạng, sử dụng và tái xuất bản. Địa phương nào chỉ chú ý đến việc phát hành thông tin mà không quản trị vòng lặp sẽ dần mất quyền kiểm soát cách mình được mô tả. Trong kiến trúc mới, sự thật không chỉ cần được công bố một lần. Nó phải đủ rõ, đủ nhất quán và đủ bền để tồn tại qua nhiều vòng tái tạo của máy.

Người ta thường hình dung quá trình học của máy như một đường truyền một chiều, dữ liệu được đưa vào hệ thống, máy xử lý dữ liệu ấy, tạo ra mô tả rồi chuyển mô tả đến người dùng. Trong hình dung này, dữ liệu là điểm khởi đầu, câu trả lời là điểm kết thúc. Máy chỉ đứng giữa, làm nhiệm vụ biến thông tin có sẵn thành nội dung dễ sử dụng hơn. Nhưng câu trả lời của máy không kết thúc tại màn hình của người dùng. Nó tiếp tục đi vào thế giới. Người dùng có thể sao chép câu trả lời vào một bài viết, đưa mô tả ấy lên trang web, bổ sung vào hồ sơ doanh nghiệp, sử dụng trong tài liệu quảng bá hoặc kể lại cho người khác. Báo chí có thể tham khảo nội dung do máy tổng hợp. Người làm truyền thông có thể dùng AI để viết phần giới thiệu cho một địa danh. Doanh nghiệp có thể lấy chính mô tả ấy làm dữ liệu chính thức trên các nền tảng số.

Khi những nội dung mới được xuất bản, chúng lại trở thành dữ liệu cho các hệ thống tìm kiếm, bản đồ, cơ sở tri thức và mô hình trí tuệ nhân tạo. Một câu trả lời từng là đầu ra của máy nay quay trở lại hệ thống dưới hình thức một nguồn mới. Máy tiếp tục đọc nó, ghép với những nguồn khác và tạo ra một mô tả tiếp theo. Vì vậy, chuỗi vận động thực tế là một vòng khép kín chứ không phải một đường thẳng, trong đó đầu ra của vòng trước trở thành đầu vào của vòng sau.

**MỆNH ĐỀ VÒNG KÍNH**

*Khi đầu ra của một hệ thống được con người tái xuất bản, nó quay lại làm đầu vào cho chính hệ thống ấy hoặc cho hệ thống khác. Hệ quả là công bố không còn là hành động cuối cùng của một quy trình mà là bước khởi đầu của một chu kỳ không có điểm dừng.*

Điểm đáng chú ý là dữ liệu mới không nhất thiết chứa thêm tri thức mới. Nhiều khi nó chỉ là sự diễn đạt lại nội dung đã tồn tại. Một câu mô tả do máy tạo ra có thể được mười trang web sao chép với những biến đổi nhỏ. Khi hệ thống quay lại đọc mười trang web ấy, nó có thể hiểu rằng nhiều nguồn độc lập đang cùng xác nhận một thông tin. Nhưng trên thực tế, tất cả chỉ bắt nguồn từ một câu trả lời duy nhất. Số lượng nguồn tăng lên, trong khi số lượng bằng chứng gốc không tăng.

Vòng kín này làm xuất hiện một dạng đồng thuận nhân tạo. Một thông tin được lặp lại nhiều lần có thể trở nên đáng tin hơn trong mắt thuật toán, dù sự lặp lại không đến từ quá trình kiểm chứng độc lập. Máy tạo ra mô tả. Con người phát tán mô tả. Các nền tảng thu thập những bản sao. Máy nhìn thấy nhiều nguồn giống nhau và tạo ra một mô tả còn chắc chắn hơn. Mỗi vòng quay làm mờ thêm dấu vết của nguồn ban đầu.

Điều đó cũng có nghĩa rằng máy không chỉ học từ thế giới do con người quan sát và ghi chép. Máy ngày càng học từ một thế giới đã được chính máy tham gia mô tả. Nội dung tổng hợp bởi trí tuệ nhân tạo trở thành một phần của môi trường thông tin, rồi môi trường thông tin ấy lại được dùng để huấn luyện, truy hồi và xác nhận cho các hệ thống tiếp theo. Máy không đứng ngoài hệ sinh thái tri thức. Nó vừa là người đọc, vừa là người viết, vừa là tác nhân làm thay đổi dữ liệu mà nó sẽ đọc trong tương lai.

Nếu thông tin ban đầu chính xác, vòng lặp có thể giúp tri thức lan rộng, trở nên dễ tiếp cận và ngày càng hoàn chỉnh. Nhưng nếu thông tin ban đầu sai, vòng lặp cũng có thể khuếch đại sai lệch theo cùng một cách. Một di tích bị ghi sai niên đại. Một làng nghề bị gắn nhầm địa giới. Một doanh nghiệp bị ghép với địa chỉ cũ. Một món ăn địa phương bị quy về sai nguồn gốc. Sau nhiều

lần tái sử dụng, sai lệch ấy không còn xuất hiện như một lỗi đơn lẻ mà như một sự thật được nhiều nguồn xác nhận. Bởi vậy, câu hỏi quan trọng không còn chỉ là máy đã học từ dữ liệu nào. Cần hỏi thêm: dữ liệu ấy có bao nhiêu phần được tạo ra từ những câu trả lời trước đó của máy? Có bao nhiêu nguồn thực sự độc lập? Có bao nhiêu mô tả chỉ là bản sao của cùng một nội dung? Và liệu hệ thống đang tiếp cận thế giới thực hay đang tiếp cận hình ảnh của thế giới do chính nó góp phần dựng nên?

Với địa phương, hệ quả rất rõ ràng. Một thông tin sai không chỉ cần được sửa tại nơi nó xuất hiện đầu tiên. Nó phải được sửa ở nhiều điểm trong vòng lặp, nguồn chính thức, bản đồ, cơ sở dữ liệu, hồ sơ doanh nghiệp, bài báo, nền tảng cộng đồng và các hệ thống đang sử dụng lại thông tin ấy. Nếu chỉ sửa một trang web nhưng hàng chục nguồn phái sinh vẫn giữ nội dung cũ, mô tả sai có thể tiếp tục quay trở lại và lấn át thông tin đã được hiệu chỉnh.

Trong một hệ thống đường thẳng, công bố là hành động cuối cùng. Trong một hệ thống vòng kín, công bố chỉ là khởi đầu của một chu kỳ mới. Dữ liệu sẽ được đọc, diễn giải, sử dụng, tái xuất bản và đọc lại. Vì thế, quản trị tri thức địa phương không thể dừng ở việc tạo ra thông tin đúng. Nó còn phải theo dõi xem thông tin ấy biến đổi như thế nào sau mỗi vòng quay và liệu phiên bản đang được máy sử dụng có còn phản ánh đúng điều địa phương muốn nói về chính mình hay không.

### 4.4.1. Mỗi lần được đọc là một lần được củng cố

Đối với con người, việc đọc một thông tin không làm thay đổi bản thân thông tin ấy. Nhưng trong hệ sinh thái trí tuệ nhân tạo, mỗi lần một thông tin được đọc, sử dụng và xuất hiện trở

lại trong không gian số đều có thể làm tăng khả năng nó tiếp tục được hệ thống lựa chọn. Quá trình này tạo thành một cơ chế tự củng cố (reinforcement) thay vì chỉ đơn thuần là truyền tải tri thức.

Cần nói rõ một điều kiện trước. Việc được đọc chỉ trở thành củng cố khi nó để lại dấu vết, một liên kết, một lượt chọn, một đánh giá, một nội dung phái sinh hay một hành động tiếp theo. Một người đọc rồi rời đi mà không để lại gì gần như không tạo ra tín hiệu nào cho hệ thống. Với điều kiện ấy, bước đầu tiên của cơ chế là mức độ xuất hiện (exposure). Một địa danh, một doanh nghiệp hay một di tích càng xuất hiện trên nhiều trang web, bản đồ, bài báo, mạng xã hội và cơ sở dữ liệu thì càng có nhiều cơ hội được các hệ thống truy hồi. Khả năng được nhìn thấy trở thành điều kiện đầu tiên để được máy đọc.

Sau khi được đọc, thông tin tạo ra phản hồi (feedback). Người dùng nhấp vào kết quả, đặt câu hỏi tương tự, chia sẻ nội dung, trích dẫn mô tả hoặc sử dụng câu trả lời của AI trong tài liệu mới. Những tín hiệu này không phải lúc nào cũng được dùng trực tiếp để huấn luyện mô hình nền, nhưng chúng thường ảnh hưởng đến các hệ thống truy hồi, xếp hạng, gợi ý và tối ưu hóa dịch vụ. Hệ thống vì thế học được rằng thông tin ấy đang được sử dụng nhiều hơn.

Khi mức độ xuất hiện và phản hồi cùng tăng, quá trình tự củng cố (reinforcement) bắt đầu hình thành. Một thực thể thường xuyên được truy hồi sẽ có xu hướng tiếp tục được ưu tiên. Một mô tả được sử dụng nhiều sẽ có xác suất cao hơn xuất hiện trong các câu trả lời tiếp theo. Không phải vì hệ thống biết chắc nó đúng, mà vì có nhiều bằng chứng cho thấy nó đang được cộng đồng và các nền tảng sử dụng.

Sau nhiều vòng lặp, cơ chế này tạo nên một dạng bộ nhớ của máy (machine memory). Bộ nhớ ấy không tồn tại như một cuốn sách lưu giữ nguyên văn từng câu chữ. Nó được phân tán trong vô số thành phần: cơ sở tri thức, chỉ mục tìm kiếm, biểu diễn véc-tơ, đồ thị tri thức, trọng số xếp hạng và các mô hình thống kê về mức độ liên quan. Một thông tin được nhắc lại đủ nhiều sẽ ngày càng khó biến mất khỏi hệ sinh thái số, ngay cả khi nguồn gốc ban đầu của nó đã không còn. Đó là lý do lợi thế trong nền kinh tế tác nhân không chỉ đến từ việc có dữ liệu đúng, mà còn từ việc dữ liệu đúng liên tục được sử dụng, được xác nhận và được tái xuất hiện. Khả năng được đọc một lần là quan trọng. Nhưng khả năng tiếp tục được đọc sau hàng nghìn vòng lặp mới là yếu tố quyết định liệu một địa phương có được bộ nhớ thuật toán ghi nhận lâu dài hay không.

Hiệu ứng giàu càng giàu không phải điều mới, thị trường truyền thông xưa nay vẫn vậy. Báo chí luôn viết về nơi đã nổi tiếng, hãng lữ hành luôn bán nơi dễ bán. Cơ chế tích lũy lợi thế đã tồn tại từ rất lâu trước khi có trí tuệ nhân tạo, nên đây không phải một hiện tượng cần lý thuyết mới. Cơ chế không mới nhưng ba tham số của nó đã đổi. Thứ nhất là tốc độ, một vòng củng cố từng mất nhiều tháng nay có thể khép lại trong vài ngày. Thứ hai là nguồn nguyên liệu, hệ thống ngày càng học từ nội dung do chính hệ thống sinh ra, nên vòng lặp không còn cần thêm quan sát mới từ thế giới thực để tiếp tục quay. Thứ ba là điểm can thiệp, trong thị trường truyền thông cũ, một địa phương có thể thay đổi cục diện bằng cách thuyết phục một số ít người biên tập, còn trong hệ sinh thái hiện nay không có người biên tập nào để thuyết phục.

**MỆNH ĐỀ KHUẾCH ĐẠI BẤT ĐỐI XỨNG**

*Vì mỗi vòng lặp ưu tiên những thực thể đã có nhiều bằng chứng và nhiều tín hiệu phản hồi, chênh lệch ban đầu, dù nhỏ và dù phát sinh ngẫu nhiên, có xu hướng nới rộng qua mỗi vòng. Hệ quả là can thiệp sớm rẻ hơn nhiều so với can thiệp mạnh về sau.*

### 4.4.2. Bộ nhớ thuật toán khác bộ nhớ con người

Con người nhớ bằng trải nghiệm. Một chuyến đi đáng nhớ, một bữa ăn đặc biệt hay một cuộc gặp gỡ ý nghĩa có thể lưu lại trong ký ức hàng chục năm. Bộ nhớ ấy không nằm trong một "tệp" hay một tế bào thần kinh riêng lẻ, mà được mã hóa trong sự thay đổi của hàng triệu kết nối giữa các neuron. Cảm xúc, sự chú ý, bối cảnh và việc lặp lại trải nghiệm làm thay đổi độ mạnh của các kết nối này, khiến một ký ức trở nên bền vững hoặc dần phai nhạt. Vì vậy, hai người cùng tham dự một sự kiện có thể nhớ những điều hoàn toàn khác nhau, bởi mỗi người đã thay đổi mạng lưới thần kinh của mình theo một cách khác nhau.

Bộ nhớ của hệ thống máy cũng tồn tại dưới nhiều dạng, từ trọng số đã học trong mô hình, cơ sở dữ liệu và chỉ mục dùng để truy hồi, đến đồ thị tri thức, lịch sử tương tác và các tín hiệu xếp hạng tích lũy trong toàn bộ hệ sinh thái. Điểm chung của chúng không nằm ở một cơ chế lưu trữ thống nhất, mà ở chỗ chúng đều duy trì những cấu trúc có thể được tái sử dụng trong lần xử lý tiếp theo. Đặc biệt, ở các mô hình học sâu, phần lớn tri thức không được lưu dưới dạng các bản ghi riêng lẻ mà được phân bố

trong hàng tỷ trọng số của mạng nơ-ron nhân tạo. Ở điểm này, chúng có một nét tương đồng đáng chú ý với hệ thần kinh sinh học, thông tin được mã hóa trong trạng thái của cả mạng lưới hơn là trong một đơn vị đơn lẻ.

Sự tương đồng này chỉ dừng ở mức nguyên lý biểu diễn phân tán (distributed representation). Não người thay đổi các kết nối thần kinh thông qua trải nghiệm sống, cảm xúc, cơ thể và quá trình củng cố ký ức diễn ra liên tục suốt đời. Ngược lại, mô hình AI chủ yếu thay đổi trọng số trong quá trình huấn luyện và sau đó thường giữ cố định, phần lớn "bộ nhớ" khi vận hành đến từ các cấu trúc bên ngoài như cơ sở dữ liệu, chỉ mục truy hồi và lịch sử tương tác. Vì vậy, AI không ghi nhớ theo nghĩa trải nghiệm, mà ghi nhận những cấu trúc thống kê có thể tái sử dụng để dự đoán và hành động. Vì vậy, khi nói AI "nhớ" một địa phương, điều đó không có nghĩa máy đang lưu giữ một câu chuyện hoàn chỉnh về nơi ấy. Máy chỉ đang duy trì những phân bố xác suất về cách các thực thể, thuộc tính và quan hệ thường xuất hiện trong dữ liệu. Nếu "Hội An" thường đi cùng "phố cổ", "đèn lồng" và "di sản thế giới", xác suất của những liên kết này sẽ cao hơn nhiều so với các mô tả hiếm gặp.

Sự khác biệt này tạo ra một hệ quả quan trọng. Con người có thể thay đổi ký ức chỉ sau một trải nghiệm mạnh. Một chuyến đi thực tế có thể làm thay đổi hoàn toàn ấn tượng về một địa phương. Máy thì không. Để thay đổi "ký ức" của hệ thống cần một lượng lớn dữ liệu mới, nhiều lần xuất hiện, nhiều nguồn xác nhận và nhiều vòng cập nhật liên tiếp. Bộ nhớ xác suất có tính ì lớn hơn nhiều so với ký ức của con người.

Chính vì được xây dựng từ xác suất, bộ nhớ thuật toán cũng không phân biệt giữa điều được lặp lại nhiều và điều có ý nghĩa sâu sắc. Một lễ hội diễn ra hàng trăm năm nhưng rất ít dữ liệu số

có thể bị đánh giá thấp hơn một sự kiện mới được truyền thông mạnh trên nhiều nền tảng. Một truyền thống chỉ tồn tại trong ký ức cộng đồng gần như không để lại dấu vết trong bộ nhớ của máy nếu không được chuyển thành dữ liệu khả đọc.

Đối với địa phương, đây là một thay đổi căn bản trong tư duy quản trị. Muốn được lưu giữ trong bộ nhớ của con người cần tạo ra trải nghiệm đáng nhớ. Muốn được lưu giữ trong bộ nhớ của thuật toán cần tạo ra dữ liệu nhất quán, có cấu trúc, được nhiều nguồn xác nhận và liên tục được cập nhật. Hai loại bộ nhớ này không thay thế nhau. Chúng bổ sung cho nhau. Một địa phương chỉ có dữ liệu mà không có trải nghiệm sẽ dễ được máy nhớ nhưng khó được con người yêu mến. Ngược lại, một địa phương chỉ có trải nghiệm mà không có dữ liệu sẽ được con người nhớ nhưng rất dễ bị máy quên.

### 4.4.3. Vì sao vòng lặp luôn khuếch đại

Vòng lặp không chỉ làm thông tin quay trở lại. Nó còn làm cho chênh lệch ban đầu ngày càng lớn hơn. Một thực thể được biết đến nhiều sẽ có nhiều cơ hội được tìm kiếm. Được tìm kiếm nhiều, nó lại xuất hiện trong nhiều bài viết, bản đồ, đánh giá và câu trả lời của AI hơn. Mỗi lần xuất hiện mới tiếp tục làm tăng khả năng được truy hồi ở vòng sau. Lợi thế nhỏ ở điểm khởi đầu vì thế có thể trở thành lợi thế lớn sau nhiều chu kỳ.

Cơ chế này gần với Hiệu ứng Matthew, ai đã có thì sẽ có thêm, ai có ít thì ngày càng khó được nhìn thấy. Trong môi trường thuật toán, điều "đã có" không chỉ là tiền bạc hay danh tiếng, mà còn là dữ liệu, liên kết, lượt nhắc đến, đánh giá, lịch sử tương tác và sự hiện diện trên nhiều nền tảng. Một địa phương đã có vị

trí mạnh trong bộ nhớ số sẽ dễ tiếp tục tích lũy bằng chứng về chính vị trí mạnh ấy. Đây cũng là logic giàu càng giàu. Một điểm đến nổi tiếng được báo chí viết nhiều hơn, được du khách đánh giá nhiều hơn, được nền tảng ưu tiên hơn và được AI đề xuất thường xuyên hơn. Mỗi đề xuất tạo ra thêm lượt ghé thăm. Mỗi lượt ghé thăm tạo thêm ảnh, bài đăng, đánh giá và giao dịch. Những dữ liệu mới lại quay về củng cố thứ hạng ban đầu. Danh tiếng không chỉ phản ánh thành công. Nó còn trở thành nguyên liệu để sản xuất thêm thành công. Ngược lại, một nơi ít được biết đến không chỉ thiếu sự chú ý. Nó còn thiếu dữ liệu phản hồi để chứng minh rằng mình đáng được chú ý. Ít người tìm kiếm dẫn đến ít nội dung. Ít nội dung dẫn đến ít bằng chứng. Ít bằng chứng khiến hệ thống càng ít đề xuất. Sự vô hình vì thế cũng tự củng cố theo vòng lặp, giống như danh tiếng.

Trong quá trình này xuất hiện thiên lệch phổ biến. Hệ thống có xu hướng ưu tiên những gì đã phổ biến bởi sự phổ biến tạo ra nhiều tín hiệu dễ đo lường gắn với lượt xem, lượt nhấp, đánh giá, liên kết và mức độ được nhắc đến. Nhưng phổ biến không đồng nghĩa với tốt hơn, đúng hơn hay phù hợp hơn. Một nơi đông khách có thể tiếp tục được đề xuất chỉ vì đã đông khách. Một mô tả quen thuộc có thể tiếp tục được lặp lại chỉ vì nó phổ biến, dù đã giản lược hoặc làm sai lệch bản sắc địa phương. Thiên lệch phổ biến khiến thuật toán dễ chọn phương án an toàn. Khi phải quyết định giữa một thực thể nổi tiếng với nhiều bằng chứng và một thực thể ít được biết đến với dữ liệu mỏng, hệ thống thường nghiêng về phương án thứ nhất. Quyết định ấy có vẻ hợp lý ở từng trường hợp riêng lẻ, nhưng khi lặp lại hàng triệu lần, nó làm thị trường ngày càng tập trung vào một số ít lựa chọn.

Vì vậy, vòng lặp không khuếch đại mọi thứ như nhau. Nó ưu tiên những gì đã có vị trí thuận lợi trong mạng lưới dữ liệu. Một

sai lệch nhỏ về khả năng được nhìn thấy, mức độ chứng thực hoặc số lượng phản hồi ban đầu có thể dẫn đến khoảng cách rất lớn về sau. Đây là lý do cạnh tranh trong nền kinh tế tác nhân không diễn ra trên một mặt phẳng. Các địa phương bước vào hệ thống với những mức vốn thuật toán rất khác nhau. Muốn phá vỡ vòng lặp bất lợi, địa phương không thể chỉ tạo thêm truyền thông ngắn hạn. Nó phải tạo đủ dữ liệu có cấu trúc, đủ nguồn xác nhận, đủ trải nghiệm tạo phản hồi và đủ liên kết với những thực thể có thẩm quyền. Mục tiêu không chỉ là được nhìn thấy một lần, mà là tạo ra một chuỗi bằng chứng đủ mạnh để vòng lặp bắt đầu tự vận hành theo hướng có lợi.

> ▪ ***Nếu bạn là giám đốc tổ chức tiếp thị điểm đến:***
> • *Xác định ba thuộc tính mà địa phương đang bị gán sai hoặc bị gán thiếu, thay vì xây dựng một thông điệp mới hoàn toàn.*
> • *Đầu tư vào việc để các chủ thể độc lập cùng xác nhận ba thuộc tính ấy, thay vì tăng ngân sách cho kênh do chính mình sở hữu.*
> • *Chấp nhận rằng chỉ số cần theo dõi là mức độ nhất quán giữa các nguồn chứ không phải lượt tiếp cận của chiến dịch.*
> • *Đặt chu kỳ đo bằng năm chứ không bằng quý, vì ký ức thuật toán không phản ứng theo nhịp của một đợt truyền thông.*

### 4.4.4. Một mô hình động của niềm tin

Niềm tin trong hệ thống trí tuệ nhân tạo không phải là một nhãn cố định gắn vĩnh viễn với một nguồn hay một thực thể. Nó là một trạng thái luôn được cập nhật. Mỗi dữ liệu mới, mỗi mâu thuẫn, mỗi lần xác nhận và mỗi thay đổi về thời gian đều có thể làm trọng số niềm tin tăng lên hoặc giảm xuống. Một nguồn được xem là đáng tin hôm nay có thể mất uy tín ngày mai nếu thông

tin không còn cập nhật. Một trang web ít được biết đến có thể trở nên có thẩm quyền hơn nếu liên tục cung cấp dữ liệu chính xác và được nhiều nguồn độc lập xác nhận. Một thực thể từng rõ ràng có thể trở nên mơ hồ khi địa giới thay đổi, doanh nghiệp đổi tên hoặc địa chỉ không còn thống nhất. Niềm tin vì thế không phải tài sản bất biến. Nó là kết quả tạm thời của quá trình đánh giá liên tục.

Hệ thống thường hình thành niềm tin từ nhiều tín hiệu cùng lúc: nguồn gốc, độ nhất quán, độ mới, mức độ được trích dẫn, số lượng nguồn xác nhận, lịch sử chính xác và mức phù hợp với ngữ cảnh. Không tín hiệu nào đủ để quyết định một mình. Một nguồn chính thức có thể có thẩm quyền cao nhưng dữ liệu đã cũ. Một nguồn cộng đồng có thể thiếu vị thế nhưng lại cập nhật nhanh. Một thông tin được nhiều nơi lặp lại có thể phổ biến nhưng cùng bắt nguồn từ một sai lệch duy nhất.

Bởi vậy, niềm tin luôn vận động giữa các lực đối nghịch. Thẩm quyền kéo niềm tin lên. Mâu thuẫn kéo niềm tin xuống. Độ mới có thể thay thế thông tin cũ. Sự lặp lại có thể củng cố một mô tả. Một bằng chứng mạnh có thể đảo ngược kết luận trước đó. Tri thức máy hiếm khi đạt tới một trạng thái cân bằng vĩnh viễn. Nó có thể ổn định tương đối trong thời gian dài, nhưng luôn còn khả năng bị điều chỉnh khi nguồn, hành vi người dùng hoặc cơ chế truy hồi thay đổi.

Điều này đặc biệt rõ với địa phương. Một điểm đến có thể được đánh giá cao trong mùa này nhưng suy giảm trong mùa khác. Một doanh nghiệp có thể từng đáng tin nhưng thay đổi chủ sở hữu. Một lễ hội có thể được tổ chức đều đặn nhiều năm rồi đột ngột dừng lại. Một di tích có thể đổi trạng thái bảo tồn. Một tuyến đường mới có thể làm thay đổi toàn bộ khả năng tiếp cận. Mỗi thay đổi thực tế đều làm cấu trúc niềm tin phải được tính lại.

Mô hình động của niềm tin cũng có nghĩa rằng quá khứ tạo lợi thế nhưng không bảo đảm tương lai. Danh tiếng tích lũy giúp một thực thể dễ được tin hơn, nhưng không miễn trừ nó khỏi việc bị đánh giá lại. Ngược lại, một địa phương từng ít được biết đến vẫn có thể xây dựng niềm tin nếu liên tục cung cấp thông tin chính xác, nhất quán và có thể kiểm chứng. Niềm tin thuật toán vì vậy giống một dòng chảy hơn là một kho lưu trữ. Nó được hình thành, tiêu hao, phục hồi và chuyển dịch. Không có thời điểm nào địa phương có thể tuyên bố rằng mình đã "hoàn thành" việc xây dựng niềm tin số. Mỗi lần dữ liệu thay đổi, mỗi lần nguồn mới xuất hiện và mỗi lần hệ thống tạo ra mô tả mới, cấu trúc niềm tin lại được sắp xếp lại. Trong một mô hình tĩnh, quản trị niềm tin là xây dựng một hồ sơ tốt rồi duy trì nó. Trong một mô hình động, quản trị niềm tin là liên tục quan sát tín hiệu, sửa mâu thuẫn, cập nhật dữ liệu và tạo thêm bằng chứng. Địa phương không chỉ cần được tin. Nó phải liên tục tạo ra lý do để tiếp tục được tin.

- ***Trường hợp dựng lại:*** Câu ấy được một blogger đưa vào bài viết, một công ty lữ hành đưa vào phần giới thiệu tuyến, một trang tổng hợp điểm đến sao chép lại. Ba nội dung mới này quay vào hệ thống như ba nguồn độc lập, dù thực chất chúng cùng bắt nguồn từ một mệnh đề do máy sinh ra. Vòng lặp khép lại đúng như mô tả, mỗi vòng làm mệnh đề ấy nặng thêm.

## 4.5. KÝ ỨC THUẬT TOÁN VÀ HIỆN TƯỢNG KHÓA CHẶT

Nếu Chương 3 giải thích cách máy đọc thế giới và bốn mục đầu của chương này chỉ ra cách một địa phương trở thành đối tượng có thể được máy đọc, thì mục này đặt ra một câu hỏi quan trọng hơn: điều gì xảy ra sau khi AI đã đọc? Câu trả lời là thông tin không biến mất sau mỗi lần truy vấn, mà tiếp tục tồn tại trong một hệ thống ký ức phân tán của các mô hình, công cụ tìm kiếm, nền tảng và tác nhân AI. Mỗi lần được thu thập, trích dẫn, xếp hạng hay đề xuất đều để lại một dấu vết, góp phần củng cố cách mà máy sẽ nhìn nhận thực thể đó trong tương lai.

Khác với ký ức của con người vốn có thể phai nhạt hoặc được thay đổi bởi trải nghiệm mới, ký ức thuật toán được tích lũy thông qua dữ liệu, xác suất và các vòng phản hồi liên tục. Một thông tin đã được nhiều hệ thống cùng ghi nhận sẽ ngày càng có khả năng được truy hồi trở lại. Khi một mô tả xuất hiện đủ nhiều, nó không chỉ phản ánh thực tế mà còn bắt đầu định hình thực tế, AI tiếp tục đọc, người dùng tiếp tục nhìn thấy, các nền tảng tiếp tục sao chép và những dữ liệu mới lại được tạo ra dựa trên chính mô tả cũ. Ký ức vì thế không còn là kho lưu trữ thụ động mà trở thành một cơ chế tự khuếch đại.

Từ đây xuất hiện hiện tượng khóa chặt (lock-in). Khi một thực thể đã được gắn với một tập mô tả, một danh mục hay một hình ảnh nhất định, việc thay đổi nhận thức của hệ thống trở nên ngày càng khó khăn. Một địa phương từng được gắn với du lịch giá rẻ có thể tiếp tục bị AI ưu tiên mô tả như vậy ngay cả khi chất lượng dịch vụ đã thay đổi. Một doanh nghiệp từng bị gắn với thông tin sai lệch có thể mất nhiều năm mới đảo ngược được vị trí của mình trong ký ức thuật toán. Ngược lại, những thực thể đã đạt được lợi thế ban đầu thường tiếp tục được củng cố bởi chính các vòng lặp của hệ thống.

Đối với quản trị địa phương, đây là một thay đổi mang tính chiến lược. Cạnh tranh không còn chỉ là tạo ra trải nghiệm tốt hơn trong hiện tại, mà còn là quản trị ký ức số mà các hệ thống AI sẽ sử dụng trong tương lai. Mỗi dữ liệu được công bố, mỗi thực thể được định danh, mỗi lần được trích dẫn hay xác nhận đều là một khoản đầu tư vào ký ức thuật toán. Một khi ký ức ấy đã hình thành và bị khóa chặt, chi phí để thay đổi sẽ lớn hơn rất nhiều so với chi phí xây dựng đúng ngay từ đầu. Năm tiểu mục tiếp theo sẽ phân tích cơ chế hình thành của ký ức thuật toán, lý do vì sao các hệ thống AI có xu hướng bảo lưu những mô tả cũ, hiện tượng khóa chặt trong tri thức số, những con đường để phá vỡ vòng lặp này và vì sao quản trị ký ức thuật toán sẽ trở thành một năng lực cạnh tranh cốt lõi của địa phương trong nền kinh tế tác nhân.

### 4.5.1. Ký ức thuật toán là gì

Khi nói đến "ký ức" của AI, nhiều người hình dung đến một cơ sở dữ liệu khổng lồ nơi mọi thông tin được lưu giữ như trong bộ nhớ của máy tính. Cách hình dung này chỉ đúng một phần. Điều mà các hệ thống AI hiện đại duy trì không phải là một bộ nhớ theo nghĩa con người hiểu, mà là một lịch sử tích lũy của những gì đã nhiều lần được lựa chọn, xác nhận và sử dụng. Đó chính là ký ức thuật toán (algorithmic memory).

Con người nhớ bằng trải nghiệm. Một sự kiện để lại dấu ấn vì nó gắn với cảm xúc, bối cảnh hay ý nghĩa cá nhân. Máy thì không có trải nghiệm. Điều máy "nhớ" là các dấu vết xác suất được tạo nên từ hàng triệu lần truy hồi, xếp hạng, trích dẫn và phản hồi. Mỗi lần một thực thể được AI đề xuất, mỗi lần một bài viết được nhiều hệ thống thu thập, mỗi lần một nguồn được coi là đáng tin cậy, hệ thống lại có thêm bằng chứng để tiếp tục lựa chọn nó trong tương lai.

Có thể định nghĩa như sau. Ký ức thuật toán là phần quá khứ tiếp tục ảnh hưởng đến các lần truy hồi, mô tả, xếp hạng và lựa chọn trong tương lai. Nó không nằm ở một nơi duy nhất mà phân tán trong dữ liệu huấn luyện, trọng số của các tham số trong mô hình, chỉ mục tìm kiếm, đồ thị tri thức, thứ hạng và lịch sử sử dụng. Một thông tin có thể bền vững vì nằm trong dữ liệu huấn luyện, vì được neo trong đồ thị tri thức, vì xuất hiện ở một nguồn có thẩm quyền, vì được xếp hạng cao hay đơn giản vì được người dùng lặp lại nhiều lần. Trong số các cơ chế ấy, lịch sử được chọn là cơ chế mà chương này đặc biệt quan tâm. Ký ức thuật toán không ghi lại mọi thứ mà ghi đậm những gì nhiều lần thắng trong quá trình lựa chọn thông qua những mô tả được truy hồi thường xuyên, được trích dẫn lại, được dùng làm cơ sở cho các suy luận tiếp theo. Mỗi lần được chọn là một lần xác suất xuất hiện trong lần chọn sau tăng thêm một chút.

Đối với địa phương, điều này có ý nghĩa đặc biệt. Một lễ hội, một làng nghề hay một di sản không trở thành một phần của ký ức thuật toán chỉ vì nó tồn tại trong thực tế. Nó chỉ bước vào ký ức ấy khi liên tục được định danh, được chuẩn hóa, được nhiều nguồn độc lập xác nhận và được các hệ thống AI thường xuyên truy hồi. Trong nền kinh tế tác nhân, cạnh tranh vì thế không chỉ là cạnh tranh để được xuất hiện một lần, mà là cạnh tranh để liên tục được lựa chọn, bởi chính lịch sử những lần được chọn sẽ quyết định khả năng được chọn trong tương lai.

### 4.5.2. Đường dẫn tạo nên tương lai

Một trong những phát hiện quan trọng của kinh tế học thể chế và khoa học hệ thống là quá khứ không chỉ giải thích hiện tại mà còn giới hạn những tương lai có thể xảy ra. Hiện tượng này được gọi là

phụ thuộc đường dẫn (path dependence): Những quyết định ban đầu, dù được đưa ra trong những điều kiện rất khác, có thể tiếp tục định hình kết quả nhiều năm, thậm chí nhiều thập kỷ sau đó.

Lịch sử công nghệ có rất nhiều ví dụ. Bàn phím QWERTY không nhất thiết là bố cục tối ưu xét theo mọi tiêu chí, nhưng sau hơn một thế kỷ đào tạo, sản xuất và tiêu chuẩn hóa, chi phí chuyển sang một bố cục khác đã trở nên quá lớn. Đáng chú ý là cơ chế đằng sau câu chuyện công nghệ ấy. Trong không gian số của một địa phương, tiêu chuẩn thắng cuộc không phải bàn phím hay định dạng băng hình, mà là mô tả đầu tiên đủ phổ biến để trở thành điểm quy chiếu cho mọi mô tả về sau. Trong cuộc chiến giữa hai chuẩn băng video, Betamax thường được đánh giá cao ở một số tiêu chí chất lượng hình ảnh, nhưng VHS giành được thị trường nhờ thời lượng ghi hình, mạng lưới phân phối và sự ủng hộ rộng hơn của các nhà sản xuất. Khi một tiêu chuẩn đã vượt qua một ngưỡng nhất định, lợi thế không còn nằm ở chất lượng kỹ thuật mà ở chính số lượng người đang sử dụng nó.

Điều tương tự đang diễn ra trong hệ sinh thái AI. Một địa phương được nhiều hệ thống biết đến từ sớm sẽ có nhiều dữ liệu hơn, nhiều lượt truy hồi hơn, nhiều trích dẫn hơn và nhiều cơ hội tiếp tục được AI đề xuất. Mỗi vòng lặp mới lại củng cố kết quả của vòng lặp trước, khiến khoảng cách giữa các thực thể ngày càng lớn. Đây chính là sự kết hợp giữa phụ thuộc đường dẫn và hiệu ứng Matthew đã được trình bày ở mục trước. Vì vậy, AI không chỉ đọc trạng thái hiện tại của thế giới mà còn đọc cả lịch sử tích lũy của thế giới đó. Một địa phương mới xây dựng thương hiệu hôm nay không cạnh tranh trên một mặt bằng hoàn toàn bằng phẳng. Họ phải cạnh tranh với những ký ức thuật toán đã được tích lũy suốt nhiều năm của các đối thủ. Càng để chậm, chi phí thay đổi nhận thức của hệ thống càng cao.

Đây là lý do vì sao quản trị dữ liệu trong nền kinh tế tác nhân phải được nhìn như một chiến lược dài hạn. Những gì được công bố hôm nay sẽ trở thành một phần của đường dẫn ngày mai. Những gì bị bỏ qua hôm nay có thể rất khó bổ sung trong tương lai. Khi ký ức thuật toán ngày càng tích lũy và hiện tượng khóa chặt ngày càng mạnh, lợi thế cạnh tranh không chỉ thuộc về người làm tốt nhất ở hiện tại mà còn thuộc về người xây dựng đúng đường dẫn ngay từ đầu. Mô tả xuất hiện sớm và đủ phổ biến sẽ trở thành điểm quy chiếu cho các mô tả sau. Vì mỗi lần được tham chiếu là một lần được củng cố, thứ tự xuất hiện có ảnh hưởng đến thứ hạng về sau ngay cả khi không có nguyên nhân mới nào được thêm vào.

### 4.5.3. Hiệu ứng khóa chặt

Một trong những đặc tính quan trọng nhất của ký ức thuật toán là hiệu ứng khóa chặt (lock-in). Đây là trạng thái mà sau khi một thực thể đã được hệ thống chấp nhận như một "mặc định", việc thay đổi nhận thức của máy trở nên rất khó khăn, ngay cả khi thực tế bên ngoài đã thay đổi đáng kể. Hiệu ứng này xuất hiện vì các hệ thống AI không đưa ra quyết định dựa trên một nguồn thông tin duy nhất. Chúng tổng hợp từ hàng loạt tín hiệu, dữ liệu huấn luyện, kết quả tìm kiếm, đồ thị tri thức, đánh giá của người dùng, các trang web có thẩm quyền, bài báo, mạng xã hội và lịch sử tương tác. Khi tất cả các tín hiệu này cùng chỉ về một mô tả giống nhau, hệ thống có xu hướng coi đó là trạng thái đáng tin cậy nhất. Mỗi lần được lựa chọn lại làm tăng thêm xác suất tiếp tục được lựa chọn trong tương lai.

Do đó, khóa chặt là hệ quả tự nhiên của quá trình tối ưu hóa, không phải một lỗi của AI. Một hệ thống liên tục thay đổi nhận

thức theo từng thông tin mới sẽ trở nên thiếu ổn định và dễ bị thao túng. Ngược lại, một hệ thống ưu tiên những tri thức đã được xác nhận nhiều lần sẽ có độ nhất quán cao hơn. Chính cơ chế tạo ra sự ổn định ấy đồng thời cũng làm cho sự thay đổi trở nên chậm chạp. Điều này giải thích vì sao việc "đính chính" một thông tin thường khó hơn rất nhiều so với việc tạo ra thông tin ban đầu. Một bài viết sai có thể chỉ mất vài phút để xuất bản, nhưng để toàn bộ hệ sinh thái AI ngừng coi thông tin đó là đúng có thể phải mất nhiều tháng hoặc nhiều năm. Khóa chặt vì thế là một dạng quán tính của tri thức số, nơi quá khứ tiếp tục chi phối cách máy nhìn nhận hiện tại.

### 4.5.4. Địa phương cũng bị khóa chặt

Hiện tượng khóa chặt không chỉ xảy ra với sản phẩm, doanh nghiệp hay công nghệ mà còn xảy ra với thương hiệu của địa phương. Một khi một hình ảnh đã được hệ thống AI và các nền tảng số chấp nhận như cách mô tả mặc định, địa phương rất khó thoát khỏi khuôn mẫu đó.

Một địa phương từng nổi tiếng với du lịch giá rẻ có thể tiếp tục được AI ưu tiên đề xuất cho nhóm khách tìm kiếm chi phí thấp, ngay cả khi nhiều năm sau nơi đó đã đầu tư vào hệ thống nghỉ dưỡng cao cấp. Một thành phố từng được biết đến là điểm đến của du lịch bụi có thể mất rất lâu mới được nhận diện như một trung tâm hội nghị hay điểm đến nghỉ dưỡng sang trọng. Tương tự, một vùng đất gắn với ẩm thực, di sản, biển, hay làng nghề thường tiếp tục được AI nhấn mạnh chính những đặc điểm đó, trong khi các giá trị mới khó có cơ hội nổi bật.

Cần phân biệt ba trường hợp ở đây. Mô tả bị khóa chặt có thể sai hoàn toàn, có thể từng đúng nhưng đã lỗi thời, hoặc vẫn đang

đúng nhưng quá hẹp so với hiện tại. Trường hợp thứ ba khó xử lý nhất, bởi địa phương không thể phản bác một điều không sai. Nguyên nhân không nằm ở việc AI "cố chấp", mà vì toàn bộ hệ sinh thái thông tin vẫn đang kể cùng một câu chuyện cũ. Hàng nghìn bài báo, bài viết, đánh giá, video, bản đồ và dữ liệu có cấu trúc đều lặp lại cùng một mô tả. Khi AI tổng hợp các nguồn này, kết quả hợp lý nhất về mặt xác suất vẫn là hình ảnh đã được tích lũy trong nhiều năm.

Đối với quản trị địa phương, đây là một thách thức lớn. Xây dựng một sản phẩm du lịch mới hay một ngành kinh tế mới chỉ là bước đầu. Quan trọng không kém là làm cho ký ức thuật toán của thế giới cũng thay đổi theo. Điều đó đòi hỏi một chiến lược dài hạn về dữ liệu, truyền thông, định danh thực thể và sự xuất hiện nhất quán trên nhiều nguồn độc lập. Thương hiệu địa phương vì thế không chỉ tồn tại trong tâm trí con người mà còn phải được tái cấu trúc trong trí nhớ của các hệ thống AI.

Các mô hình được huấn luyện lại liên tục nên khóa chặt chỉ là tạm thời? Mỗi thế hệ mô hình mới lại học từ dữ liệu mới hơn. Nếu địa phương thay đổi thực chất, chỉ cần chờ vài vòng huấn luyện là mô tả sẽ được cập nhật. Thực tế, huấn luyện lại thay dữ liệu chứ không thay phân bố của dữ liệu. Nếu hệ sinh thái bên ngoài vẫn đang tạo ra hàng nghìn nội dung mới lặp lại mô tả cũ thì đợt huấn luyện kế tiếp sẽ học đúng mô tả ấy một lần nữa, thậm chí với bằng chứng dày hơn trước. Việc chờ đợi chỉ hiệu quả khi trong khoảng thời gian chờ, tỷ lệ bằng chứng mới trên bằng chứng cũ đã đảo chiều. Nếu địa phương không chủ động tạo ra sự đảo chiều ấy thì mỗi vòng huấn luyện lại là một lần khóa chặt được gia cố. Ký ức thuật toán thay đổi chậm hơn thực tại vì nó dựa trên khối bằng chứng đã tích lũy chứ không dựa trên trạng thái mới nhất. Trong khoảng chênh giữa hai tốc độ ấy, mô tả cũ vẫn tiếp tục được sử dụng như thể nó đang đúng.

### 4.5.5. Ký ức mạnh hơn sự thật hiện tại

Trong đời sống thường nhật, con người thường tin rằng sự thật hiện tại sẽ tự động thay thế những nhận thức cũ. Đối với hệ thống AI, điều này không phải lúc nào cũng đúng. Thực tại có thể đã thay đổi, nhưng ký ức thuật toán vẫn tiếp tục phản ánh quá khứ.

Một khách sạn đã được cải tạo toàn diện, một khu phố đã trở nên an toàn hơn, một điểm đến đã phát triển hệ thống dịch vụ chất lượng cao, hay một địa phương đã thay đổi chiến lược phát triển du lịch. Nếu những thay đổi này chưa được phản ánh rộng rãi trên các nguồn dữ liệu mà AI thường sử dụng, máy vẫn có xu hướng trả lời theo hình ảnh cũ. Không phải vì AI không biết thông tin mới tồn tại, mà vì bằng chứng tích lũy của quá khứ vẫn mạnh hơn bằng chứng còn hạn chế của hiện tại.

Đây là một nguyên lý quen thuộc trong khoa học nhận thức và học máy, niềm tin chỉ thay đổi khi lượng bằng chứng mới đủ lớn để vượt qua trọng số của bằng chứng cũ. Trong hệ sinh thái AI, điều đó đồng nghĩa với việc một vài bài báo mới hay một chiến dịch truyền thông ngắn hạn hiếm khi đủ sức thay đổi ký ức thuật toán đã được xây dựng suốt nhiều năm. Vì vậy, quản trị địa phương trong nền kinh tế tác nhân không thể chỉ tập trung vào việc thay đổi thực tế, mà còn phải thay đổi ký ức số về thực tế đó. Hai quá trình này không diễn ra đồng thời. Một địa phương có thể đã đổi mới rất nhanh, nhưng AI và hàng triệu người dùng vẫn tiếp tục nhìn thấy phiên bản của nhiều năm trước. Khoảng trễ giữa thực tại và ký ức thuật toán chính là một trong những nguyên nhân khiến nhiều nỗ lực tái định vị thương hiệu không đạt được hiệu quả như mong đợi.

Đó cũng là lý do vì sao các địa phương không thể chỉ quản lý hạ tầng vật chất, dịch vụ hay trải nghiệm, mà còn phải quản lý lịch sử dữ liệu của chính mình. Trong nền kinh tế tác nhân, chiến thắng không chỉ thuộc về nơi thay đổi nhanh nhất, mà còn thuộc về nơi làm cho hệ sinh thái AI nhận ra sự thay đổi ấy nhanh nhất. Ký ức thuật toán, vì thế, không chỉ là quá khứ của dữ liệu mà còn là một yếu tố quyết định tương lai của năng lực cạnh tranh.

***Dựng lại:*** Ba năm sau, làng đã có mười hai hộ quay lại nghề, một hợp tác xã và một xưởng trưng bày. Nhưng mô tả mặc định vẫn là làng nghề đang mai một. Đây thuộc trường hợp thứ ba đã nói ở trên, mệnh đề ấy không sai khi ra đời, nhưng không còn đủ để phản ánh thực tại hiện nay. Chính vì không sai, địa phương không có gì để đính chính. Cũng vì thế mà việc thay đổi nó khó hơn nhiều so với sửa một con số.

## 4.6. ĐỊNH VỊ LẠI TRONG MỘT HỆ THỐNG CÓ KÝ ỨC

Một trong những ngộ nhận phổ biến của quản trị thương hiệu là tin rằng hình ảnh của một địa phương có thể được thay đổi chỉ bằng một chiến dịch truyền thông mới. Thực tế, điều đó có thể đúng khi đối tượng tiếp nhận là con người, nhưng không còn đúng trong nền kinh tế tác nhân. Đối với các hệ thống AI, mỗi địa phương không chỉ có một hình ảnh hiện tại mà còn mang theo toàn bộ "lịch sử số" của chính mình: những mô tả đã từng được xuất bản, những đánh giá đã từng được ghi nhận, những lần được đề xuất, những thực thể đã được liên kết và cả những chuỗi suy luận đã nhiều lần được lặp lại.

Chính vì vậy, định vị lại không còn đơn thuần là xây dựng một câu chuyện mới, mà là thay đổi một ký ức đã được tích lũy trong các hệ thống thuật toán. Mỗi lần AI truy hồi dữ liệu cũ, mỗi lần một bài báo cũ được trích dẫn, mỗi lần một người dùng lặp lại mô tả quen thuộc, hình ảnh trước đây lại được củng cố thêm một lần. Quá khứ không biến mất khi hiện tại thay đổi. Nó tiếp tục tồn tại trong bộ nhớ của các mô hình, các cơ sở tri thức và các mạng lưới dữ liệu mà AI sử dụng để suy luận.

Vì vậy, mục này phân tích một nghịch lý của quản trị địa phương trong nền kinh tế tác nhân, xây dựng thương hiệu đã khó, nhưng thay đổi thương hiệu còn khó hơn. Khi thuật toán tích lũy ký ức qua dữ liệu và các vòng lặp suy luận, mỗi định vị hôm nay có thể trở thành lợi thế hoặc gánh nặng của nhiều năm sau. Vì thế, quản trị danh tiếng số không còn là truyền thông ngắn hạn, mà là quá trình tích lũy và điều chỉnh ký ức thuật toán.

Mức độ khó của việc tái định vị cũng khác nhau. Thay đổi logo, khẩu hiệu hay bộ nhận diện là dễ nhất. Thay đổi thông tin

thực thể như tên gọi, địa giới hay giờ mở cửa khó hơn vì phải được cập nhật đồng bộ trên các nguồn chính thức. Khó hơn nữa là thay đổi thuộc tính chức năng, chẳng hạn từ điểm đến mùa vụ thành điểm đến quanh năm, vì cần dữ liệu vận hành mới trong thời gian dài. Khó nhất là thay đổi liên tưởng cốt lõi như từ "giá rẻ" sang "cao cấp" hay từ "ô nhiễm" sang "sống xanh", bởi điều đó đòi hỏi xây dựng lại toàn bộ tập bằng chứng mà máy dùng để hình thành niềm tin.

### 4.6.1. Thay đổi thương hiệu dễ hơn thay đổi ký ức

Một địa phương có thể thay đổi logo chỉ trong vài tháng. Một khẩu hiệu mới có thể được công bố trong một buổi lễ. Một bộ nhận diện mới có thể xuất hiện đồng loạt trên biển hiệu, trang web và các tài liệu truyền thông. Về mặt kỹ thuật, việc thay đổi thương hiệu chưa bao giờ dễ dàng hơn thế.

Nhưng những gì khó thay đổi lại không phải là biểu tượng, mà là ký ức. Trong tâm trí con người, niềm tin được hình thành qua nhiều lần trải nghiệm. Một người từng có ấn tượng xấu với một điểm đến sẽ không dễ thay đổi chỉ vì nhìn thấy một logo mới. Họ cần những trải nghiệm mới, những câu chuyện mới và những bằng chứng mới đủ sức phủ định ký ức cũ. Niềm tin vì vậy luôn thay đổi chậm hơn hình ảnh nhận diện.

Đối với AI, khoảng cách ấy còn lớn hơn. Máy có thể nhìn thấy logo, nhưng logo chỉ là một dữ kiện rất nhỏ và không đủ sức thay đổi lịch sử dữ liệu đứng phía sau nó. Khi được hỏi về một địa phương, AI truy hồi hàng nghìn nguồn thông tin đã được tích lũy trong nhiều năm: bài báo, đánh giá của người dùng, dữ liệu bản đồ, hồ sơ thực thể, cơ sở tri thức, các lần được nhắc đến

và vô số liên kết giữa chúng. Tất cả những dữ liệu đó hợp thành một "ký ức thuật toán" về địa phương.

Nếu một địa phương từng được biết đến như một điểm đến giá rẻ, một khu công nghiệp, một nơi chỉ nổi tiếng với một loại đặc sản hay một vùng thường xuyên ô nhiễm, thì những mô tả ấy sẽ tiếp tục xuất hiện trong quá trình suy luận của AI, ngay cả khi thực tế đã thay đổi. Logo mới không làm thay đổi xác suất mà mô hình đã học. Chỉ có một lượng lớn bằng chứng mới, được tích lũy trong thời gian dài, mới có thể làm điều đó.

Đây là lý do thương hiệu luôn có hai tầng tồn tại. Tầng thứ nhất là thương hiệu mà địa phương tự công bố thông qua chiến lược truyền thông. Tầng thứ hai là thương hiệu mà hệ thống tri thức của thế giới đang ghi nhớ. Trong nền kinh tế tác nhân, tầng thứ hai mới là yếu tố quyết định địa phương có được AI đề xuất hay không. Có thể nói, thay đổi thương hiệu là thay đổi lời giới thiệu. Thay đổi ký ức là thay đổi toàn bộ lịch sử mà máy dùng để giới thiệu. Việc đầu tiên có thể hoàn thành trong một năm, nhưng việc thứ hai thường cần nhiều năm tích lũy dữ liệu, bằng chứng và sự nhất quán.

### 4.6.2. Vì sao tái định vị luôn tốn kém

Các chiến dịch tái định vị thường được hiểu như một hoạt động truyền thông, thay đổi thông điệp, làm mới hình ảnh, tổ chức sự kiện và quảng bá rộng rãi. Tuy nhiên, trong nền kinh tế tác nhân, truyền thông chỉ là phần nổi của vấn đề. Điều thực sự phải thay đổi là hệ tri thức mà các hệ thống AI đang sử dụng để suy luận.

Một AI không tin một địa phương vì đã xem quảng cáo của địa phương đó. Nó tin vì nhiều nguồn độc lập cùng xác nhận một mô tả

giống nhau. Do đó, tái định vị không thể dừng lại ở việc phát đi một câu chuyện mới, mà phải tạo ra một mạng lưới bằng chứng mới đủ mạnh để thay thế mạng lưới bằng chứng cũ. Điều này đồng nghĩa với việc địa phương phải đồng thời thay đổi ở nhiều tầng. Hệ thống phân loại phải được cập nhật để phản ánh đúng bản sắc mới. Hồ sơ thực thể trên các nền tảng phải được bổ sung. Dữ liệu có cấu trúc và phi cấu trúc phải nhất quán. Báo chí, trang web chính thức, bản đồ số, cộng đồng, doanh nghiệp và các nền tảng du lịch phải dần sử dụng cùng một cách mô tả. Chỉ khi các nguồn tri thức độc lập cùng hội tụ, AI mới bắt đầu điều chỉnh mô hình suy luận của mình.

Quá trình đó tiêu tốn nhiều nguồn lực hơn rất nhiều so với một chiến dịch truyền thông truyền thống. Nó đòi hỏi đầu tư vào dữ liệu, chuẩn hóa bản thể luận, quản trị thực thể, xây dựng nguồn dữ liệu có thẩm quyền, duy trì nội dung trong nhiều năm và liên tục tạo ra các bằng chứng mới để củng cố định vị mong muốn. Nói cách khác, chi phí lớn nhất của tái định vị không nằm ở quảng cáo, mà nằm ở việc tái cấu trúc toàn bộ hạ tầng tri thức số của địa phương. Đó cũng là lý do nhiều địa phương tuy đã đổi khẩu hiệu, đổi bộ nhận diện và đầu tư mạnh cho truyền thông nhưng vẫn không thay đổi được cách AI mô tả mình. Hình ảnh mới tồn tại trên các ấn phẩm quảng bá, còn ký ức cũ vẫn tồn tại trong các hệ thống tri thức. Hai thế giới ấy song song tồn tại và trong quá trình suy luận của AI, ký ức được tích lũy lâu dài thường có trọng số lớn hơn lời khẳng định mới xuất hiện.

Vì vậy, tái định vị trong nền kinh tế tác nhân thực chất là quá trình sửa đổi một hệ tri thức, chứ không chỉ là làm mới một chiến dịch truyền thông. Khi nhìn từ góc độ này, truyền thông chỉ là điểm khởi đầu. Nhiệm vụ quan trọng hơn là tạo ra một tập hợp bằng chứng mới đủ lớn để dần thay thế ký ức cũ trong bộ nhớ của thế giới số.

### 4.6.3. Một thuộc tính sai kéo theo hàng trăm mô tả sai

Trong những hệ thống sử dụng đồ thị tri thức, dữ liệu không tồn tại dưới dạng những bài viết rời rạc mà được tổ chức thành một mạng lưới thực thể và quan hệ. Mỗi địa phương, mỗi di tích, mỗi doanh nghiệp hay mỗi sản phẩm đều là một thực thể. Các đặc điểm của chúng được biểu diễn bằng các thuộc tính và được kết nối với vô số thực thể khác trong đồ thị tri thức (knowledge graph). Chính cấu trúc này cho phép AI không chỉ truy hồi thông tin mà còn suy luận từ những gì đã biết. Điều đó cũng có nghĩa là một sai sót nhỏ ở tầng dữ liệu gốc có thể lan truyền rất xa. Nếu một thuộc tính nền tảng của thực thể bị ghi nhận sai, mọi phép suy luận dựa trên thuộc tính đó đều có nguy cơ sai theo.

Hãy hình dung một địa phương được phân loại là "điểm đến du lịch bụi giá rẻ". Thuộc tính này không chỉ xuất hiện trong hồ sơ của chính địa phương đó. Nó còn ảnh hưởng đến cách AI lựa chọn khách sạn để đề xuất, mức chi tiêu mà AI dự đoán, loại hình trải nghiệm được ưu tiên, nhóm du khách được nhắm tới, thậm chí cả những địa phương "tương tự" được đưa vào cùng một tập so sánh. Một nhãn phân loại ban đầu dần trở thành hàng trăm mô tả khác nhau thông qua quá trình suy luận.

Đây chính là đặc điểm của đồ thị tri thức, thông tin không đứng yên mà liên tục tạo ra thông tin mới. Một thuộc tính sai không chỉ tạo ra một câu trả lời sai. Nó có thể tạo ra cả một chuỗi suy luận sai. Trong khoa học dữ liệu, hiện tượng này thường được gọi là lan truyền sai lệch (propagation). Sai lệch ban đầu được truyền qua các mối quan hệ, được sử dụng trong các mô hình xếp hạng, được AI tổng hợp lại thành các mô tả mới và tiếp tục trở thành nguồn dữ liệu cho những lần suy luận sau.

Đối với quản trị thương hiệu địa phương, đây là một thay đổi mang tính nền tảng. Điều cần bảo vệ không chỉ là hình ảnh bên ngoài, mà còn là tính chính xác của những thuộc tính cốt lõi cấu thành thực thể số của địa phương. Bởi một sai sót ở tầng định danh có thể âm thầm ảnh hưởng đến hàng trăm quyết định mà không ai nhận ra nguyên nhân ban đầu.

Nếu chi phí sửa cao đến vậy thì mọi tái định vị đều vô vọng hay không? Một địa phương đã bị gán nhãn từ nhiều năm sẽ không bao giờ đủ nguồn lực để tạo ra khối lượng bằng chứng ngược lại. Kết luận hợp lý là nên chấp nhận nhãn cũ và tối ưu hóa trong phạm vi đó. Thực tế, chi phí cao không có nghĩa là không thể trả, mà có nghĩa là phải trả đúng chỗ. Sai lầm phổ biến là phân bổ ngân sách theo tỷ lệ ngược, phần lớn cho chiến dịch truyền thông ngắn hạn, phần nhỏ cho dữ liệu nền. Thứ tự đúng là ngược lại, vì một thuộc tính được sửa ở tầng dữ liệu gốc sẽ tự lan sang hàng trăm mô tả dẫn xuất, còn một chiến dịch dừng lại khi ngân sách dừng lại. Ngoài ra, tái định vị không đòi hỏi xóa nhãn cũ mà đòi hỏi bổ sung một nhãn mới đủ bằng chứng để cùng tồn tại. Đó là bài toán khả thi hơn nhiều.

### 4.6.4. Quản trị thương hiệu trở thành quản trị ký ức

Trong nhiều thập kỷ, quản trị thương hiệu chủ yếu được hiểu là quản trị truyền thông. Công việc của các tổ chức là xây dựng thông điệp, thiết kế nhận diện, tổ chức chiến dịch quảng bá và tạo ra những câu chuyện hấp dẫn để định hình nhận thức của công chúng. Trong nền kinh tế tác nhân, phạm vi ấy không còn đủ rộng. Các hệ thống AI không ghi nhớ chiến dịch theo cách con người nhớ một thông điệp. Chúng hấp thụ những dấu vết và bằng chứng mà chiến dịch để lại và một chiến dịch tạo ra báo chí, liên kết và nội dung mới thì cũng tạo ra bằng chứng. Điều

được tích lũy không phải là khẩu hiệu quảng cáo mà là các thực thể, thuộc tính, mối quan hệ, dữ liệu có cấu trúc, bài báo, đánh giá, hồ sơ bản đồ, cơ sở tri thức và hàng triệu dấu vết số khác. Toàn bộ chúng hợp thành ký ức thuật toán của một địa phương.

Điều đó làm thay đổi bản chất của quản trị thương hiệu. Nhiệm vụ không còn dừng ở việc tạo ra một hình ảnh hấp dẫn, mà là duy trì một hệ tri thức đúng, nhất quán và có thẩm quyền. Mỗi thông tin sai cần được phát hiện sớm. Mỗi thuộc tính mới cần được cập nhật đồng bộ. Mỗi thay đổi của địa phương phải nhanh chóng phản ánh trong các hệ thống dữ liệu mà AI đang sử dụng. Theo cách nhìn này, thương hiệu không còn là một tài sản chỉ tồn tại trong nhận thức của con người. Nó trở thành một tài sản tri thức tồn tại trong các mạng lưới dữ liệu của thế giới số. Giá trị của thương hiệu không chỉ phụ thuộc vào mức độ nổi tiếng, mà còn phụ thuộc vào độ chính xác, tính nhất quán và khả năng được các hệ thống AI truy hồi như một nguồn đáng tin cậy.

Vì vậy, trong nền kinh tế tác nhân, quản trị thương hiệu thực chất là quản trị ký ức. Đó là quá trình liên tục xây dựng, kiểm chứng, cập nhật và bảo vệ ký ức số của địa phương để mỗi lần AI truy hồi đều phản ánh đúng thực tế hiện tại. Khi thương hiệu được hiểu như một hệ tri thức sống, công việc của nhà quản trị không còn kết thúc sau một chiến dịch truyền thông, mà chỉ kết thúc khi ký ức của cả hệ sinh thái số đã thực sự thay đổi.

- ***Dựng lại:*** Khi tỉnh muốn đưa làng vào tuyến du lịch trải nghiệm, họ phát hiện việc phải làm không nằm ở cấp độ biểu tượng. Đây là thay đổi liên tưởng cốt lõi, từ một nơi đang tàn lụi thành một nơi đang hồi sinh. Một video quảng bá không đủ. Cần một dòng bằng chứng mới, đều đặn, từ nhiều nguồn có thẩm quyền, kéo dài đủ lâu để làm thay đổi tương quan giữa bằng chứng mới và khối bằng chứng đã tích lũy ba năm.

## 4.7. SỬA SỚM HAY KHÔNG SỬA ĐƯỢC

Một trong những khác biệt lớn nhất giữa quản trị thông tin truyền thống và quản trị tri thức trong nền kinh tế tác nhân nằm ở yếu tố thời gian. Nếu trước đây, một thông tin sai có thể được cải chính bằng một thông cáo báo chí hoặc một bài viết mới, thì ngày nay, mỗi ngày chậm trễ đều làm tăng số lượng hệ thống đã ghi nhớ sai lệch đó.

Thuật toán không chỉ đọc dữ liệu. Thuật toán còn tích lũy dữ liệu. Mỗi lần một thông tin được truy hồi, được trích dẫn hay được sử dụng trong suy luận, nó lại trở thành một phần của ký ức số. Vì vậy, tốc độ phát hiện và sửa sai quan trọng không kém chất lượng của việc sửa sai. Trong nhiều trường hợp, sai sót không nguy hiểm vì bản thân nó, mà vì nó đã có đủ thời gian để lan truyền.

Chính vì vậy, quản trị thương hiệu trong nền kinh tế tác nhân không thể chỉ là quản trị hình ảnh sau khi khủng hoảng xảy ra. Nó phải trở thành một hệ thống giám sát và bảo trì tri thức hoạt động liên tục. Mục tiêu là ngăn chặn sai lệch trước khi chúng trở thành ký ức mặc định của các hệ thống AI, không phải xử lý khủng hoảng truyền thông.

### 4.7.1. Sai lúc mới xuất hiện dễ sửa nhất

Mọi sai lệch đều có một "thời điểm vàng" để sửa chữa. Khi một thông tin sai mới xuất hiện trên một trang web, một bài báo hoặc một cơ sở dữ liệu, phạm vi ảnh hưởng của nó còn rất nhỏ. Nó chưa được các công cụ tìm kiếm lập chỉ mục đầy đủ, chưa được các nền tảng khác sao chép và cũng chưa được AI sử dụng rộng rãi trong quá trình truy hồi hay suy luận. Đây chính là giai đoạn của hiệu chỉnh sớm (early correction). Chỉ cần sửa tại

nguồn và cập nhật lại dữ liệu, sai lệch có thể bị loại bỏ trước khi hình thành chuỗi lan truyền.

Trong quản trị dữ liệu, thời gian phản ứng thường quan trọng hơn quy mô của lỗi. Một lỗi nhỏ được sửa trước khi nó kịp được lập chỉ mục, sao chép hoặc dùng làm nguồn cho nội dung phái sinh có thể gần như không để lại dấu vết. Điều quyết định không phải số ngày mà là giai đoạn lan truyền: một nguồn có lưu lượng lớn có thể bị thu thập trong vài phút, trong khi một trang ít người đọc có thể nằm yên nhiều tháng. Cùng một lỗi, nếu vượt qua ngưỡng lan truyền, sẽ bắt đầu xuất hiện trên các công cụ tìm kiếm, hệ thống bản đồ, bài viết tổng hợp và cơ sở tri thức của nhiều nền tảng khác nhau.

Điều này cho thấy quản trị thương hiệu số không còn là hoạt động phản ứng theo sự kiện, mà phải là một cơ chế phát hiện sớm. Địa phương càng phát hiện sai lệch nhanh thì chi phí sửa chữa càng thấp và khả năng khôi phục sự thật càng cao.

Các nhà làm chính sách có thể phản biện rằng giám sát liên tục là gánh nặng mà địa phương nhỏ không kham nổi. Việc quan sát thường xuyên cách địa phương được mô tả trên các hệ thống đòi hỏi nhân sự chuyên trách, công cụ đo lường và ngân sách duy trì. Phần lớn địa phương không có những thứ đó. Tuy nhiên, quy mô của công việc này tỷ lệ với số thực thể lõi chứ không tỷ lệ với quy mô địa phương. Một xã có thể chỉ có vài chục thực thể thực sự quan trọng. Việc rà soát chúng theo quý là công việc của một người trong vài ngày, không phải của một phòng ban. Chi phí thật sự không nằm ở giám sát mà nằm ở việc không giám sát, một sai lệch bị bỏ qua trong hai năm sẽ tiêu tốn nhiều hơn toàn bộ chi phí giám sát của hai năm ấy cộng lại. Đây là lập luận quen thuộc của bảo trì dự phòng, áp dụng cho dữ liệu.

▪ ***Nếu bạn là Chủ tịch Ủy ban nhân dân cấp tỉnh:***

• *Yêu cầu một báo cáo định kỳ về việc các hệ thống trí tuệ nhân tạo đang mô tả tỉnh như thế nào, đặt cạnh báo cáo về lượt khách và vốn đầu tư.*

• *Giao trách nhiệm dữ liệu cho một đầu mối có thẩm quyền điều phối liên ngành, vì sai lệch thường nằm ở khoảng giữa các sở.*

• *Đưa yêu cầu về định danh và dữ liệu có cấu trúc vào điều kiện nghiệm thu của mọi dự án chuyển đổi số, thay vì xử lý sau khi hệ thống đã bàn giao.*

• *Không phê duyệt một đề án tái định vị nếu đề án đó không nêu được hiện trạng mô tả mà máy đang lưu giữ.*

### 4.7.2. Sau khi được trích dẫn nhiều lần, chi phí tăng phi tuyến

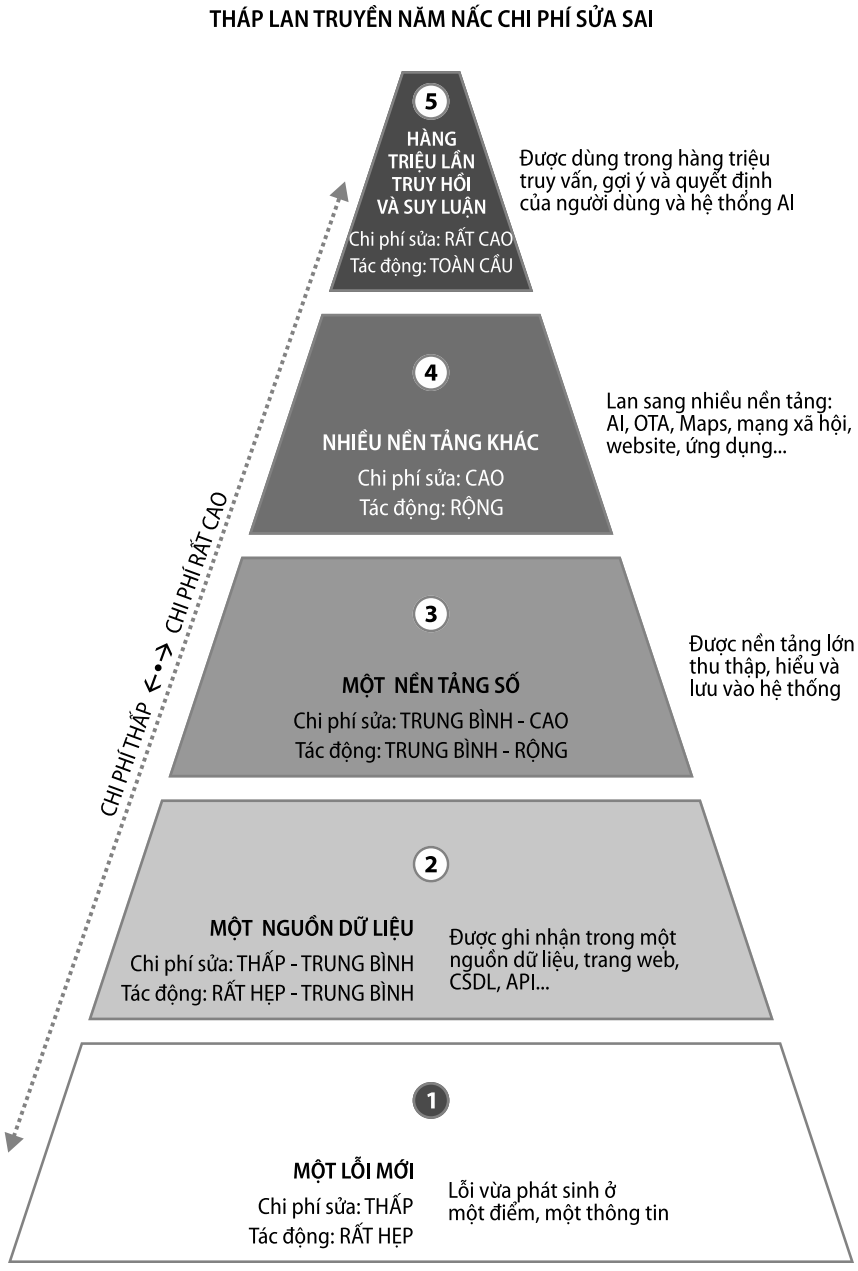
Một thông tin chỉ thực sự trở nên "nguy hiểm" khi nó được nhiều nguồn khác nhau lặp lại. Trong thế giới số, việc trích dẫn không chỉ có nghĩa là sao chép nguyên văn. Một bài báo có thể dựa trên bài báo trước, một trang web tổng hợp có thể lấy dữ liệu từ nhiều nguồn, một AI có thể tổng hợp từ hàng chục tài liệu, còn người dùng lại tiếp tục sử dụng câu trả lời của AI để tạo ra nội dung mới. Mỗi lần lặp lại như vậy đều làm tăng độ tin cậy biểu kiến của thông tin. Đây là cơ chế khiến chi phí sửa chữa tăng nhanh hơn nhiều so với mức tỷ lệ thuận với thời gian.

Ban đầu chỉ cần sửa một nguồn. Sau đó phải sửa nhiều nguồn. Tiếp theo là phải làm việc với nhiều nền tảng khác nhau. Cuối cùng, ngay cả khi tất cả các nguồn đã được cập nhật, các hệ thống AI vẫn có thể tiếp tục truy hồi những bản sao cũ hoặc các nội dung đã được lưu trữ trước đó.

Một sai lệch hiếm khi dừng lại ở nơi nó xuất hiện. Ban đầu, lỗi chỉ tồn tại trong một bản ghi và có thể được sửa ngay tại nguồn với chi phí rất thấp. Khi được một nền tảng thu thập, việc sửa không còn dừng ở dữ liệu gốc mà phải chờ quá trình đồng bộ và cập nhật của nền tảng đó. Nếu thông tin tiếp tục lan sang nhiều nền tảng khác, mỗi hệ thống lại có quy trình xác minh và cập nhật riêng, khiến chi phí sửa sai tăng lên đáng kể.

Khó hơn nữa là khi sai lệch đã trở thành tri thức dẫn xuất. Lúc này, chỉ sửa dữ liệu gốc không còn đủ, cần tạo ra một khối lượng bằng chứng mới đủ lớn để thay đổi cách các hệ thống suy luận và xếp hạng. Mức độ khó cao nhất xuất hiện khi thông tin đã đi vào ký ức thuật toán, chẳng hạn được sử dụng trong dữ liệu huấn luyện hoặc các cấu trúc tri thức tích lũy của hệ thống. Khi đó, việc sửa sai không thể diễn ra tức thời mà phụ thuộc vào các chu kỳ cập nhật, tái lập chỉ mục hoặc huấn luyện mới, đồng thời đòi hỏi phải duy trì nguồn bằng chứng chính xác trong thời gian dài. Hình 4.1 tóm tắt năm nấc chi phí sửa sai theo mức độ lan truyền của sai lệch.

Càng nhiều vòng lặp được hình thành, quán tính của ký ức thuật toán càng lớn. Đến một thời điểm nhất định, chi phí lớn nhất không còn là sửa dữ liệu mà là tạo ra đủ bằng chứng mới để lấn át toàn bộ mạng lưới bằng chứng cũ. Đó là lý do các tổ chức quản trị dữ liệu luôn coi việc ngăn chặn lan truyền quan trọng hơn xử lý hậu quả sau khi lan truyền đã xảy ra. Chi phí sửa một sai lệch tăng theo số hệ thống đã tiếp nhận nó, số nội dung phái sinh đã được tạo ra và độ sâu của ký ức đã hình thành. Vì vậy, sửa tại nguồn và sửa sớm thường rẻ hơn nhiều so với sửa sau khi sai lệch đã lan rộng.




**Hình 4.1.** Năm nấc chi phí sửa sai

### 4.7.3. Khi AI đã học từ dữ liệu cũ

Một trong những hiểu lầm phổ biến nhất là cho rằng chỉ cần sửa dữ liệu thì AI sẽ ngay lập tức trả lời đúng. Điều này không hoàn toàn chính xác. Xét riêng bài toán sửa sai, có thể tạm chia tri thức mà một hệ thống sử dụng thành hai tầng lớn. Cách chia này bỏ qua nhiều thành phần khác như bộ nhớ hội thoại, bộ nhớ tác nhân, chỉ mục tìm kiếm hay dữ liệu thời gian thực do công cụ trả về, nhưng đủ để thấy vì sao việc sửa không có một độ trễ chung. Tầng thứ nhất là dữ liệu được truy hồi theo thời gian thực từ các trang web, cơ sở dữ liệu hoặc kho tri thức. Tầng thứ hai là tri thức đã được mô hình học trong quá trình huấn luyện.

Nếu sai lệch chỉ tồn tại ở tầng truy hồi, việc cập nhật nguồn dữ liệu sẽ nhanh chóng cải thiện chất lượng câu trả lời. Nhưng nếu thông tin sai đã được hấp thụ trong quá trình huấn luyện của mô hình, việc sửa dữ liệu gốc không đồng nghĩa với việc mô hình ngay lập tức thay đổi những gì đã học. Đó là lý do nhiều người vẫn gặp các câu trả lời cũ dù trang web chính thức đã được cập nhật từ lâu. Mô hình có thể vẫn dựa vào những mẫu xác suất được hình thành từ dữ liệu trước đó, đặc biệt khi các nguồn mới chưa đủ nhiều hoặc chưa đủ mạnh để thay đổi cân bằng bằng chứng.

Điều này cho thấy sửa sai trong nền kinh tế tác nhân là một quá trình tích lũy, không phải một hành động tức thời. Cập nhật dữ liệu là điều kiện cần, nhưng chưa phải điều kiện đủ. Địa phương còn phải liên tục tạo ra các nguồn thông tin mới, gia tăng mức độ được trích dẫn và củng cố những bằng chứng đúng để dần thay đổi ký ức của các hệ thống AI.

Mỗi lớp lưu trữ của hệ sinh thái số đòi hỏi một cách sửa sai khác nhau. Với trang web và cổng dữ liệu chính thức, việc sửa

thường đơn giản nhất, chỉ cần cập nhật tại nguồn và chờ các hệ thống lập chỉ mục lại. Khi thông tin đã được các nền tảng bản đồ hoặc cơ sở dữ liệu bên thứ ba thu thập, cần gửi yêu cầu chỉnh sửa kèm theo bằng chứng và trải qua quy trình xác minh của từng nền tảng. Nếu nội dung sai đã bị sao chép và lan truyền rộng, việc sửa không còn dừng ở nguồn gốc mà phải đồng thời hiệu chỉnh nhiều nguồn phái sinh và tạo đủ bằng chứng mới để lấn át thông tin cũ. Khó nhất là khi tri thức đã được phản ánh trong các mô hình AI hoặc các cấu trúc tri thức tích lũy của hệ thống. Khi đó, không thể sửa trực tiếp mà phải duy trì nguồn dữ liệu chính xác trong thời gian dài và chờ các chu kỳ cập nhật, tái lập chỉ mục hoặc huấn luyện mới của từng hệ thống.

### 4.7.4. Chiến lược sửa sai của địa phương

Nếu thương hiệu được hiểu là một hệ thống ký ức, thì sửa sai phải trở thành một quy trình quản trị thường xuyên chứ không phải hoạt động xử lý khủng hoảng. Một chiến lược sửa sai hiệu quả cần vận hành như một chu trình liên tục gồm năm bước.

- ***Thứ nhất, theo dõi:*** Địa phương phải thường xuyên quan sát cách mình được mô tả trên công cụ tìm kiếm, nền tảng bản đồ, cơ sở tri thức, mạng xã hội và các hệ thống AI. Không thể sửa điều mình không biết đang tồn tại.
- ***Thứ hai, kiểm toán:*** Không phải mọi thông tin đều có giá trị như nhau. Cần xác định nguồn nào có thẩm quyền cao, thuộc tính nào đang sai, mức độ lan truyền ra sao và sai lệch nào ảnh hưởng trực tiếp đến khả năng được AI truy hồi và đề xuất.
- ***Thứ ba, đối chiếu:*** Mọi dữ liệu quan trọng cần được so sánh với nguồn chính thức để phát hiện sự khác biệt. Tên

gọi, địa chỉ, danh mục, tọa độ, lịch sử, thuộc tính và các mối quan hệ giữa các thực thể phải được kiểm tra định kỳ nhằm bảo đảm tính nhất quán.

- ***Thứ tư, cập nhật:*** Khi phát hiện sai lệch, cần sửa tại nguồn có thẩm quyền trước, sau đó mở rộng sang các nền tảng và cơ sở dữ liệu liên quan. Việc cập nhật phải đồng thời tạo ra các bằng chứng mới thay vì chỉ xóa bỏ thông tin cũ.
- ***Thứ năm, đồng bộ:*** Giá trị của một thông tin không chỉ nằm ở việc đúng, mà còn ở việc đúng ở mọi nơi. Trang web chính thức, hồ sơ thực thể, bản đồ số, dữ liệu mở, báo chí, doanh nghiệp và các nền tảng số cần phản ánh cùng một hệ tri thức. Chỉ khi sự nhất quán được duy trì trên toàn bộ hệ sinh thái, AI mới có đủ cơ sở để điều chỉnh cách suy luận của mình.

Quy trình sửa sai hiệu quả bắt đầu từ theo dõi. Địa phương cần quan sát thường xuyên cách mình đang được mô tả trên các công cụ tìm kiếm, bản đồ, OTA và hệ thống AI. Nếu chỉ biết có sai lệch khi người bên ngoài phản ánh, nghĩa là năng lực theo dõi đã thất bại. Bước tiếp theo là kiểm toán. Không phải mọi sai lệch đều quan trọng như nhau, cần xác định đâu là nguồn có thẩm quyền và thuộc tính nào ảnh hưởng trực tiếp đến quyết định của máy. Nếu mọi lỗi đều được xử lý với cùng mức ưu tiên, nguồn lực sẽ bị phân tán. Sau đó là đối chiếu. Các dữ liệu quan trọng phải được so sánh với nguồn chính thức để xác định đâu là bản đúng. Khi không còn biết bản gốc nằm ở đâu, việc sửa sai gần như mất điểm tựa. Tiếp theo là cập nhật. Nguyên tắc là sửa trước tại nguồn có thẩm quyền rồi mới lan sang các hệ thống khác. Nếu chỉ sửa ở kênh dễ chỉnh sửa nhất mà bỏ qua nguồn gốc, sai lệch sẽ nhanh chóng xuất hiện trở lại. Cuối cùng là đồng bộ. Mục tiêu không phải đúng ở một nơi, mà đúng ở mọi nơi mà máy có

thể đọc. Nếu mỗi trang web, nền tảng hay cơ sở dữ liệu vẫn hiển thị một phiên bản khác nhau, thì về bản chất, địa phương vẫn chưa được sửa đúng.

Trong nền kinh tế tác nhân, quản trị thương hiệu vì thế không còn là công việc của riêng bộ phận truyền thông. Nó trở thành một năng lực quản trị dữ liệu, quản trị tri thức và quản trị ký ức số. Địa phương nào phát hiện sai sớm, sửa nhanh và duy trì được sự nhất quán lâu dài sẽ có lợi thế lớn trong cuộc cạnh tranh để được các hệ thống AI tin cậy và lựa chọn. Năm giai đoạn trên mô tả một chu trình quản trị chạy liên tục, không gắn với bất kỳ sai lệch cụ thể nào.

Khi phát hiện một sai lệch cụ thể, việc sửa không nên diễn ra theo phản xạ mà theo một giao thức gồm bảy bước liên tiếp:

- ***Bước 1. Xác định sai lệch:*** Trước hết phải xác định chính xác thực thể nào và mệnh đề nào đang sai. Một nhận định chung như "hình ảnh địa phương đang xấu đi" không đủ để kiểm chứng hay sửa chữa.
- ***Bước 2. Truy nguồn:*** Lần theo nguồn xuất hiện sớm nhất còn có thể xác định được. Nếu không xử lý được nguồn phát, sai lệch sẽ tiếp tục lan truyền dù đã sửa ở nhiều nơi khác.
- ***Bước 3. Sửa tại nguồn có thẩm quyền:*** Luôn cập nhật trước ở nguồn dữ liệu có thẩm quyền cao nhất, vì đây là nơi nhiều hệ thống khác sử dụng để đối chiếu và xác minh.
- ***Bước 4. Tạo bằng chứng cập nhật:*** Công bố phiên bản đúng kèm thời gian cập nhật và bằng chứng xác thực. Máy chỉ nhận biết có thay đổi khi xuất hiện những tín hiệu cập nhật rõ ràng.

- ***Bước 5. Đồng bộ hệ sinh thái:*** Tiếp tục cập nhật các nền tảng trung gian theo thứ tự ảnh hưởng, thay vì chỉ ưu tiên những nơi dễ chỉnh sửa nhất.
- ***Bước 6. Kiểm chứng kết quả:*** Theo dõi câu trả lời của các công cụ tìm kiếm, bản đồ và hệ thống AI trước và sau khi sửa để xác nhận rằng sai lệch đã thực sự được khắc phục.
- ***Bước 7. Duy trì bằng chứng:*** Tiếp tục tạo và duy trì các bằng chứng mới trong một khoảng thời gian đủ dài để thay thế ký ức cũ của hệ thống. Đây là bước thường bị bỏ qua nhưng lại quyết định việc sửa sai có bền vững hay không.

Áp dụng giao thức bảy bước, thứ tự trở nên rõ ràng. Trước hết hợp nhất hai tên về một mã định danh và sửa tọa độ. Tiếp đó công bố trên cổng tỉnh một bản ghi có nhãn thời gian, nêu số hộ đang sản xuất và lịch mở cửa xưởng. Sau đó cập nhật nền tảng bản đồ kèm bằng chứng. Rồi để hợp tác xã, các hộ và đơn vị lữ hành cùng công bố dữ liệu riêng của họ, tạo ra những nguồn thực sự độc lập. Cuối cùng là theo dõi mô tả do máy tổng hợp theo từng quý. Không bước nào trong số này là một chiến dịch truyền thông.

Trước khi sang phần kết, một khoảng nghỉ. Giao thức bảy bước ở trên có thể chạy thử trên đúng một thực thể của địa phương mình và người đọc nên chạy thử một lần trước khi đồng ý với phần còn lại.

## 4.8. KẾT LUẬN CHƯƠNG

Chương này bắt đầu từ một câu hỏi tưởng như thuần túy kỹ thuật, làm thế nào để AI biết một địa phương là ai? Nhưng càng đi sâu, câu trả lời càng cho thấy đây không phải là vấn đề của công nghệ, mà là vấn đề của quản trị tri thức.

Một địa phương không trở nên khả đọc chỉ vì có nhiều dữ liệu. Địa phương trở nên khả đọc khi dữ liệu ấy có thể được phân giải thành đúng thực thể, được các nguồn có thẩm quyền cùng xác nhận, được tổ chức trong một hệ tri thức nhất quán và được duy trì đủ lâu để hình thành ký ức thuật toán. Từ định danh, đối chiếu, xếp hạng, tổng hợp đến ghi nhớ, mỗi bước đều góp phần quyết định địa phương sẽ được AI hiểu như thế nào và cuối cùng có được AI lựa chọn hay không.

Điều quan trọng nhất mà chương này muốn chỉ ra là niềm tin của máy không được xây dựng trong một lần truy hồi, mà trong hàng nghìn lần xác nhận. Mỗi lần một thực thể được nhận diện đúng, mỗi lần một nguồn được đối chiếu, mỗi lần một thuộc tính được cập nhật và mỗi lần một mô tả được AI tổng hợp đều để lại một dấu vết trong ký ức của hệ thống. Những dấu vết ấy tích lũy theo thời gian, tạo thành một lịch sử số mà các thuật toán sử dụng để suy luận về tương lai.

Chính vì vậy, bài toán của địa phương không chỉ là được nhìn thấy, mà còn là được nhìn thấy đúng, được mô tả nhất quán và được sửa sai đủ sớm. Một sai lệch nhỏ nếu bị bỏ mặc sẽ lan truyền qua đồ thị tri thức, qua các nền tảng số, qua các mô hình AI và dần trở thành một phần của ký ức mặc định. Khi hiện tượng khóa chặt đã hình thành, việc sửa sai không còn là sửa một bài viết, một trang web hay một cơ sở dữ liệu. Đó là quá trình thay đổi cả lịch sử tích lũy của niềm tin.

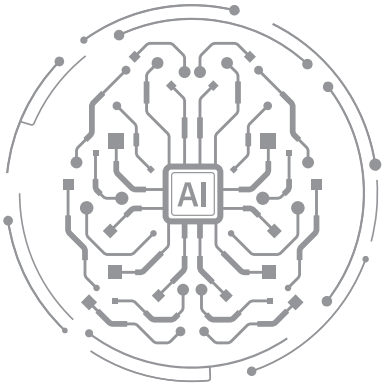
Nhìn từ góc độ này, quản trị thương hiệu địa phương trong nền kinh tế tác nhân đã bước sang một giai đoạn hoàn toàn mới. Trọng tâm không còn là sản xuất thật nhiều nội dung hay triển khai những chiến dịch truyền thông ngắn hạn, mà là xây dựng và bảo vệ một hạ tầng tri thức số đủ chính xác, đủ thống nhất và đủ bền vững để các hệ thống AI khác nhau có khả năng đọc, kiểm chứng và có xu hướng hội tụ về một mô tả đáng tin cậy. Thương hiệu vì thế không còn chỉ là hình ảnh trong tâm trí con người. Nó còn là trạng thái của ký ức thuật toán.

Nếu Chương 3 giải thích máy đọc thế giới bằng kiến trúc nào, thì Chương 4 đã trả lời câu hỏi niềm tin của máy được hình thành và khóa chặt ra sao. Chương 5 sẽ theo dõi điều xảy ra khi niềm tin ấy được đưa vào cơ chế xếp hạng. Bởi một mô tả không chỉ nằm lại trong ký ức của máy. Nó quyết định thực thể nào được đưa vào tập ứng viên, địa phương nào được đề xuất, dòng người, dòng tiền cùng sự chú ý sẽ được phân bổ về đâu.

## CHƯƠNG 5

# CUỘC CẠNH TRANH GIÀNH QUYỀN ĐƯỢC ĐỀ XUẤT

*Bộ lọc xếp hạng. Kinh tế học của sự đề xuất, tập ứng viên, hiệu ứng Matthew của thuật toán và nghịch lý máy càng thông minh thì đám đông càng dễ tập trung, trong điều kiện không có cơ chế đa dạng hóa hoặc điều tiết. Xem xét ba tầng can thiệp vào chuỗi quyết định.*


AI

Nếu Chương 4 ở trong hai bộ lọc đầu, chương này ở trọn trong bộ lọc thứ ba. Một địa phương đã được phân giải và đã được chứng thực vẫn có thể không bao giờ được nhắc tới, vì xếp hạng không thưởng cho việc tồn tại mà thưởng cho việc phù hợp. Vì bộ lọc thứ ba nằm bên trong cái vòng, mỗi lần được xếp trên lại tạo thêm căn cứ để lần sau được xếp trên nữa. Cơ chế tái cấu trúc mô tả ở Chương 2 trở lại ở đây dưới dạng cụ thể nhất của nó, khi thứ hạng trở thành thứ được phân bổ, các địa phương bắt đầu tự sửa mình theo tiêu chí xếp hạng chứ không theo nhu cầu của cư dân. Được nhìn thấy là điều kiện. Được đề xuất là phần thưởng. Hai thứ đó không tự chuyển hóa cho nhau.

Nếu Chương 4 trả lời câu hỏi máy hình thành niềm tin như thế nào, thì Chương 5 trả lời câu hỏi tiếp theo về niềm tin ấy được chuyển hóa thành lựa chọn ra sao. Đây là nơi bộ lọc thứ ba, xếp hạng (ranking), bắt đầu vận hành. Sau khi một thực thể đã được nhận diện, dữ liệu đã được đối chiếu và tri thức đã được hình thành, hệ thống vẫn còn phải giải quyết bài toán quan trọng nhất, trong hàng nghìn khả năng đều có thể chấp nhận được, nên đề xuất cái nào trước.

Trong thời đại tìm kiếm, xếp hạng quyết định thứ tự xuất hiện trên màn hình. Người dùng vẫn có quyền kéo xuống, mở thêm kết quả hoặc tự mình tiếp tục tìm kiếm. Nhưng trong nền kinh tế tác nhân, nơi con người ngày càng giao quyền lựa chọn cho AI, xếp hạng không còn chỉ là một thao tác kỹ thuật mà ngày càng trở thành một cơ chế phân bổ của nền kinh tế. Mỗi lần AI đề xuất một khách sạn, một điểm đến, một nhà hàng hay một hành trình, nó đồng thời loại bỏ vô số phương án khác khỏi tập lựa chọn. Không được đề xuất không còn chỉ có nghĩa là ít hiển thị hơn, mà có thể đồng nghĩa với việc hoàn toàn biến mất khỏi quá trình ra quyết định.

Một điểm ít được nhắc. Chúng được tối ưu để tìm phương án có xác suất thành công cao nhất đối với từng người dùng trong từng hoàn cảnh cụ thể. Thành công ở đây có thể là khả năng người dùng sẽ nhấp vào gợi ý, đặt phòng, hoàn thành chuyến đi, để lại đánh giá tích cực hoặc tiếp tục sử dụng dịch vụ trong tương lai. Nói cách khác, thuật toán không tối đa hóa chất lượng của địa phương mà tối đa hóa xác suất đạt được mục tiêu của hệ thống.

Đó là lý do vì sao nền kinh tế tác nhân không tự động phân bổ khách du lịch đồng đều hơn, dù AI hiểu người dùng tốt hơn trước. Nó có xu hướng khuếch đại những tín hiệu vốn đã mạnh, theo một vòng lặp. Hệ quả là cuộc cạnh tranh giữa các địa phương đổi tính chất. Họ không chỉ cạnh tranh để thu hút sự chú ý của khách du lịch mà còn cạnh tranh để trở thành lựa chọn mặc định của các hệ thống AI. Cùng lúc, các nền tảng cạnh tranh với nhau để trở thành nơi người dùng sẵn sàng giao quyền lựa chọn. Quyền lực dịch chuyển từ việc sở hữu tài nguyên du lịch sang khả năng định hình thứ tự ưu tiên trong quá trình đề xuất.

Chương được chia thành ba phần. Phần A hỏi quyền đề xuất được hình thành thế nào, từ giá trị kinh tế của một vị trí đến những tín hiệu mà địa phương và nền tảng dùng để cạnh tranh. Phần B hỏi vì sao lợi thế ban đầu tự khuếch đại qua thời gian. Phần C hỏi điều gì xảy ra ngoài thực địa và bộ công cụ điều tiết hiện có thiếu hụt ở đâu không. Một điểm cần nói trước để tránh ngộ nhận. Chương này không lập luận rằng các hệ thống đề xuất là nguyên nhân của quá tải du lịch. Quá tải là hiện tượng có trước AI, sinh ra từ hàng không giá rẻ, tàu biển du lịch, thu nhập tăng, chính sách thị thực, thị trường lưu trú ngắn hạn và giới hạn hạ tầng. Điều chương này lập luận là các hệ thống đề xuất đang trở thành bộ khuếch đại và bộ định tuyến của dòng chảy ấy, làm nó nhanh hơn, tập trung hơn và khó điều tiết hơn bằng những công cụ sẵn có.

> ▪ ***Quy ước thuật ngữ dùng trong chương:***
>
> 1. *Mô hình là hệ thống sinh ra hoặc chấm điểm nội dung.*
> 2. *Thuật toán đề xuất là cơ chế chọn và sắp xếp các phương án.*
> 3. *Tác nhân AI là hệ thống có khả năng lập kế hoạch và thực thi, không chỉ trả lời.*
> 4. *Nền tảng là chủ thể vận hành giao diện, dữ liệu và khả năng giao dịch.*
> 5. *Trợ lý AI là cách gọi cùng đối tượng ấy từ góc nhìn người dùng.*
> 6. *Tập ứng viên là nhóm phương án được đưa vào vòng cân nhắc. Tập lựa chọn là nhóm phương án cuối cùng được trình bày hoặc được thực thi. Nhóm K phương án đứng đầu, viết gọn là Top-K, là phần đầu của tập lựa chọn.*

▪ ***Về cách dùng thuật ngữ:***

• ***Cơ chế*** *là quan hệ giải thích, nó nói vì sao một hiện tượng xảy ra và được dùng khi lập luận nhắm vào nguyên nhân.*

• ***Mệnh đề*** *là phát biểu khái quát về xu hướng vận hành của hệ thống, quan sát được và có thể bị bác bỏ khi kiến trúc thay đổi, mỗi mệnh đề trung tâm trong sách đều kèm điều kiện bác bỏ.*

• ***Nguyên tắc*** *là lời khuyên chuẩn tắc dành cho người quản trị, không phải mô tả thực tại.*

• ***Giả thuyết*** *là mệnh đề tác giả cho là đúng nhưng chưa có bằng chứng trực tiếp trong chương. Chương này có bốn giả thuyết.*

• ***Khung phân tích*** *là cách sắp xếp khái niệm do sách này đề xuất, không phải mô tả cơ chế của mọi hệ thống. Chương này có một khung.*

Không mệnh đề nào trong các loại trên mang nghĩa của định luật khoa học tự nhiên. Sách không dùng chữ định luật, vì không phát biểu nào ở đây có địa vị của một định luật khoa học tự nhiên.

PHẦN A

# QUYỀN ĐỀ XUẤT ĐƯỢC HÌNH THÀNH THẾ NÀO

## 5.1. KINH TẾ HỌC CỦA SỰ ĐỀ XUẤT

Nếu trong nền kinh tế tìm kiếm, giá trị lớn nhất là được xuất hiện, thì trong nền kinh tế tác nhân, giá trị lớn nhất là được đề xuất. Hai khái niệm này tưởng như gần giống nhau nhưng khác nhau về bản chất. Xuất hiện chỉ mang lại cơ hội được lựa chọn. Còn được đề xuất đồng nghĩa với việc hệ thống đã thay người dùng thực hiện phần quan trọng nhất của quá trình ra quyết định. Khi AI không còn đưa ra một danh sách dài để con người tự đọc mà chỉ đề xuất một vài phương án hoặc tự động lựa chọn một phương án duy nhất, cạnh tranh cũng chuyển từ việc giành sự chú ý sang giành quyền được ưu tiên.

Đây là nơi bộ lọc thứ ba, xếp hạng, bắt đầu phát huy vai trò quyết định. Bộ lọc này không đánh giá đúng hay sai, nó đánh giá mức độ ưu tiên. Một thông tin hoàn toàn chính xác vẫn có thể không được đề xuất nếu nó không vượt qua được cơ chế xếp hạng.

### 5.1.1. Từ được nhìn thấy đến được đề xuất

Lịch sử của internet cũng có thể được nhìn như lịch sử của quyền lựa chọn. Trong mỗi giai đoạn, công nghệ đều cố gắng giảm bớt lượng công việc mà con người phải tự thực hiện khi ra quyết định. Sự thay đổi ấy có thể được mô tả qua ba cấp độ: tìm kiếm (search), đề xuất (recommendation) và tác nhân (agent). Đây là ba nấc tiến hóa của cùng một quá trình chứ không phải ba công nghệ tách biệt, máy ngày càng đảm nhận nhiều hơn vai trò lựa chọn thay con người.

Ở giai đoạn tìm kiếm, hệ thống chủ yếu đóng vai trò người thủ thư. Người dùng phải biết mình muốn gì, tự đặt câu hỏi, tự đọc kết quả và tự đưa ra quyết định cuối cùng. Công cụ tìm kiếm chỉ sắp xếp thế giới theo mức độ liên quan, còn trách nhiệm lựa chọn vẫn thuộc về con người. Vì vậy, mục tiêu của doanh nghiệp hay địa phương khi ấy là được nhìn thấy. Chỉ cần xuất hiện trong những vị trí đầu của kết quả tìm kiếm là đã có cơ hội được người dùng cân nhắc.

Giai đoạn tiếp theo là các hệ thống đề xuất. Thay vì chờ người dùng hỏi, nền tảng bắt đầu dự đoán nhu cầu và chủ động gợi ý. Netflix đề xuất bộ phim có thể bạn sẽ thích. Spotify tạo danh sách nhạc phù hợp với thói quen nghe. TikTok quyết định video tiếp theo xuất hiện trên màn hình. Các nền tảng du lịch đề xuất khách sạn, nhà hàng và điểm tham quan dựa trên lịch sử tìm kiếm, vị trí, ngân sách và hành vi của những người tương tự. Quyền lựa chọn chưa mất đi, nhưng phạm vi lựa chọn đã được thuật toán thu hẹp đáng kể. Mục tiêu cạnh tranh lúc này không còn là xuất hiện trong hàng nghìn kết quả, mà là lọt vào danh sách rất ngắn mà hệ thống quyết định giới thiệu.

Nền kinh tế tác nhân đưa quá trình ấy tiến thêm một bước. AI không chỉ đề xuất mà còn bắt đầu thực hiện công việc thay con người. Người dùng không còn hỏi "nên đi đâu?" mà giao cho tác nhân: "Hãy lên kế hoạch cho chuyến đi bốn ngày với ngân sách này." Hệ thống sẽ tự tìm kiếm, so sánh, xếp hạng, lựa chọn, đặt dịch vụ và trong nhiều trường hợp hoàn thành giao dịch mà không cần từng bước xác nhận. Khi đó, phần lớn các phương án cạnh tranh chưa bao giờ xuất hiện trước mắt người dùng. Cuộc cạnh tranh thực sự diễn ra bên trong quá trình suy luận của máy.

Mức độ ủy nhiệm ấy đang tăng nhanh, dù các con số đo lường nó thì mâu thuẫn nhau đến mức đáng ngờ. Khảo sát của Allianz

Partners trên 2.001 người trưởng thành ở Mỹ, thực hiện từ ngày 20/3 đến ngày 14/4/2026, cho thấy 37% du khách Mỹ dùng AI để lập kế hoạch chuyến đi.[1] Khảo sát thường niên của nền tảng Klook trên 11.000 người dùng toàn cầu lại đưa ra con số 91%.[2] Khoảng cách giữa 37% và 91% không phải sai sót thống kê mà là hệ quả của việc hỏi ai. Klook khảo sát chính người dùng của mình, vốn là nhóm thạo công nghệ hơn mặt bằng chung. Các khảo sát công bố trong giai đoạn 2024 tới 2026 cho thấy tỷ lệ người dùng AI để lập kế hoạch du lịch dao động rất rộng, từ dưới 10% đến trên 90%, tùy đối tượng khảo sát, phương pháp chọn mẫu và cách đặt câu hỏi. Vì vậy, điều quan trọng đối với nhà hoạch định chính sách không phải là tìm một con số "đúng", mà là hiểu bối cảnh của từng khảo sát trước khi sử dụng nó làm căn cứ ra quyết định.

Ba giai đoạn ấy phản ánh ba cấp độ của quyền lựa chọn. Tìm kiếm giúp con người nhìn thấy nhiều khả năng hơn. Đề xuất giúp con người lựa chọn nhanh hơn. Tác nhân giúp con người không cần trực tiếp lựa chọn nữa. Cần nói rõ rằng đây là ba mức độ chuyển giao quyền lựa chọn thường cùng tồn tại trong một hệ sinh thái, không phải ba thời đại tách biệt nối tiếp nhau. Một tác nhân vẫn gọi tìm kiếm và vẫn gọi hệ thống đề xuất bên trong quy trình của nó, còn người dùng thì chuyển qua lại giữa ba chế độ tùy mức hệ trọng của việc mình đang làm. Điều thay đổi theo thời gian không phải chế độ nào tồn tại, mà tỷ trọng của từng chế độ trong tổng số quyết định.

1. Allianz Partners, khảo sát về việc dùng trí tuệ nhân tạo lập kế hoạch chuyến đi, dẫn lại theo Hotel Dive. Đây là khảo sát do một doanh nghiệp bảo hiểm du lịch công bố và được trang tin ngành dẫn lại.

2. Klook, khảo sát thường niên trên khoảng 11.000 người dùng toàn cầu, dẫn lại theo CNBC. Klook khảo sát chính người dùng nền tảng của mình, tức một mẫu tự chọn thiên mạnh về nhóm thạo công nghệ, nên con số 91% không đại diện cho du khách nói chung.

Ba giai đoạn này phản ánh sự dịch chuyển dần của quyền lựa chọn và sự thu hẹp của tập phương án mà con người nhìn thấy. Ở giai đoạn tìm kiếm, máy chủ yếu đóng vai trò như một thủ thư, giúp truy xuất thông tin nhưng vẫn để người dùng tiếp cận một tập lựa chọn rất rộng. Sang giai đoạn đề xuất, máy trở thành người gợi ý, sử dụng dữ liệu và thuật toán để sàng lọc, khiến người dùng chỉ còn nhìn thấy một số ít phương án được ưu tiên. Đến giai đoạn tác nhân, máy không còn dừng ở việc gợi ý mà trực tiếp lựa chọn và thực hiện hành động thay người dùng. Khi đó, tập phương án nhìn thấy gần như thu hẹp xuống còn một lựa chọn duy nhất. Vì vậy, sự chuyển dịch từ tìm kiếm sang đề xuất rồi đến tác nhân không chỉ là thay đổi về công nghệ, mà còn là quá trình quyền lựa chọn từng bước được ủy nhiệm cho máy.

Mỗi bước chuyển giao thêm quyền lựa chọn từ con người sang máy đều làm thu hẹp tập phương án mà người dùng thực sự cân nhắc. Ở giai đoạn tìm kiếm, máy chủ yếu thực hiện việc truy hồi và sắp xếp thông tin. Người dùng vẫn phải tự đặt câu hỏi, tự so sánh các kết quả và tự đưa ra quyết định cuối cùng. Vì vậy, đơn vị cạnh tranh của một địa phương là khả năng xuất hiện trên một trang kết quả tìm kiếm. Sang giai đoạn đề xuất, máy bắt đầu dự đoán nhu cầu và chủ động sàng lọc các phương án phù hợp. Người dùng không còn xem toàn bộ tập lựa chọn mà chỉ cân nhắc trong danh sách ngắn đã được thuật toán ưu tiên. Khi đó, các địa phương không còn cạnh tranh để được tìm thấy, mà cạnh tranh để giành một vị trí trong danh sách được đề xuất. Đến giai đoạn tác nhân, máy không chỉ đề xuất mà còn lập kế hoạch và thực hiện hành động thay người dùng. Vai trò của con người chủ yếu còn lại là xác nhận hoặc phủ quyết quyết định của tác nhân. Lúc này, đơn vị cạnh tranh của địa phương không còn là một vị trí trên trang tìm kiếm hay trong danh sách gợi ý, mà là một trong rất ít suất xuất hiện trong câu trả lời hoặc kế hoạch hành động mà tác nhân lựa chọn.

### 5.1.2. Giá trị kinh tế của một vị trí

Trong hầu hết các thị trường số, sự khác biệt giữa vị trí thứ nhất và vị trí thứ mười không chỉ là chín bậc xếp hạng, mà có thể là nhiều lần về lượng người dùng, doanh thu và giá trị kinh tế. Điều đó khiến thứ hạng trở thành một loại tài sản vô hình.[1] Một vị trí càng cao trong danh sách đề xuất càng có khả năng thu hút sự chú ý, tạo ra giao dịch và tích lũy thêm dữ liệu cho những lần xếp hạng tiếp theo.

Thực tế này đã được quan sát từ thời kỳ công cụ tìm kiếm. Phần lớn người dùng chỉ nhấp vào vài kết quả đầu tiên và rất ít khi chuyển sang trang thứ hai. Vì vậy, doanh nghiệp cạnh tranh quyết liệt để lọt vào ba vị trí đầu, hoặc ít nhất là mười vị trí đầu. Không phải vì các vị trí thấp hơn hoàn toàn vô giá trị, mà vì xác suất được lựa chọn giảm rất nhanh khi thứ hạng đi xuống. Giá trị kinh tế của mỗi bậc xếp hạng không phân bố tuyến tính mà thường tập trung ở một số vị trí đầu.

Mức độ dốc của đường phân phối này đã được đo từ lâu. Phân tích của Sistrix trên hơn 80 triệu từ khóa cho thấy kết quả tự nhiên đứng đầu trên Google nhận tỷ lệ nhấp trung bình khoảng 28,5%, vị trí thứ hai khoảng 15%, vị trí thứ ba khoảng 11% và vị trí thứ mười chỉ còn khoảng 2,5%. Nói cách khác, chín bậc xếp hạng làm bốc hơi hơn chín phần mười lượng nhấp. Báo cáo của First Page Sage cho năm 2025 đo được ba kết quả tự nhiên đầu tiên gom khoảng 68,7% toàn bộ lượt nhấp trên một trang kết quả.[2]

1. First Page Sage, "Google Click-Through Rates (CTRs) by Ranking Position", bản cập nhật năm 2026, https://firstpagesage.com/reports/google-click-through-rates-ctrs-by-ranking-position/. Số liệu ngành tổng hợp từ nhiều nguồn, phương pháp không công khai đầy đủ, chỉ nên đọc theo bậc độ lớn. Con số này đo trang kết quả không có tóm tắt sinh ra nên không áp dụng cho truy vấn có tóm tắt.

2. First Page Sage: vị trí thứ nhất 39,8%, thứ hai 18,7%, thứ ba 10,2%. Đối chiếu với SISTRIX, "Why (almost) everything you knew about Google CTR is no longer valid", ngày 14 tháng 7 năm 2020, trên hơn 80 triệu từ khoá: 28,5%, 15,7% và 11%. Khoảng cách đến từ chênh lệch sáu năm và từ việc SISTRIX chỉ đo dữ liệu di động.

Trong nền kinh tế tác nhân, hiện tượng này còn trở nên mạnh hơn. AI thường không trình bày mười phương án để người dùng cân nhắc, mà chỉ đưa ra ba gợi ý, thậm chí chỉ một lựa chọn duy nhất. Điều đó có nghĩa là khoảng cách giữa việc đứng thứ ba và thứ tư có thể không còn là chênh lệch nhỏ về lưu lượng truy cập, mà là khoảng cách giữa được tham gia thị trường và hoàn toàn bị loại khỏi quá trình ra quyết định. Khi danh sách đề xuất càng ngắn, giá trị kinh tế của từng vị trí càng tăng.

Trong nền kinh tế tác nhân, độ dốc ấy còn dựng đứng hơn. Nghiên cứu của Pew Research Center theo dõi hành vi duyệt web thực tế của khoảng 900 người trưởng thành ở Mỹ trong tháng 3 năm 2025, trên 68.879 lượt tìm kiếm Google, cho thấy khi có bản tóm tắt do AI sinh ra ở đầu trang, người dùng chỉ nhấp vào một kết quả truyền thống trong 8% số lượt.[1] Khi không có bản tóm tắt ấy, tỷ lệ là 15%, tức gần gấp đôi. Đáng chú ý hơn, tỷ lệ người dùng nhấp vào chính các nguồn được trích dẫn bên trong bản tóm tắt chỉ là 1%. Con số 1% ấy là hình ảnh cô đọng nhất của điều chương này muốn nói. Một địa phương có thể được AI đọc, được AI trích dẫn và vẫn gần như không nhận được lượt truy cập nào.

Đây là logic của hiện tượng thường được gọi là người thắng giành phần lớn (Winner Takes Most). Cần giữ đúng tên gọi ấy. Người đứng đầu thường không giành toàn bộ thị trường. Nhu cầu vẫn phân mảnh giữa vài phương án. Điều xảy ra là phần lớn lợi ích dồn về một nhóm rất nhỏ, còn phần đuôi thì mỏng dần tới mức không đủ nuôi một hoạt động kinh doanh.

1. Pew Research Center, nghiên cứu công bố ngày 22 tháng 7 năm 2025 trên dữ liệu duyệt web tháng 3 năm 2025, đã dẫn đầy đủ ở chú thích 4. Mẫu 900 người trưởng thành tại Mỹ với 68.879 lượt tìm kiếm.

Thu hẹp. Mỗi lần quyền lựa chọn được chuyển thêm một phần từ người sang máy, tập phương án mà người dùng thực sự cân nhắc có xu hướng thu hẹp mạnh. Tìm kiếm trình bày một trang kết quả, hệ thống đề xuất trình bày một danh sách ngắn, tác nhân thường trình bày một phương án kèm lời đề nghị xác nhận. Hệ quả là giá trị kinh tế của thứ hạng không giảm đều theo từng bậc mà sụp quanh ngưỡng của danh sách.

Dữ liệu từ Sistrix, First Page Sage và Pew Research Center cho thấy hai thay đổi quan trọng. Thứ nhất, lưu lượng truy cập luôn tập trung rất mạnh vào các vị trí đầu của kết quả tìm kiếm, riêng ba vị trí đầu chiếm gần 70% tổng số lượt nhấp.[1] Thứ hai, khi Google hiển thị AI Overview, tỷ lệ người dùng nhấp vào các kết quả truyền thống giảm gần một nửa (từ khoảng 15% xuống còn 8%), trong khi chỉ khoảng 1% nhấp trực tiếp vào các nguồn được AI trích dẫn.[2] Điều này phản ánh sự chuyển dịch từ mô hình "đọc danh sách kết quả" sang mô hình "đọc câu trả lời do AI tổng hợp". Dữ liệu quan sát hành vi không tự chứng minh đầy đủ quan hệ nhân quả.

Đối với địa phương, điều này đổi thứ tự ưu tiên trong tư duy cạnh tranh. Có mặt trên bản đồ số và có một trang web đẹp là điều kiện tối thiểu, không phải mục tiêu. Mục tiêu là có đủ tín hiệu để được xếp vào nhóm nhỏ các lựa chọn đáng đề xuất. Ở nghĩa đó, một vị trí trong danh sách đề xuất là một nguồn lực kinh tế, có thể so sánh với hạ tầng hoặc với thương hiệu về mức độ ảnh hưởng tới dòng khách.

1. First Page Sage, "Google Click-Through Rates by Ranking Position", bản cập nhật năm 2026: ba kết quả tự nhiên đầu tiên nhận 68,7% tổng số lượt nhấp. SISTRIX, nghiên cứu trên hơn 80 triệu từ khoá: vị trí thứ nhất 28,5%, thứ hai 15,7%, thứ ba 11%. Số liệu ngành, phương pháp không công khai đầy đủ.

2. Pew Research Center, ngày 22 tháng 7 năm 2025: khi bản tóm tắt sinh ra xuất hiện, 8% lượt tìm kiếm dẫn tới một lần nhấp vào kết quả thường, so với 15% khi không có bản tóm tắt và chỉ 1% số lượt tìm kiếm có một lần nhấp vào nguồn được dẫn ngay trong bản tóm tắt.

Người dùng vẫn có thể hỏi lại. Nếu không hài lòng với phương án đầu tiên, họ chỉ cần yêu cầu tác nhân đưa ra lựa chọn khác. Vậy tập lựa chọn đâu có thu hẹp thật, nó chỉ được trình bày theo từng lượt. Phản bác này đúng về khả năng kỹ thuật, nhưng có lẽ đánh giá cao hơn thực tế mức độ người dùng sử dụng khả năng ấy. Việc hỏi lại đòi hỏi người dùng phải nghi ngờ phương án đang có, phải biết rằng còn phương án khác đáng xem xét và phải bỏ thêm công sức. Ba điều kiện ấy phụ thuộc vào loại người dùng, vào mức hệ trọng của quyết định và vào thiết kế giao diện, nên tần suất hỏi lại thay đổi rất nhiều giữa các tình huống. Điểm quan trọng hơn nằm ở chỗ khác. Phương án thứ hai mà tác nhân đưa ra khi bị hỏi lại thường lấy từ cùng một tập ứng viên đã được lọc trước đó. Hỏi lại làm thay đổi thứ tự trình bày nhiều hơn là mở rộng tập ứng viên. Với một địa phương chưa lọt vào tập ấy, số lượt hỏi lại của người dùng không giúp được gì.

### 5.1.3. Mọi đề xuất đều là một quyết định phân bổ

Mỗi lần AI đưa ra một đề xuất, điều xảy ra không chỉ là một câu trả lời xuất hiện trên màn hình. Đồng thời, một chuỗi nguồn lực của nền kinh tế cũng bắt đầu được phân bổ lại. Một nhà hàng được đề xuất sẽ nhận thêm lượt đặt bàn. Một khách sạn được ưu tiên sẽ có thêm đơn đặt phòng. Một địa phương xuất hiện trong kế hoạch du lịch của hàng triệu tác nhân sẽ nhận thêm dòng khách, doanh thu và dữ liệu. Vì vậy, mỗi đề xuất của AI thực chất là một quyết định phân bổ nguồn lực.

Nguồn lực đầu tiên được phân bổ là sự chú ý (attention). Trong nền kinh tế số, sự chú ý luôn khan hiếm. Người dùng chỉ có vài phút để quyết định một chuyến đi, một nhà hàng hay một khách sạn. Khi AI chỉ hiển thị ba phương án thay vì ba mươi phương án,

nó đang quyết định ai được quyền sử dụng quỹ thời gian hữu hạn ấy. Những lựa chọn không được đề xuất gần như không còn cơ hội được cân nhắc.

Sau sự chú ý là nhu cầu (demand). Một đề xuất không chỉ tạo ra lượt xem mà còn chuyển hóa thành nhu cầu thực tế. Hàng triệu quyết định nhỏ của AI sẽ cộng lại thành lượng khách đến một điểm đến, lượng phòng được đặt hay lượng vé được bán. Trong nền kinh tế tác nhân, nhu cầu không còn chỉ do quảng cáo hay truyền miệng tạo ra mà ngày càng được hình thành thông qua các hệ thống đề xuất.

Tiếp theo là dòng lưu chuyển (traffic). Dòng người, dòng dữ liệu và dòng giao dịch đều chảy theo những con đường mà thuật toán mở ra. Một thay đổi nhỏ trong cơ chế xếp hạng có thể làm lưu lượng truy cập của một địa phương tăng hoặc giảm mạnh chỉ trong thời gian ngắn. Điều này khiến thuật toán trở thành một loại hạ tầng vô hình, định hướng sự dịch chuyển của cả nền kinh tế.

Cuối cùng, mọi sự phân bổ ấy đều kết tinh thành vốn (capital). Nơi có nhiều khách hơn sẽ có doanh thu cao hơn. Doanh thu tạo ra lợi nhuận, đầu tư mới, giá trị bất động sản cao hơn và khả năng tiếp tục cải thiện dịch vụ. Trong khi đó, những địa phương liên tục bị loại khỏi danh sách đề xuất sẽ khó tích lũy vốn để thay đổi vị thế của mình. Vì thế, AI không trực tiếp tạo ra của cải, nhưng nó ngày càng quyết định của cải sẽ được tạo ra ở đâu. Bộ lọc xếp hạng vì vậy không chỉ là cơ chế lựa chọn thông tin mà còn là cơ chế phân phối nguồn lực của nền kinh tế số.

Phân bổ kèm theo. Không có đề xuất nào chỉ là thông tin. Mỗi đề xuất khởi động một chuỗi chuyển hóa có điều kiện gồm bốn khâu: sự chú ý, nhu cầu, dòng lưu chuyển và vốn. Chuỗi này

không tự động chạy hết. Mỗi khâu chỉ chuyển sang khâu sau nếu một điều kiện được thỏa mãn. Tỷ lệ rơi ở từng khâu là nơi phần lớn giá trị bị mất. Vì bốn khâu nối tiếp nhau, can thiệp vào khâu đầu thường rẻ hơn can thiệp vào khâu cuối.

Bốn tầng giá trị không nối với nhau một cách tự động. Mỗi tầng chỉ chuyển sang tầng tiếp theo khi đáp ứng những điều kiện riêng, đồng thời có độ trễ và mức độ khó đảo ngược khác nhau. Sự chú ý có thể thay đổi gần như tức thời, nhưng để chuyển thành nhu cầu, dòng lưu chuyển và cuối cùng là vốn tích lũy cho địa phương cần vượt qua nhiều rào cản về niềm tin, năng lực cung ứng và khả năng giữ lại giá trị. Càng đi xuống các tầng sâu hơn, thời gian tích lũy càng dài và chi phí sửa sai càng lớn.

### 5.1.4. Khi AI trở thành nhà môi giới lớn nhất

Trong lịch sử kinh tế, những chủ thể có quyền kết nối cung và cầu luôn nắm giữ vị trí đặc biệt. Chợ truyền thống kết nối người bán với người mua. Đại lý du lịch kết nối du khách với khách sạn. Các nền tảng số kết nối người dùng với hàng triệu dịch vụ. Nền kinh tế tác nhân đang tạo ra một lớp trung gian mới: AI trở thành nhà môi giới của các quyết định. Điểm đặc biệt là AI hầu như không sở hữu những tài sản mà nó phân phối. Nó không xây khách sạn, không vận hành nhà hàng, không quản lý bảo tàng và cũng không tổ chức lễ hội. Tuy nhiên, AI ngày càng quyết định khách sẽ đến khách sạn nào, ăn ở nhà hàng nào và ghé thăm điểm đến nào. Quyền lực kinh tế vì thế bắt đầu tách khỏi quyền sở hữu tài sản vật chất.

Trong nền kinh tế trải nghiệm, địa phương cạnh tranh chủ yếu bằng việc tạo ra sản phẩm và quảng bá sản phẩm ấy tới

khách du lịch. Nhưng trong nền kinh tế tác nhân, giữa sản phẩm và khách hàng xuất hiện thêm một lớp trung gian có khả năng sàng lọc và lựa chọn. Địa phương không còn giao tiếp trực tiếp với người dùng mà trước hết phải thuyết phục được AI rằng mình là một phương án đáng được ưu tiên. Cuộc cạnh tranh vì thế dịch chuyển từ thị trường tiêu dùng sang thị trường của các hệ thống đề xuất.

Điều này tạo ra một dạng quyền lực mới. Một tác nhân AI không cần sở hữu bất kỳ phòng khách sạn nào nhưng vẫn có thể ảnh hưởng đến tỷ lệ lấp đầy của hàng triệu phòng trên toàn thế giới. Nó không cần vận hành hãng hàng không nào nhưng có thể tác động đến luồng hành khách giữa các thành phố. Cũng giống như các nhà môi giới tài chính không trực tiếp sản xuất hàng hóa nhưng có thể quyết định dòng vốn chảy về đâu, AI trở thành nhà môi giới của dòng khách, dòng dữ liệu và dòng doanh thu. Đối với địa phương, điều đó có nghĩa là cạnh tranh trong tương lai không chỉ là xây dựng tài sản du lịch tốt hơn mà còn là xây dựng khả năng được các hệ thống AI tin cậy và ưu tiên. Trong nhiều tác vụ, bên kiểm soát cơ chế đề xuất có ảnh hưởng tới dòng nhu cầu lớn hơn bên trực tiếp sở hữu tài sản được đề xuất.

Môi giới không tài sản. Bên phân phối nhu cầu không cần sở hữu tài sản tạo ra nhu cầu ấy. Khi quyền định tuyến tách khỏi quyền sở hữu, cấu trúc rủi ro của hai bên khác nhau về loại. Bên sở hữu tài sản chịu rủi ro của vốn cố định và của địa điểm. Bên định tuyến chịu rủi ro danh tiếng, rủi ro pháp lý và rủi ro mất người dùng, nhưng ít bị ràng buộc bởi tài sản cố định, nên có khả năng dịch chuyển và thương lượng lớn hơn. Người xây khách sạn chịu rủi ro của tài sản cố định. Người xếp hạng khách sạn chịu ít rủi ro vật chất hơn, nhưng lại có thể quyết định tài sản ấy được tiếp cận nhu cầu đến mức nào.

### 5.1.5. Kinh tế học của xác suất

Một trong những hiểu lầm phổ biến nhất về AI là cho rằng nó luôn cố gắng tìm ra câu trả lời đúng nhất. Trên thực tế, phần lớn các hệ thống hiện đại được thiết kế để tìm câu trả lời có xác suất thành công cao nhất theo mục tiêu mà chúng được giao. Hai mục tiêu này thường gần nhau nhưng không hoàn toàn trùng khớp.

Đối với một hệ thống lập kế hoạch du lịch, thành công có thể là người dùng chấp nhận lịch trình được đề xuất, hoàn thành việc đặt dịch vụ và hài lòng sau chuyến đi. Đối với một nền tảng thương mại điện tử, thành công có thể là khả năng người dùng nhấp vào sản phẩm và hoàn tất thanh toán. Đối với một hệ thống đề xuất nội dung, thành công có thể là thời gian người dùng tiếp tục tương tác. Trong phần lớn các hệ thống hiện đại, cơ chế lựa chọn được thiết kế để tăng xác suất đạt một hoặc nhiều mục tiêu đã định, chứ không tìm kiếm một chân lý tuyệt đối.

Chính vì vậy, AI thường ưu tiên những phương án đã được kiểm chứng nhiều lần hơn là những phương án có thể tốt hơn nhưng thiếu dữ liệu. Một khách sạn với hàng chục nghìn đánh giá đáng tin cậy thường được ưu tiên hơn một khách sạn mới mở dù chất lượng thực tế có thể tương đương hoặc thậm chí cao hơn. Một điểm đến nổi tiếng với lượng dữ liệu phong phú cũng có nhiều khả năng được đề xuất hơn một địa phương ít được biết đến nhưng sở hữu trải nghiệm độc đáo. Thuật toán không nhất thiết thiên vị cái cũ. Nó thiên vị điều có xác suất thất bại thấp hơn.

Cách phát biểu chặt hơn là như sau. Hệ thống không trực tiếp hỏi địa phương nào tốt nhất. Nó ước tính phương án nào có xác suất đạt mục tiêu cao nhất với lượng thông tin hiện có. Chất

lượng không bị loại khỏi phép tính, nó là một biến trong phép tính, một biến chỉ đo được qua đại diện. Khi dữ liệu về một nơi còn mỏng, phần đóng góp của biến chất lượng vào kết quả cũng mỏng theo. Những biến dễ đo hơn sẽ chiếm chỗ.

Tối ưu theo mục tiêu quan sát được. Hệ thống xếp hạng không tối đa hóa chất lượng của phương án mà tối đa hóa xác suất đạt mục tiêu của chính hệ thống, đo bằng những đại lượng nó quan sát được. Chất lượng vẫn là một biến trong phép tính, nhưng chỉ ở mức mà nó có đại diện trong dữ liệu. Giả thuyết này cần được kiểm chứng riêng cho từng loại hệ thống, vì hàm mục tiêu của một công cụ tìm kiếm và của một nền tảng đặt chỗ không giống nhau.

> ▪ ***Về cách đọc một thứ hạng:***
> • *Không được coi thứ hạng của một địa phương trên bất kỳ hệ thống nào là thước đo chất lượng của địa phương ấy. Thứ hạng đo mức độ chắc chắn của máy, không đo giá trị của nơi chốn.*
> • *Không được đặt chỉ tiêu quản lý cho một cơ quan công dựa trên thứ hạng của một nền tảng thương mại mà cơ quan ấy không kiểm soát phương pháp tính.*
> • *Không được giải thích một cú tụt hạng là suy giảm chất lượng trước khi loại trừ khả năng dữ liệu về địa phương đã cũ hoặc đã đứt kết nối.*

## 5.2. ĐỊA PHƯƠNG CẠNH TRANH TRONG TẬP ỨNG VIÊN

Trong nền kinh tế trải nghiệm, các địa phương chủ yếu cạnh tranh để thu hút sự chú ý của con người. Họ đầu tư vào thương hiệu, truyền thông, sự kiện, hạ tầng và các sản phẩm du lịch với hy vọng tạo ra một hình ảnh đủ hấp dẫn để du khách lựa chọn. Mỗi địa phương có thể kể câu chuyện của riêng mình và chờ người dùng tự khám phá. Nhưng khi AI trở thành chủ thể đọc và đề xuất đầu tiên, cuộc cạnh tranh ấy chuyển sang một không gian khác hẳn, không gian xếp hạng của thuật toán.

Đây là một sự thay đổi có ý nghĩa căn bản. Địa phương không còn cạnh tranh trước hết trong tâm trí của du khách mà cạnh tranh trong quá trình suy luận của máy. Trước khi một con người nhìn thấy ba phương án du lịch trên màn hình, AI có thể đã loại bỏ hàng nghìn địa phương khác khỏi tập lựa chọn. Cuộc cạnh tranh thực sự vì thế diễn ra ở giai đoạn mà phần lớn con người rất khó quan sát được.

Điều đó khiến mỗi quyết định đề xuất mang tính tương đối. Một địa phương được lựa chọn không chỉ vì bản thân nó đáp ứng tiêu chí của người dùng, mà còn vì nó được đánh giá cao hơn những địa phương còn lại trong cùng thời điểm và cùng ngữ cảnh. Với một danh sách có độ dài cố định, mỗi vị trí trao cho một địa phương là một vị trí không còn cho những địa phương khác. Danh sách càng ngắn, tính loại trừ của cơ chế càng mạnh.

Từ góc nhìn đó, sự cạnh tranh giữa các địa phương không chỉ là cuộc đua về tài nguyên, cảnh quan hay ngân sách quảng bá. Đó là cuộc cạnh tranh về tín hiệu mà thuật toán có thể đọc, tin tưởng và ưu tiên. Những địa phương có dữ liệu đầy đủ hơn, thẩm quyền cao hơn, khả năng giao dịch tốt hơn và lịch sử được lựa

chọn dài hơn sẽ có lợi thế ngày càng lớn trong cơ chế xếp hạng. Ngược lại, những địa phương thiếu dữ liệu hoặc chưa từng được đề xuất nhiều sẽ phải đối mặt với một rào cản vô hình: rất khó được lựa chọn nếu chưa từng được lựa chọn trước đó.

### 5.2.1. Một nhóm nhỏ được thấy, phần lớn bị loại trước khi xuất hiện

Một trong những thay đổi căn bản nhất của nền kinh tế tác nhân là phần lớn địa phương không có cơ hội được con người nhìn thấy. Trước khi một hành trình được đề xuất, AI đã phải thu hẹp hàng chục nghìn khả năng xuống chỉ còn một nhóm rất nhỏ các phương án có thể chấp nhận được. Trong khoa học máy tính, tập hợp này thường được gọi là tập ứng viên (candidate set). Chỉ những thực thể vượt qua các bộ lọc ban đầu mới được đưa vào danh sách để tiếp tục xếp hạng và lựa chọn.

Ngay cả khi đã nằm trong tập ứng viên, cơ hội cũng rất hạn chế. Hầu hết các hệ thống AI cuối cùng chỉ trả về một số lượng rất nhỏ kết quả, thường được gọi là Top-K, tức chỉ giữ lại K phương án tốt nhất. Nếu K=3, nghĩa là chỉ ba địa phương xuất hiện trong câu trả lời cuối cùng dù phía sau có thể đã có hàng nghìn địa phương đủ điều kiện. Từ góc nhìn của người dùng, dường như chỉ tồn tại ba lựa chọn. Từ góc nhìn của thuật toán, hàng nghìn lựa chọn còn lại đã bị loại bỏ.

Đây là sự khác biệt quan trọng giữa thời đại tìm kiếm và thời đại tác nhân. Trước đây, doanh nghiệp hoặc địa phương đứng ở vị trí thứ năm mươi trên kết quả tìm kiếm vẫn có khả năng được một số người dùng nhìn thấy. Trong nền kinh tế tác nhân, nếu AI chỉ đề xuất ba phương án thì vị trí thứ tư gần như không còn ý nghĩa

thực tế. Khoảng cách giữa người được chọn cuối cùng và người bị loại đầu tiên không chỉ là một bậc xếp hạng, nó là khoảng cách giữa hiện diện và vô hình. Điều này đổi hình dạng của cuộc cạnh tranh. Thị trường có xu hướng dồn phần lớn nhu cầu vào một nhóm nhỏ các địa phương được lựa chọn, thay vì phân phối cho tất cả người tham gia theo những tỷ lệ khác nhau. Với phần lớn địa phương, câu hỏi không còn là giành thêm một phần thị phần mà là có tồn tại trong tập lựa chọn của máy hay không.

### 5.2.2. Không còn cạnh tranh với hàng xóm

Trong phần lớn lịch sử của du lịch, địa phương thường cạnh tranh với những nơi ở gần mình. Một tỉnh ven biển cạnh tranh với các tỉnh ven biển khác. Một thành phố di sản cạnh tranh với các thành phố di sản trong cùng quốc gia. Khoảng cách địa lý, chi phí di chuyển và khả năng tiếp cận khiến phạm vi cạnh tranh tương đối giới hạn. Internet đã mở rộng phạm vi ấy, nhưng AI còn đi xa hơn. Khi người dùng giao cho tác nhân một yêu cầu như hãy lên kế hoạch cho kỳ nghỉ bốn ngày phù hợp với ngân sách này, hệ thống không bị giới hạn bởi ranh giới hành chính hay địa lý. Nó có thể đồng thời cân nhắc Kyoto, Hội An, Luang Prabang, Chiang Mai hay Seville nếu chúng cùng đáp ứng mục tiêu của người dùng. Đối thủ của một địa phương không đơn thuần là địa phương bên cạnh mà là những nơi có khả năng thay thế nó trong cùng một ngữ cảnh sử dụng, ở bất kỳ đâu.

Điều này làm thay đổi đáng kể khái niệm thị trường. Một lễ hội truyền thống của một địa phương không chỉ cạnh tranh với lễ hội của địa phương lân cận mà còn cạnh tranh với hàng trăm sự kiện văn hóa trên toàn cầu. Một khu nghỉ dưỡng ven biển không chỉ cạnh tranh với những bãi biển trong nước mà còn cạnh tranh với Bali, Maldives hay Hawaii nếu người dùng có đủ ngân sách

và thời gian. AI không quan tâm đến biên giới quốc gia. Nó quan tâm đến xác suất thỏa mãn nhu cầu của người dùng.

Số liệu thống kê của ngành du lịch Việt Nam cho thấy mức tập trung này rõ. Năm 2025, cả nước đón gần 21,2 triệu lượt khách quốc tế, mức cao nhất từ trước đến nay, tăng 20,4% so với năm 2024 và cao hơn 17,8% so với năm 2019. Trong đó Thành phố Hồ Chí Minh ghi nhận khoảng 8,5 triệu lượt khách quốc tế và Hà Nội khoảng 7,8 triệu lượt. Cần đọc hai con số ấy một cách thận trọng. Thống kê theo địa phương có tính trùng, vì một du khách đi qua nhiều nơi sẽ được đếm nhiều lần, nên chúng không cộng thẳng vào tổng quốc gia được. Số liệu hiện có cũng không cho phép xác định chính xác tỷ lệ hành trình đi qua cả hai. Điều nói được là hai đô thị này chiếm vị trí nổi bật trong thống kê địa phương và trong nhiều hành trình quốc tế. Trong khi đó, phần lớn đơn vị hành chính cấp tỉnh còn lại ghi nhận quy mô khách quốc tế thấp hơn đáng kể. Từ ngày 1 tháng 7 năm 2025, cả nước có 34 đơn vị hành chính cấp tỉnh theo Nghị quyết 202/2025/QH15, gồm 28 tỉnh và 6 thành phố trực thuộc trung ương.[1] Mức tập trung ấy không phải kết quả của một quyết định quy hoạch nào. Nó có thể được giải thích một phần bởi việc hai thành phố có sân bay quốc tế lớn, có dữ liệu dày và có lịch sử xác nhận dài hơn phần còn lại.

Vì vậy, lợi thế địa lý dần nhường chỗ cho lợi thế thuật toán. Địa phương không còn được bảo vệ bởi khoảng cách hay sự thiếu thông tin của du khách. Chỉ cần hệ thống đánh giá rằng một điểm đến khác phù hợp hơn, một phần đáng kể dòng khách có thể được điều hướng sang nơi đó. Cuộc cạnh tranh giữa các địa phương trở thành một thị trường toàn cầu theo thời gian thực, nơi phần lớn điểm đến đều có thể trở thành đối thủ của nhau chỉ sau một lần truy vấn.

1. Nghị quyết 202/2025/QH15, Quốc hội khóa XV, thông qua ngày 12 tháng 6 năm 2025.

Đối với các địa phương nhỏ, đây vừa là cơ hội vừa là thách thức. Họ có thể tiếp cận người dùng ở bất kỳ đâu trên thế giới mà không cần mạng lưới đại lý truyền thống. Nhưng đồng thời, họ cũng phải đối mặt với sự cạnh tranh từ những thương hiệu mạnh nhất hành tinh. Trong nền kinh tế tác nhân, quy mô của đối thủ không còn được xác định bởi khoảng cách địa lý mà bởi sự tương đồng trong khả năng đáp ứng nhu cầu. Một bãi biển miền Trung giờ cạnh tranh với Đà Nẵng ít hơn là với Bali, vì cỗ máy so sánh không đọc bản đồ mà đọc ý định.

Thay thế theo ngữ cảnh. Khi tiêu chí lựa chọn là mức tương hợp với ý định người dùng chứ không phải khoảng cách địa lý, tập đối thủ của một địa phương được xác định lại theo khả năng thay thế trong cùng ngữ cảnh. Khoảng cách giảm vai trò hàng rào bảo vệ, đồng thời giảm vai trò trần giới hạn. Giả thuyết này mạnh hơn với những tác vụ mở như chọn điểm đến, yếu hơn với những tác vụ bị ràng buộc bởi thời gian di chuyển.

### 5.2.3. Địa phương cạnh tranh bằng tín hiệu

Nếu trong nền kinh tế trải nghiệm, địa phương cạnh tranh bằng câu chuyện, hình ảnh và cảm xúc, thì trong nền kinh tế tác nhân, họ còn phải cạnh tranh bằng những tín hiệu mà thuật toán có thể đọc và định lượng. AI không trực tiếp trải nghiệm một điểm đến. Nó chỉ có thể quan sát những dấu vết số mà địa phương để lại trên internet và trong các hệ thống dữ liệu. Chính những tín hiệu ấy trở thành cơ sở để thuật toán đánh giá và xếp hạng.

Tín hiệu đầu tiên là thẩm quyền (authority). Một thông tin xuất hiện trên trang web chính thức của cơ quan quản lý, được nhiều nguồn độc lập xác nhận và duy trì ổn định trong thời gian dài sẽ có trọng lượng lớn hơn một bài đăng ngắn trên mạng xã

hội. Thẩm quyền không chỉ đến từ một nguồn mà được tích lũy từ sự đồng thuận của nhiều nguồn đáng tin cậy. Địa phương càng xây dựng được hệ sinh thái dữ liệu nhất quán, tín hiệu thẩm quyền càng mạnh.

Tín hiệu thứ hai là mức độ phổ biến (popularity). Lượng đánh giá, số lượt ghé thăm, số bài viết, số hình ảnh, mức độ tương tác và nhiều chỉ báo hành vi khác giúp AI ước lượng xác suất một địa phương sẽ làm hài lòng người dùng. Phổ biến không đồng nghĩa với chất lượng, nhưng nó là một bằng chứng thống kê quan trọng. Những điểm đến càng được nhiều người lựa chọn càng có xu hướng tiếp tục được lựa chọn vì thuật toán coi đó là tín hiệu của sự thành công đã được kiểm chứng.

Tín hiệu thứ ba là tính cập nhật (freshness). Một địa phương có dữ liệu được cập nhật thường xuyên về giờ mở cửa, giá vé, sự kiện, tình trạng dịch vụ hay các thay đổi trong hạ tầng sẽ được đánh giá đáng tin cậy hơn một địa phương có thông tin cũ hoặc mâu thuẫn. Trong môi trường mà AI luôn phải giảm thiểu rủi ro, dữ liệu mới là một yếu tố quan trọng để duy trì thứ hạng.

Cuối cùng là khả năng sẵn sàng phục vụ (availability). Một điểm đến có thể rất hấp dẫn, nhưng nếu khách sạn đã hết phòng, vé tham quan không còn, không có khả năng đặt chỗ trực tuyến hoặc không thể xác nhận giao dịch theo thời gian thực, AI sẽ có xu hướng ưu tiên một phương án khác. Trong nền kinh tế tác nhân, khả năng được lựa chọn không chỉ phụ thuộc vào chất lượng của điểm đến mà còn phụ thuộc vào khả năng đáp ứng ngay lập tức nhu cầu của hệ thống.

Bốn nhóm tín hiệu này tạo nên ngôn ngữ cạnh tranh mới giữa các địa phương. Chúng không thay thế bản sắc, trải nghiệm hay

tài sản văn hóa, nhưng quyết định liệu những giá trị ấy có được AI nhìn thấy và đưa vào quá trình đề xuất hay không. Trong kỷ nguyên tác nhân, địa phương không chỉ phải tốt hơn, mà còn phải phát ra những tín hiệu đủ rõ, đủ mạnh và đủ nhất quán để thuật toán nhận ra rằng mình tốt hơn.

Có thể hình dung máy đọc một địa phương qua bốn nhóm tín hiệu bổ trợ lẫn nhau: thẩm quyền, mức độ phổ biến, tính cập nhật và khả năng sẵn sàng phục vụ. Bốn tín hiệu mà AI sử dụng để đánh giá một địa phương không chỉ khác nhau về bản chất mà còn khác nhau về nguồn dữ liệu, điểm can thiệp và tốc độ có thể cải thiện. Thẩm quyền được hình thành từ các nguồn chính thức, dữ liệu có định danh và xuất xứ rõ ràng, địa phương có thể tác động thông qua cổng dữ liệu công, lược đồ đánh dấu và hồ sơ thực thể, với thời gian cải thiện thường chỉ vài tháng. Mức độ phổ biến lại được phản ánh qua số lượng đánh giá, lượt nhắc đến và hình ảnh do cộng đồng tạo ra, đây là tín hiệu chỉ có thể cải thiện gián tiếp bằng việc nâng cao chất lượng trải nghiệm thực, nên thường cần nhiều năm mới tạo ra khác biệt đáng kể. Tính cập nhật phụ thuộc vào những thông tin thay đổi thường xuyên như giờ mở cửa, giá, lịch sự kiện hay tình trạng phục vụ, chỉ cần thiết lập quy trình cập nhật định kỳ và phân công rõ đầu mối chịu trách nhiệm thì có thể cải thiện trong vài tuần. Trong khi đó, khả năng sẵn sàng phản ánh việc còn phòng, còn vé, còn suất hay khả năng đặt chỗ theo thời gian thực, để nâng cao tín hiệu này, địa phương cần kết nối các hệ thống đặt chỗ thông qua giao diện lập trình và thường mất vài tháng để hoàn thiện. Trong những tác vụ đòi hỏi thực thi, một tín hiệu quá yếu, đặc biệt là khả năng sẵn sàng, có thể trở thành điểm nghẽn khiến ba tín hiệu còn lại không chuyển hóa thành giao dịch. Sự khác biệt về thời gian cải thiện cho thấy một điều quan trọng trong hoạch định chính sách, ba trong bốn tín hiệu có thể được nâng lên ngay

trong một chu kỳ ngân sách nếu có quy trình và hạ tầng phù hợp, còn mức độ phổ biến là tài sản chỉ có thể tích lũy dần qua thời gian. Vì vậy, chiến lược đúng không phải là chờ địa phương trở nên nổi tiếng rồi mới đầu tư vào khả đọc, mà là ưu tiên cải thiện trước những tín hiệu có thể quản trị được để tạo nền tảng cho danh tiếng bền vững về sau.

Một số người cho rằng tín hiệu là thứ mua được. Một địa phương có ngân sách lớn có thể thuê dịch vụ tạo đánh giá, tạo bài viết và tạo lượt nhắc đến. Nếu vậy thì bốn nhóm tín hiệu chỉ đo khả năng chi tiêu chứ không đo điều gì có thật. Phản bác này đúng một phần và cần được nói thẳng. Tín hiệu phổ biến là nhóm dễ tác động bằng tiền nhất. Thị trường mua bán đánh giá tồn tại thật. Tuy vậy, việc mua một nhóm tín hiệu không đồng nghĩa với việc tạo được một hồ sơ đáng tin cậy trên toàn bộ hệ thống. Tín hiệu phổ biến có thể bị thao túng tương đối dễ. Tính cập nhật đòi hỏi một quy trình vận hành liên tục, tức một cam kết tổ chức chứ không phải một khoản chi. Khả năng sẵn sàng đòi hỏi năng lực phục vụ có thể kiểm chứng qua giao dịch, nên tồn kho ảo hay dữ liệu không đồng bộ thường bộc lộ ngay ở lần đặt chỗ thất bại đầu tiên. Còn thẩm quyền đòi hỏi sự xác nhận từ những nguồn có tư cách và có lịch sử ổn định, thứ mà tiền tác động được nhưng chỉ một cách gián tiếp và chậm. Một hồ sơ có mức độ phổ biến cao nhưng dữ liệu giao dịch, dữ liệu cập nhật và nguồn xác nhận đều yếu sẽ tạo ra một sự mất cân đối mà các hệ thống có thể nhận diện.

### 5.2.4. Danh tiếng trở thành dữ liệu

Danh tiếng từng là một khái niệm mềm. Nó tồn tại trong ký ức của con người, trong lời truyền miệng, trong các bài báo, trong

câu chuyện của du khách và trong hình ảnh mà một địa phương tích lũy qua nhiều năm. Danh tiếng có thể rất mạnh nhưng khó đo lường, khó chuẩn hóa và khó đưa vào một phép tính cụ thể. Trong nền kinh tế tác nhân, điều đó bắt đầu thay đổi. Danh tiếng ngày càng được chuyển hóa thành dữ liệu có thể quan sát, so sánh và xếp hạng. Quá trình ấy diễn ra trước hết thông qua đánh giá (review). Một nhận xét của du khách vốn là một trải nghiệm cá nhân. Nhưng khi hàng nghìn nhận xét được tập hợp, chấm điểm, phân loại và phân tích, chúng trở thành một chỉ báo thống kê về chất lượng. AI không cần trực tiếp ngủ trong một khách sạn hay dùng bữa tại một nhà hàng. Nó có thể suy ra xác suất hài lòng từ điểm đánh giá trung bình, số lượng đánh giá, mức độ phân tán giữa các ý kiến, nội dung thường xuyên được nhắc tới và xu hướng thay đổi theo thời gian.

Tuy nhiên, danh tiếng không chỉ được hình thành từ lời người dùng viết ra. Nó còn được suy ra từ hành vi (behavior). Người dùng có thể nói rằng họ thích một điểm đến, nhưng hành vi của họ cung cấp những tín hiệu cụ thể hơn, họ có đặt phòng hay không, có hủy chuyến hay không, ở lại bao lâu, có quay lại, có giới thiệu cho người khác, có tiếp tục tìm kiếm về địa phương đó sau chuyến đi hay không. Đối với hệ thống AI, hành vi thường đáng tin hơn lời tuyên bố vì nó phản ánh một quyết định đã thực sự xảy ra.

Từ đây, danh tiếng trở thành một tập hợp nhiều lớp dữ liệu. Điểm đánh giá cho biết mức độ hài lòng được công khai. Nội dung nhận xét cho biết nguyên nhân của sự hài lòng hoặc thất vọng. Tỷ lệ chuyển đổi cho biết khả năng một gợi ý biến thành giao dịch. Tỷ lệ hủy cho biết mức độ không chắc chắn. Thời gian lưu trú, mức chi tiêu và tỷ lệ quay lại cho biết chiều sâu của trải nghiệm. Mỗi tín hiệu riêng lẻ đều không hoàn hảo, nhưng khi được kết hợp, chúng tạo ra một hồ sơ xác suất về địa phương.

Nhưng khi danh tiếng trở thành dữ liệu, nó cũng bị giản lược. Một thành phố với lịch sử phức tạp, đời sống văn hóa sâu và nhiều tầng trải nghiệm có thể bị nén thành một điểm số, một nhóm từ khóa và một xác suất hài lòng. Những gì dễ đo sẽ có xu hướng được coi trọng hơn những gì chỉ có thể cảm nhận. Sự hiếu khách có thể được phản ánh qua đánh giá, nhưng cảm giác thuộc về một nơi chốn thì khó hơn. Một cảnh quan đẹp dễ tạo ra hình ảnh và lượt chia sẻ, trong khi chiều sâu của ký ức cộng đồng rất khó trở thành tín hiệu xếp hạng.

Thách thức của địa phương không chỉ là có danh tiếng tốt, mà còn là bảo đảm danh tiếng ấy được ghi lại dưới dạng dữ liệu đủ đầy và đúng bản chất. Một trải nghiệm tốt nhưng không để lại dấu vết số sẽ đóng góp rất ít cho thứ hạng tương lai. Một trải nghiệm bình thường nhưng được ghi nhận qua hàng triệu lượt đánh giá lại có thể tạo ra lợi thế lớn. Trong nền kinh tế tác nhân, danh tiếng không còn chỉ là điều con người nghĩ về một địa phương, mà là những gì hệ thống có thể quan sát được từ cách con người phản ứng với địa phương ấy.

### 5.2.5. Cuộc đua tích lũy niềm tin

Quảng cáo là một lời khẳng định do chính chủ thể đưa ra. Niềm tin lại hình thành từ việc lời khẳng định ấy được nhiều nguồn độc lập xác nhận và được thực tế lặp lại. Một địa phương có thể quảng bá rằng mình an toàn, thuận tiện và giàu trải nghiệm. Nhưng AI sẽ tìm kiếm những tín hiệu khác, dữ liệu giao thông có ổn định không, khách du lịch có đánh giá tích cực không, thông tin giữa các nền tảng có nhất quán không, dịch vụ có thực sự đặt được không và những người từng được đề xuất có hoàn thành chuyến đi một cách thuận lợi không.

Điều này làm cho niềm tin trở thành một tài sản tích lũy. Mỗi thông tin chính xác là một khoản gửi nhỏ. Mỗi giao dịch được hoàn thành đúng cam kết là một lần củng cố. Mỗi đánh giá tích cực, mỗi nguồn độc lập xác nhận và mỗi dữ liệu được cập nhật đúng hạn đều làm tăng độ tin cậy của thực thể. Không có một chiến dịch duy nhất nào có thể tạo ra toàn bộ tài sản ấy. Nó chỉ được hình thành qua nhiều năm vận hành ổn định.

Ngược lại, niềm tin có thể mất nhanh hơn nhiều so với tốc độ được xây dựng. Một thông tin sai về giờ mở cửa, một dịch vụ không thể đặt dù hệ thống báo còn chỗ, một chuỗi đánh giá cho thấy trải nghiệm không giống quảng bá hoặc một sự cố bị nhiều nguồn xác nhận đều có thể làm giảm trọng số của địa phương. Trong môi trường xác suất, hệ thống không cần kết luận rằng một nơi hoàn toàn không đáng tin. Chỉ cần xác suất xảy ra trục trặc cao hơn một chút, nó đã có lý do để ưu tiên phương án khác.

Cuộc cạnh tranh thực sự không phải là ai nói lớn hơn mà là ai tạo ra được lịch sử xác nhận dày hơn. Những địa phương có dữ liệu ổn định, trải nghiệm nhất quán và khả năng thực hiện đúng cam kết sẽ tích lũy lợi thế qua từng vòng lặp. Những địa phương chỉ xuất hiện mạnh trong một chiến dịch nhưng thiếu nền tảng vận hành sẽ sớm mất vị trí khi các tín hiệu thực tế không xác nhận hình ảnh được quảng bá.

Cuộc đua này cũng tạo ra bất lợi lớn cho những địa phương mới hoặc ít được biết đến. Họ có thể sở hữu tài nguyên tốt nhưng thiếu lịch sử dữ liệu, thiếu đánh giá và thiếu bằng chứng về khả năng phục vụ. Vấn đề là hệ thống chưa có đủ căn cứ để tin rằng lựa chọn ấy sẽ thành công, chứ không phải AI biết chắc rằng họ kém. Trong điều kiện không chắc chắn, thuật toán thường chọn phương án đã được kiểm chứng. Niềm tin đã tích lũy vì thế trở thành một rào cản gia nhập vô hình.

Điều đó không có nghĩa quảng cáo không còn giá trị. Quảng cáo vẫn có thể tạo ra lượt thử đầu tiên, kéo thêm người dùng và mở ra cơ hội hình thành dữ liệu mới. Nhưng quảng cáo chỉ là điểm khởi đầu của vòng lặp, không phải đích đến. Nếu trải nghiệm thực tế không tạo ra đánh giá tích cực, hành vi quay lại và các nguồn xác nhận độc lập, sự chú ý mua được sẽ không chuyển thành niềm tin bền vững.

Trong nền kinh tế tác nhân, địa phương có lợi thế không nhất thiết là địa phương chi nhiều nhất cho truyền thông. Đó là địa phương xây dựng được một chuỗi bằng chứng mà máy có thể liên tục kiểm tra và khó bác bỏ. Quảng cáo có thể đưa một địa phương vào tầm nhìn trong một mùa. Chuỗi bằng chứng ấy mới là thứ giữ nó trong tập lựa chọn qua nhiều năm.

Bất đối xứng của niềm tin. Niềm tin của hệ thống vào dữ liệu về một địa phương thường được xây dựng chậm qua nhiều lần xác nhận, nhưng có thể suy giảm nhanh sau một số lỗi có tính hệ thống hoặc được nhiều nguồn độc lập ghi nhận. Chương này không có số liệu đo tốc độ của hai chiều ấy, nên phát biểu được để ở dạng giả thuyết chờ kiểm chứng. Danh tiếng cần mười năm để dựng và một mùa lễ hội sai giờ mở cửa để lung lay.

## 5.3. NỀN TẢNG CẠNH TRANH ĐỂ ĐỊNH NGHĨA TIÊU CHUẨN

Trong nền kinh tế tìm kiếm, các nền tảng chủ yếu cạnh tranh để trở thành nơi người dùng nhập truy vấn đầu tiên. Google cạnh tranh với Bing, Facebook cạnh tranh với TikTok, Amazon cạnh tranh với các sàn thương mại điện tử khác. Mục tiêu của họ là thu hút sự chú ý, kéo dài thời gian sử dụng và gia tăng số lần nhấp chuột. Nhưng khi AI tác nhân xuất hiện, bản chất của cuộc cạnh tranh thay đổi. Điều các nền tảng muốn giành lấy không còn là sự chú ý của con người, mà là quyền được thay mặt con người đưa ra quyết định.

Khi một người dùng giao cho AI nhiệm vụ lên kế hoạch chuyến đi, lựa chọn nhà hàng, đặt khách sạn hay mua sản phẩm, họ đồng thời trao cho nền tảng AI đó quyền xác định tập lựa chọn ban đầu. Từ thời điểm ấy, mọi địa phương, doanh nghiệp hay điểm đến muốn xuất hiện trong phiên quyết định đó đều phải đi qua cánh cửa của nền tảng được người dùng ủy nhiệm. Được AI lựa chọn đồng nghĩa với cơ hội được con người biết đến. Bị loại khỏi quá trình lựa chọn đồng nghĩa với việc gần như biến mất khỏi thị trường, dù chất lượng thực tế có thể không hề thua kém.

Chính vì vậy, các địa phương không còn chỉ cạnh tranh với nhau trong cùng một bảng xếp hạng. Họ còn phải thích ứng với nhiều hệ thống AI khác nhau, mỗi hệ thống có cách thu thập dữ liệu, đánh giá bằng chứng, xây dựng niềm tin và xếp hạng riêng. Cuộc cạnh tranh giữa các nền tảng vì thế không chỉ là cuộc đua về công nghệ, mà còn là cuộc đua để thiết lập các tiêu chuẩn mà toàn bộ thế giới số phải tuân theo.

### 5.3.1. Quyền đề xuất không còn tập trung ở một cổng

Trong gần hai thập niên, Google gần như đồng nghĩa với việc tìm kiếm thông tin trên internet. Muốn biết điều gì, con người mở Google. Muốn tìm khách sạn, nhà hàng hay điểm tham quan, hành trình đầu tiên thường cũng bắt đầu từ một ô tìm kiếm. Chính vị trí "điểm khởi đầu" này giúp Google trở thành cổng phân phối lưu lượng lớn nhất của nền kinh tế số. Nhưng nền kinh tế tác nhân đang làm thay đổi cấu trúc ấy. Quyền đề xuất không còn tập trung vào một công cụ tìm kiếm duy nhất mà được phân tán cho nhiều loại nền tảng khác nhau. Người dùng có thể hỏi một mô hình ngôn ngữ để lập kế hoạch chuyến đi, nhờ một siêu ứng dụng đặt toàn bộ dịch vụ, hoặc để một nền tảng du lịch trực tuyến tự động gợi ý khách sạn và hoàn tất việc đặt phòng. Mỗi nền tảng đều cố gắng trở thành nơi đầu tiên mà người dùng giao quyền lựa chọn.

Điều này tạo ra một hệ sinh thái đề xuất đa trung tâm. Công cụ tìm kiếm vẫn rất quan trọng, nhưng không còn là cánh cửa duy nhất. Các mô hình AI tổng hợp tri thức từ nhiều nguồn khác nhau trước khi trả lời. Các nền tảng du lịch trực tuyến tối ưu cho việc chuyển đổi thành giao dịch. Các siêu ứng dụng lại tích hợp di chuyển, thanh toán, lưu trú và tiêu dùng vào cùng một hành trình. Người dùng càng ít phải chuyển đổi giữa các ứng dụng, nền tảng đó càng có nhiều cơ hội giữ quyền quyết định.

Vì thế, quyền đề xuất đang chuyển từ một cổng tìm kiếm thống trị sang một mạng lưới các hệ thống AI và nền tảng số cùng cạnh tranh. Cuộc chơi không còn là "được Google tìm thấy", mà là "được nhiều hệ thống AI khác nhau đồng thời tin tưởng và lựa chọn".

### 5.3.2. Mỗi nền tảng có một logic tối ưu

Mặc dù đều thực hiện chức năng đề xuất, các nền tảng không theo đuổi cùng một mục tiêu. Chính mục tiêu kinh doanh khác nhau khiến chúng xây dựng những hệ thống đánh giá và xếp hạng khác nhau. Không tồn tại một thuật toán chung cho mọi nền tảng, cũng không tồn tại một bảng xếp hạng duy nhất phản ánh toàn bộ thế giới.

Một công cụ tìm kiếm hướng tới việc cung cấp câu trả lời đáng tin cậy và toàn diện nhất. Một mô hình ngôn ngữ tối ưu cho khả năng tổng hợp tri thức và tạo ra câu trả lời mạch lạc. Một nền tảng đặt phòng ưu tiên những lựa chọn có khả năng chuyển đổi thành giao dịch cao. Một siêu ứng dụng lại tối ưu cho việc giữ người dùng ở lại trong hệ sinh thái càng lâu càng tốt. Khi mục tiêu khác nhau, tín hiệu được coi trọng cũng khác nhau.

Nguồn dữ liệu mà mỗi nền tảng sử dụng cũng không giống nhau. Có hệ thống dựa nhiều vào dữ liệu mở trên internet, có hệ thống khai thác dữ liệu giao dịch nội bộ, có hệ thống học từ hành vi của hàng trăm triệu người dùng và có hệ thống kết hợp tất cả các nguồn này. Cùng một khách sạn, một bảo tàng hay một lễ hội có thể được đánh giá rất khác nhau tùy theo loại dữ liệu mà từng nền tảng đang sở hữu.

Đây là lý do địa phương không thể chỉ theo đuổi một chiến lược tối ưu duy nhất. Muốn hiện diện trong nền kinh tế tác nhân, địa phương phải xây dựng năng lực đáp ứng đồng thời nhiều hệ thống đánh giá khác nhau: dữ liệu phải đáng tin cậy, thông tin phải nhất quán, dịch vụ phải khả hành động và trải nghiệm phải tạo ra tín hiệu tích cực sau mỗi lần được lựa chọn. Trên thực tế, mỗi hệ thống thường tối ưu đồng thời nhiều mục tiêu, còn trọng số giữa chúng thì không được công khai.

Mỗi hệ thống AI tối ưu cho một mục tiêu khác nhau nên cũng đánh giá địa phương theo những tiêu chí khác nhau. Công cụ tìm kiếm ưu tiên tạo ra câu trả lời đáng tin cậy và bao quát, vì vậy chủ yếu dựa vào dữ liệu mở trên internet và thường đánh giá cao những nguồn chính thức cùng dữ liệu có cấu trúc rõ ràng. Mô hình ngôn ngữ lại hướng đến việc tạo ra câu trả lời mạch lạc và hữu ích, kết hợp giữa tri thức đã được huấn luyện với dữ liệu truy hồi theo thời gian thực, nên có xu hướng ưu tiên những thông tin được nhiều nguồn độc lập cùng xác nhận và nhắc lại. Các nền tảng đặt chỗ không tối ưu cho tri thức mà cho khả năng tạo ra giao dịch, vì thế, chúng dựa nhiều vào dữ liệu tồn kho, tình trạng sẵn sàng và lịch sử đặt chỗ để ưu tiên những điểm đến còn phòng, còn vé, giá cạnh tranh và có khả năng phục vụ ngay. Trong khi đó, bản đồ và hệ thống dẫn đường tối ưu cho độ chính xác về không gian và thời gian, nên đặc biệt coi trọng dữ liệu vị trí, giờ mở cửa và hành vi di chuyển thực tế của người dùng. Chính sự khác biệt về mục tiêu tối ưu này giải thích vì sao cùng một địa phương có thể được xếp hạng rất cao trên một hệ thống nhưng gần như không xuất hiện trên hệ thống khác, dù bản thân địa phương không hề thay đổi. Thứ thay đổi là hàm mục tiêu (objective function) mà mỗi hệ thống sử dụng để lựa chọn và xếp hạng thông tin.

Chi phí đa chuẩn. Khi không còn một cổng đề xuất duy nhất, chi phí khả đọc của địa phương tăng theo số nền tảng phải phục vụ, còn lợi ích chỉ tăng theo số nền tảng thực sự có người dùng của địa phương ấy. Nếu hai đường này không song song thì tối ưu cho một nền tảng duy nhất là chiến lược yếu, còn tối ưu cho mọi nền tảng là chiến lược vượt ngân sách. Điểm cần kiểm chứng là độ dốc thực tế của từng đường trong từng bối cảnh.

### 5.3.3. Người dùng không còn nhìn thấy quá trình lựa chọn

Trong thời đại công cụ tìm kiếm, người dùng không nhìn thấy bản thân thuật toán nhưng vẫn nhìn thấy kết quả xếp hạng dưới dạng một danh sách tương đối rộng. Một trang kết quả hiển thị hàng chục liên kết, quảng cáo, bản đồ, hình ảnh và đánh giá. Họ có thể cuộn xuống, mở nhiều nguồn, so sánh và tự đưa ra phán xét cuối cùng. Logic, trọng số và tiêu chí thì vẫn không công khai, nhưng kết quả của quá trình chọn lọc được trình bày một phần trước mắt người dùng.

Tác nhân AI thay đổi đáng kể trải nghiệm đó. Thay vì một danh sách lựa chọn, người dùng ngày càng nhận được một câu trả lời đã được tổng hợp. Họ không nhìn thấy hàng trăm nguồn dữ liệu mà chỉ nhìn thấy ba gợi ý khách sạn, hai điểm đến hoặc một lịch trình hoàn chỉnh. Phần lớn quá trình truy hồi, đối chiếu, đánh giá và loại bỏ ứng viên đều diễn ra bên trong hệ thống.

Điều này tạo ra một lớp "vô hình thuật toán". Người dùng không biết có bao nhiêu lựa chọn đã bị loại bỏ, tiêu chí nào được sử dụng để xếp hạng hay vì sao một địa phương được ưu tiên hơn địa phương khác. Những quyết định từng diễn ra công khai trên trang kết quả tìm kiếm giờ được nén lại thành một câu trả lời ngắn gọn, khiến quá trình lựa chọn trở nên khó quan sát hơn bao giờ hết.

Lớp vô hình ấy có thể đo được. Cũng trong nghiên cứu Pew nói trên, người dùng gặp một bản tóm tắt AI kết thúc hẳn phiên duyệt web ngay sau đó trong 26% số lượt, so với 16% ở những trang chỉ có kết quả tìm kiếm thông thường.[1] Nghĩa là bản tóm

1. Pew Research Center, ngày 22 tháng 7 năm 2025: 900 người trưởng thành tại Mỹ, 68.879 lượt tìm kiếm trong tháng 3 năm 2025.

tắt không chỉ hút bớt lượt nhấp, nó còn kết thúc sớm hành trình tìm hiểu của người dùng. Với một địa phương chưa nằm trong tóm tắt ấy, cơ hội được xem xét không giảm đi mà biến mất. Nó biến mất trước khi người dùng kịp biết là mình đã không xem xét gì.

Đối với địa phương, đây là một thay đổi có ý nghĩa chiến lược. Trước đây, chỉ cần xuất hiện ở trang đầu kết quả tìm kiếm đã là một lợi thế lớn. Ngày nay, ngay cả việc nằm trong tập dữ liệu mà AI truy hồi cũng chưa đủ. Điều quan trọng hơn là phải vượt qua toàn bộ chuỗi đánh giá nội bộ để trở thành một trong số rất ít lựa chọn cuối cùng được AI trình bày cho người dùng. Nói cách khác, thuật toán không biến mất. Ngược lại, nó trở nên quyền lực hơn vì không còn hiện diện như một bảng xếp hạng công khai mà như một "hộp đen" đưa ra quyết định thay con người. Người dùng chỉ nhìn thấy câu trả lời. Còn cuộc cạnh tranh giữa hàng nghìn thực thể đã kết thúc từ rất lâu trước khi câu trả lời ấy xuất hiện trên màn hình.

### 5.3.4. Bốn loại quyền lực trong hệ sinh thái đề xuất

Cuộc cạnh tranh trong nền kinh tế tác nhân không chỉ diễn ra giữa các mô hình AI, mà giữa những hệ sinh thái số hoàn chỉnh. Mỗi hệ sinh thái đều cố gắng kiểm soát toàn bộ chuỗi giá trị, từ dữ liệu, suy luận, đề xuất cho đến thực thi giao dịch. Ai sở hữu được chuỗi này sẽ nắm quyền phân phối các quyết định của hàng triệu người dùng.

Một hệ sinh thái có thể bắt đầu từ hệ điều hành trên điện thoại, tiếp tục với trợ lý AI, kết nối với bản đồ, lịch, thanh toán và cửa hàng ứng dụng. Một hệ sinh thái khác lại xuất phát từ

công cụ tìm kiếm, mở rộng sang bản đồ, email, video, quảng cáo và AI. Có hệ sinh thái lấy mạng xã hội làm trung tâm để học từ hành vi của hàng tỷ người dùng, trong khi những hệ sinh thái du lịch tập trung vào dữ liệu đặt phòng, đánh giá và khả năng hoàn tất giao dịch.

Điểm khác biệt không nằm ở từng sản phẩm riêng lẻ mà ở khả năng phối hợp của toàn bộ hệ thống. Một người có thể hỏi AI lập kế hoạch du lịch, xem bản đồ, đặt vé máy bay, đặt khách sạn, thanh toán và nhận hướng dẫn di chuyển mà không cần rời khỏi cùng một hệ sinh thái. Càng ít điểm phải chuyển sang nền tảng khác, hệ sinh thái đó càng có khả năng giữ quyền ra quyết định.

Chính vì vậy, cuộc cạnh tranh giữa các hệ sinh thái ngày càng khốc liệt. Tại thời điểm cuốn sách được viết, có thể quan sát thấy năm nhóm chủ thể tranh giành ở năm lớp khác nhau của chuỗi. Lớp thiết bị và hệ điều hành nắm điểm khởi đầu của mọi phiên sử dụng, nên bên kiểm soát lớp này quyết định trợ lý nào được gọi trước (Apple, Google). Lớp tìm kiếm và bản đồ nắm dữ liệu về ý định và về không gian, nên bên kiểm soát nó có khả năng suy ra một phần ý định của người dùng trước khi ý định ấy được phát biểu đầy đủ (Google). Lớp giao diện tác nhân nắm nơi nhiệm vụ được giao và được thực thi, nên bên kiểm soát nó có cơ hội trở thành cổng vào mặc định của mọi tác vụ (OpenAI cùng những bên phát triển trợ lý khác). Lớp mạng xã hội nắm dữ liệu hành vi và mạng quan hệ, thứ cho phép suy ra sở thích mà người dùng chưa từng khai báo (Meta). Lớp dữ liệu giao dịch chuyên ngành nắm tồn kho, giá và lịch sử đặt chỗ, thứ mà không mô hình ngôn ngữ nào tự sinh ra được (Booking, TripAdvisor cùng các kênh phân phối). Những cái tên trong ngoặc chỉ là ví dụ của thời điểm này. Điều bền hơn tên là năm lớp ấy và trật tự của chúng. Cần lưu ý rằng năm lớp trên không tương ứng một đối một với năm loại

quyền lực. Chúng cạnh tranh bằng bốn loại quyền lực nền tảng. Một chủ thể có thể đồng thời nắm nhiều loại quyền lực và một loại quyền lực có thể xuất hiện ở nhiều lớp.

Trong nền kinh tế tác nhân, quyền lực không chỉ nằm ở việc ai sở hữu dữ liệu mà còn ở việc ai kiểm soát các lớp hạ tầng quyết định dòng thông tin và giao dịch. ***Quyền sở hữu giao diện*** với người dùng là quyền kiểm soát nơi câu hỏi được đặt ra và nơi câu trả lời được hiển thị. Đây là lớp quyền lực khó tiếp cận nhất vì phần lớn thuộc về các nền tảng số toàn cầu và địa phương hầu như không có khả năng thương lượng trực tiếp. ***Quyền sở hữu dữ liệu*** hành vi lại quyết định ai biết người dùng đã xem gì, bỏ qua điều gì, đặt dịch vụ nào hay hủy giao dịch nào. Ở lớp này, địa phương vẫn có thể tạo ảnh hưởng thông qua các thỏa thuận chia sẻ dữ liệu hai chiều với doanh nghiệp và nền tảng. ***Quyền kiểm soát khả năng giao dịch*** liên quan đến việc một đề xuất có thể được chuyển thành hành động đặt chỗ, mua vé hay thanh toán hay không. Đây là quyền có thể được thương lượng thông qua các điều khoản hợp đồng với các kênh phân phối, nền tảng đặt chỗ và đơn vị vận hành dịch vụ. Cuối cùng, ***quyền đặt tiêu chuẩn dữ liệu và kết nối*** là quyền quyết định định dạng và quy tắc mà các bên khác phải tuân theo để có thể liên thông với hệ thống. Mặc dù khó đạt được, địa phương vẫn có thể gia tăng ảnh hưởng bằng cách tham gia sớm vào các liên minh, tổ chức hoặc nhóm xây dựng tiêu chuẩn. Bốn lớp quyền lực này cho thấy một thực tế quan trọng: địa phương không thể kiểm soát toàn bộ hệ sinh thái số, nhưng hoàn toàn có thể lựa chọn những lớp quyền mà mình có khả năng tác động và đầu tư để nâng cao vị thế trong chuỗi giá trị của nền kinh tế tác nhân. Ba loại quyền lực đầu thuộc về nền tảng gần như trọn vẹn. Loại thứ tư là loại duy nhất mà một chủ thể công có thể dự phần một cách có ý nghĩa. Đây cũng là loại quyết định lâu dài nhất.

Đối với địa phương, điều này có nghĩa là không còn một "cửa ngõ" duy nhất để tiếp cận người dùng. Mỗi hệ sinh thái đều có tiêu chuẩn dữ liệu, nguồn tín hiệu và cơ chế đánh giá riêng. Một địa phương có thể rất nổi bật trong một hệ sinh thái nhưng gần như vô hình trong hệ sinh thái khác nếu thiếu dữ liệu, thiếu kết nối hoặc không đáp ứng các điều kiện để AI có thể thực hiện hành động. Do đó, chiến lược của địa phương không thể chỉ tối ưu cho một nền tảng cụ thể. Mục tiêu dài hạn phải là xây dựng một hạ tầng dữ liệu đủ chuẩn, đủ nhất quán và đủ khả hành động để có thể được nhận diện, đánh giá và lựa chọn trong nhiều hệ sinh thái cùng lúc. Trong nền kinh tế tác nhân, lợi thế bền vững không nằm ở việc thắng trên một nền tảng, mà ở khả năng hiện diện ổn định trên toàn bộ mạng lưới các hệ sinh thái AI đang cạnh tranh để trở thành người đại diện đầu tiên của người dùng.

### 5.3.5. Không phải một chuẩn thắng mà nhiều lớp chuẩn

Trong mọi cuộc cách mạng công nghệ, cuộc cạnh tranh cuối cùng thường không phải giữa các sản phẩm mà giữa các tiêu chuẩn. Khi một tiêu chuẩn trở thành mặc định, toàn bộ hệ sinh thái còn lại buộc phải thích nghi với nó. Bàn phím QWERTY, giao thức TCP/IP hay định dạng PDF đều cho thấy giá trị lớn nhất không nằm ở việc chúng hoàn hảo nhất, mà ở việc chúng được chấp nhận rộng rãi nhất. Nền kinh tế tác nhân cũng sẽ đi theo quy luật đó.

Cuộc cạnh tranh hiện nay không chỉ là AI nào thông minh hơn, mà là chuẩn nào sẽ trở thành ngôn ngữ chung để AI và thế giới thực kết nối với nhau. Một địa phương có thể sở hữu tài nguyên du lịch hấp dẫn, dịch vụ chất lượng và thương hiệu mạnh, nhưng nếu dữ liệu không theo chuẩn mà các hệ thống AI

có thể đọc, nếu dịch vụ không có giao diện kết nối chuẩn, hoặc nếu giao dịch không thể được thực hiện tự động, địa phương đó vẫn có nguy cơ bị loại khỏi chuỗi quyết định.

Tiêu chuẩn đầu tiên là chuẩn dữ liệu. Các thực thể phải có định danh rõ ràng, thuộc tính thống nhất, quan hệ được mô tả nhất quán và có thể được nhiều hệ thống cùng hiểu theo một cách. Khi mỗi nền tảng mô tả một địa điểm theo một cấu trúc khác nhau, AI phải dành nguồn lực để diễn giải và đối chiếu. Ngược lại, khi dữ liệu được chuẩn hóa, chi phí nhận diện giảm xuống và xác suất được truy hồi tăng lên. Trong nền kinh tế tác nhân, dữ liệu không chỉ cần đầy đủ mà còn phải tương thích.

Tiêu chuẩn thứ hai là chuẩn giao diện kết nối (API). AI không thể hành động nếu không có một cách thức chuẩn để truy cập thông tin và thực hiện thao tác. Một hệ thống đặt phòng, bán vé, thanh toán hay xác nhận dịch vụ chỉ thực sự trở thành một phần của nền kinh tế tác nhân khi các tác nhân AI có thể kết nối một cách an toàn, ổn định và có cấu trúc. Khả đọc vì thế phải được mở rộng thành khả kết nối.

Tiêu chuẩn thứ ba là chuẩn giao dịch. Đề xuất chỉ tạo ra giá trị khi có thể chuyển thành hành động. Điều đó đòi hỏi những quy tắc thống nhất về đặt chỗ, thanh toán, xác nhận, hoàn hủy, cập nhật trạng thái và chia sẻ dữ liệu sau giao dịch. Một AI sẽ ưu tiên những hệ thống mà nó có thể hoàn tất toàn bộ quy trình với mức độ tin cậy cao, thay vì những nơi buộc người dùng phải thực hiện nhiều bước thủ công.

Lịch sử internet cho thấy thị trường thường không thuộc về công nghệ mạnh nhất mà thuộc về công nghệ trở thành chuẩn kết nối. Trong nền kinh tế tác nhân cũng vậy, lợi thế dài hạn sẽ

thuộc về những địa phương, doanh nghiệp và nền tảng chủ động xây dựng dữ liệu theo chuẩn chung, mở các giao diện kết nối chuẩn và tham gia vào các quy trình giao dịch chuẩn hóa. Cuộc cạnh tranh cuối cùng không phải là ai có nhiều dữ liệu hơn, mà là dữ liệu của ai dễ được phần lớn hệ thống sử dụng nhất.

Cũng cần bỏ giả định rằng chỉ có một chuẩn duy nhất sẽ thắng. Kinh nghiệm của các hạ tầng số trước đây cho thấy thứ hình thành thường là một chồng nhiều lớp chuẩn, mỗi lớp có bên thắng riêng và tốc độ hội tụ riêng. Với hạ tầng đề xuất, có thể phân biệt ít nhất bảy lớp: chuẩn nhận diện thực thể, chuẩn mô tả thuộc tính, chuẩn tồn kho và khả năng phục vụ, chuẩn giao dịch, chuẩn thanh toán, chuẩn xác nhận danh tính và chất lượng, cuối cùng là chuẩn truy xuất nguồn gốc và phân định trách nhiệm khi máy sai. Với một địa phương, cách đọc thực dụng là như sau. Ba lớp đầu quyết định mình có được đọc đúng hay không. Đây là ba lớp mình tác động được ngay. Ba lớp giữa quyết định mình có bán được hay không. Ba lớp này cần đối tác. Lớp cuối quyết định mình có kênh khiếu nại hay không khi hình ảnh của mình bị dựng sai. Đây là lớp còn trống nhất, cũng là chủ đề mà các chương sau của sách sẽ trở lại.

Đối với địa phương, đây là một thay đổi mang tính chiến lược. Thay vì chỉ đầu tư vào nội dung hay quảng bá, cần coi việc xây dựng chuẩn dữ liệu, chuẩn kết nối và chuẩn giao dịch như một phần của hạ tầng công cộng. Bởi trong nền kinh tế tác nhân, tiêu chuẩn không chỉ là vấn đề kỹ thuật. Tiêu chuẩn chính là quyền lực. Ai định nghĩa được tiêu chuẩn sẽ định hình cách AI nhìn nhận thế giới và ai thích ứng sớm với tiêu chuẩn đó sẽ có lợi thế trong cuộc cạnh tranh giành quyền được lựa chọn.

> ▪ ***Về đầu tư khả đọc đa nền tảng:***
> • *Không được xây dựng dữ liệu địa phương theo định dạng riêng của một nền tảng cụ thể. Dữ liệu phải được giữ ở một nguồn gốc trung lập rồi mới xuất ra từng nền tảng.*
> • *Không được ký thỏa thuận độc quyền dữ liệu với một nền tảng phân phối, kể cả khi thỏa thuận ấy đi kèm cam kết quảng bá.*
> • *Không được đo hiệu quả của đầu tư khả đọc bằng thứ hạng trên một nền tảng duy nhất trong một quý duy nhất.*

Lịch sử cho thấy chuẩn mở luôn thắng cuối cùng. Giao thức internet đã thắng các mạng đóng, định dạng mở đã thắng định dạng độc quyền. Vậy địa phương chỉ cần chờ, rồi chuẩn chung sẽ xuất hiện và chi phí thích ứng sẽ giảm về gần bằng không. Lập luận này bỏ qua độ dài của giai đoạn chuyển tiếp và ai phải trả giá trong giai đoạn ấy. Chuẩn mở thường thắng, nhưng thắng sau mười đến hai mươi năm cạnh tranh. Trong khoảng thời gian đó thị phần đã được phân chia xong. Một địa phương vắng mặt suốt giai đoạn chuyển tiếp sẽ bước vào kỷ nguyên chuẩn mở với một hồ sơ dữ liệu mỏng và một lịch sử xác nhận ngắn, tức là bước vào cuộc chơi mới với đúng bất lợi của cuộc chơi cũ. Thêm nữa, chuẩn mở làm giảm chi phí kết nối chứ không làm giảm chi phí tạo ra dữ liệu đúng. Phần đắt nhất của khả đọc nằm ở việc mô tả một nơi chốn cho chính xác. Phần ấy không có chuẩn nào gánh hộ.

PHẦN B

# VÌ SAO LỢI THẾ TỰ KHUẾCH ĐẠI?

## 5.4. HIỆU ỨNG MATTHEW CỦA THUẬT TOÁN

Trong nền kinh tế tác nhân, thuật toán không chỉ phản ánh sự phổ biến mà còn liên tục khuếch đại nó. Mỗi lần một địa phương được AI lựa chọn và đề xuất đều tạo ra cơ hội thu hút thêm người dùng, dữ liệu và bằng chứng mới. Những tín hiệu mới này lại trở thành cơ sở để các hệ thống AI đánh giá địa phương đó đáng tin cậy và phù hợp hơn trong những lần truy hồi tiếp theo. Vì vậy, lợi thế ban đầu không dừng lại ở một lần xuất hiện mà được tích lũy thành một vòng lặp tự củng cố.

Cơ chế này chính là biểu hiện của hiệu ứng Matthew trong môi trường thuật toán, thực thể đã có nhiều lợi thế sẽ tiếp tục nhận thêm lợi thế, còn những thực thể ít được nhìn thấy ngày càng khó bước vào tập lựa chọn. Đối với địa phương, điều này đồng nghĩa với việc cạnh tranh không chỉ diễn ra ở hiện tại mà còn diễn ra trong khả năng tích lũy niềm tin của máy theo thời gian. Một khi vòng lặp tích cực đã hình thành, khoảng cách giữa các địa phương có thể ngày càng lớn dù chất lượng thực tế không chênh lệch tương ứng. Đây là một trong những nguyên nhân quan trọng khiến khả năng được AI đề xuất có xu hướng tập trung vào một số ít địa phương và ngày càng khó đảo ngược.

### 5.4.1. Giàu càng giàu trong thế giới số

Hiệu ứng Matthew được nhà xã hội học Robert K. Merton sử dụng để mô tả hiện tượng "ai đã có sẽ được cho thêm". Trong khoa

học, những nhà nghiên cứu nổi tiếng thường nhận được nhiều trích dẫn hơn. Trong kinh tế, doanh nghiệp lớn dễ tiếp cận vốn hơn. Trong truyền thông, người có nhiều người theo dõi thường tiếp tục thu hút thêm người theo dõi. Thế giới số cũng vận hành theo cùng một logic. Merton phát biểu điều này năm 1968 trên tạp chí Science, khi ông quan sát rằng một công trình khoa học do người đã nổi tiếng ký tên được trích dẫn nhiều hơn hẳn một công trình tương đương của người chưa ai biết. Ông gọi đó là hiệu ứng Matthew, theo câu trong Phúc âm Mátthêu về việc ai đã có sẽ được cho thêm.[1] Điều đáng chú ý là Merton không mô tả một sự bất công do ai đó cố ý tạo ra. Ông mô tả một cơ chế phân bổ sự chú ý trong điều kiện người đọc có thời gian hữu hạn. Gần sáu mươi năm sau, cơ chế ấy được cài vào phần mềm.

Đối với các hệ thống AI, "giàu" không phải là giàu tiền bạc mà là giàu tín hiệu số. Một địa phương có nhiều dữ liệu, nhiều đánh giá, nhiều nguồn độc lập xác nhận, nhiều lượt được nhắc đến và nhiều người lựa chọn sẽ có nhiều bằng chứng hơn để thuật toán suy luận. Càng nhiều bằng chứng, xác suất được xếp hạng cao càng lớn. Khi được xếp hạng cao hơn, địa phương lại có cơ hội tiếp cận nhiều người dùng hơn, từ đó tạo ra thêm dữ liệu và tiếp tục củng cố vị thế của mình. Bởi vậy, trong nền kinh tế tác nhân, cạnh tranh không còn là cuộc đua bắt đầu từ cùng một vạch xuất phát. Mỗi chu kỳ lựa chọn đều làm thay đổi vị trí của các địa phương trong hệ thống và khoảng cách ấy có xu hướng ngày càng mở rộng nếu không có những tín hiệu mới đủ mạnh để phá vỡ quán tính của thuật toán.

1. Robert K. Merton, "The Matthew Effect in Science", Science, tập 159, số 3810, 1968, trang 56 tới 63.

### 5.4.2. Được chọn tạo ra thêm lý do để được chọn

Điểm đặc biệt của các hệ thống AI là quyết định hôm nay trở thành dữ liệu cho quyết định ngày mai. Mỗi lần một địa phương được AI đề xuất đều không chỉ tạo ra một lượt truy cập mà còn tạo ra hàng loạt tín hiệu mới, thêm lượt tìm kiếm, thêm đánh giá, thêm hình ảnh, thêm bài viết, thêm giao dịch, thêm thời gian lưu trú, thêm chia sẻ trên mạng xã hội và thêm dữ liệu hành vi của người dùng. Tất cả những tín hiệu này lại quay trở về làm đầu vào cho các mô hình xếp hạng.

Đó là vòng lặp phản hồi (feedback loop). Thuật toán lựa chọn một địa phương, người dùng tương tác với địa phương đó, các tương tác được ghi nhận thành dữ liệu mới và dữ liệu mới lại khiến thuật toán có thêm cơ sở để tiếp tục lựa chọn địa phương ấy trong tương lai. Quá trình này diễn ra liên tục, gần như không có điểm dừng.

Ngược lại, một địa phương đã bước vào vòng lặp tích cực sẽ liên tục nhận được những "lý do mới" để được lựa chọn. Thuật toán không chỉ nhìn vào quá khứ mà còn sử dụng chính kết quả của những lần lựa chọn trước như bằng chứng cho các lần lựa chọn tiếp theo. Vì vậy, trong nền kinh tế tác nhân, được chọn không phải là kết thúc của quá trình cạnh tranh mà là khởi đầu của một chu kỳ tích lũy lợi thế mới. Đây chính là cơ chế khiến hiệu ứng Matthew trở thành một trong những động lực mạnh mẽ nhất của cạnh tranh thuật toán.

Trong hệ thống có vòng phản hồi, kết quả của lần xếp hạng trước trở thành dữ liệu đầu vào của lần xếp hạng sau. Tự củng cố. Vì vậy một lợi thế ban đầu nhỏ có thể sinh ra chênh lệch lớn theo thời gian, kể cả khi chất lượng giữa các phương án không chênh

lệch tương ứng. Đây là dạng lợi thế tích lũy đã được ghi nhận trong nhiều lĩnh vực trước khi xuất hiện các hệ thống đề xuất.

Hai vòng lặp của cơ chế xếp hạng vận hành song song nhưng dẫn đến hai quỹ đạo hoàn toàn khác nhau. Ở vòng lặp của những nơi đã được đề xuất, mỗi lần xuất hiện trước người dùng lại tạo thêm giao dịch. Giao dịch tạo ra nhiều dữ liệu tương tác hơn, dữ liệu tích lũy thêm các tín hiệu xác nhận và những tín hiệu ấy tiếp tục làm tăng xác suất được hệ thống đề xuất trong những lần sau. Càng được chọn, khả năng tiếp tục được chọn càng lớn. Ngược lại, vòng lặp của những nơi ít được đề xuất bắt đầu từ việc hiếm khi xuất hiện trong danh sách gợi ý. Ít được nhìn thấy dẫn đến ít giao dịch, dữ liệu tương tác vì thế ngày càng mỏng, tín hiệu xác nhận không đủ mạnh, khiến xác suất được đề xuất tiếp tục giảm. Càng ít được chọn, cơ hội xuất hiện trong tương lai càng nhỏ. Điều đáng chú ý là cả hai vòng lặp đều vận hành theo cùng một cơ chế. Chúng không đòi hỏi chất lượng của điểm đến phải tốt lên hay xấu đi để tạo ra khác biệt. Chỉ cần thời gian, lợi thế ban đầu sẽ có xu hướng được khuếch đại, còn bất lợi ban đầu cũng dễ bị củng cố theo cùng một logic.

### 5.4.3. Thiên lệch theo mức độ phổ biến

Một trong những đặc điểm phổ biến của các hệ thống đề xuất là xu hướng ưu tiên những gì đã được nhiều người lựa chọn. Hiện tượng này được gọi là thiên lệch theo mức độ phổ biến (popularity bias). Khi dữ liệu còn hạn chế và có nhiều phương án cạnh tranh, thuật toán thường coi sự phổ biến là một tín hiệu giúp giảm rủi ro. Nếu nhiều người đã lựa chọn một địa phương, nhiều khả năng địa phương đó đáp ứng được nhu cầu của những người dùng tiếp theo.

Vấn đề là sự phổ biến không chỉ phản ánh chất lượng mà còn tạo ra chất lượng trong mắt thuật toán. Một điểm đến có nhiều lượt đánh giá sẽ tiếp tục thu hút thêm nhiều lượt xem. Một nhà hàng xuất hiện ở vị trí đầu sẽ có nhiều khách hơn, từ đó nhận thêm đánh giá và tiếp tục giữ vị trí cao. Một địa phương thường xuyên được AI giới thiệu sẽ có thêm du khách, thêm bài viết, thêm hình ảnh và thêm dữ liệu hành vi. Chính những dữ liệu mới ấy lại củng cố nhận định rằng địa phương này "xứng đáng" được ưu tiên.

Có một thí nghiệm cho thấy điều này gần như ở dạng thuần khiết. Năm 2006, Matthew Salganik, Peter Dodds và Duncan Watts dựng một thị trường âm nhạc nhân tạo với 14.341 người tham gia, cho họ nghe và tải những bài hát chưa ai biết.[1] Một nhóm quyết định trong im lặng, không biết người khác đã chọn gì. Các nhóm còn lại nhìn thấy số lượt tải của những người đi trước. Kết quả công bố trên Science cho thấy, khi cường độ ảnh hưởng xã hội tăng lên, cả bất bình đẳng lẫn tính khó đoán của thành công đều tăng theo. Chất lượng vẫn có vai trò, nhưng chỉ ở hai đầu mút. Bài hay nhất hiếm khi xếp bét, bài dở nhất hiếm khi lên đỉnh, còn mọi kết quả khác đều có thể xảy ra.

Đó là lý do có thể nói rằng đám đông sinh ra đám đông. Mỗi quyết định của hàng nghìn người dùng trở thành bằng chứng cho hàng triệu quyết định tiếp theo. Thuật toán không cố ý tạo ra thiên lệch, nhưng cách nó học từ dữ liệu khiến những gì đã phổ biến có xu hướng ngày càng phổ biến hơn.

Điểm khiến kết quả ấy đáng lo không phải là bất bình đẳng tăng, điều đó ai cũng đoán được. Điểm đáng lo là tính khó đoán

1. Matthew J. Salganik, Peter S. Dodds và Duncan J. Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market", Science, tập 311, số 5762, 2006, trang 854 tới 856.

cũng tăng. Nghĩa là khi tín hiệu đám đông được đưa vào hệ thống, người ta không chỉ mất công bằng mà còn mất luôn khả năng dự báo ai sẽ thắng. Một địa phương đầu tư nghiêm túc vào chất lượng vẫn có thể thua một địa phương tương đương chỉ vì vài trăm lượt tương tác ngẫu nhiên trong tuần đầu tiên.

Đối với địa phương, điều này có nghĩa rằng lợi thế cạnh tranh không chỉ đến từ tài nguyên hay trải nghiệm, mà còn đến từ khả năng tạo ra động lượng dữ liệu. Khi một địa phương đạt đến một ngưỡng phổ biến nhất định, chính quy mô của cộng đồng người dùng sẽ trở thành động cơ tiếp tục thúc đẩy sự tăng trưởng.

### 5.4.4. Bốn dạng thiếu dữ liệu của địa phương mới

Nếu hiệu ứng Matthew là lợi thế của người dẫn đầu thì vấn đề khởi đầu lạnh (cold start) là thách thức của những người mới tham gia. Thuật toán luôn cần bằng chứng để đưa ra khuyến nghị, nhưng một địa phương mới hoặc ít được biết đến lại gần như chưa có dữ liệu để chứng minh giá trị của mình.

Không có nhiều đánh giá, không có nhiều bài viết, không có nhiều hình ảnh, không có lịch sử giao dịch, không có dữ liệu hành vi và cũng chưa có nhiều nguồn độc lập xác nhận. Vì thiếu bằng chứng, hệ thống đánh giá mức độ chắc chắn thấp hơn và thường ưu tiên những lựa chọn đã được kiểm chứng. Kết quả là địa phương nhỏ ít được đề xuất, mà càng ít được đề xuất thì lại càng khó tạo ra dữ liệu mới. Sự lệch này hiện rõ ngay trong dữ liệu huấn luyện của chính các hệ thống gợi ý. Trên bộ dữ liệu MovieLens vốn là chuẩn mực của ngành, hơn 80% toàn bộ lượt

đánh giá tập trung vào nhóm mục phổ biến.[1] Một thuật toán học từ tập dữ liệu như vậy sẽ học được rất nhiều về nhóm phổ biến và gần như không học được gì về phần đuôi. Sự thiên lệch ấy có trước khi bất kỳ địa phương nào tham gia cuộc chơi.

> ▪ ***Khởi đầu lạnh gồm bốn dạng thiếu hụt dữ liệu khác nhau:***
> • ***Thiếu dữ liệu nhận diện*** *là khi hệ thống không xác định được chính xác điểm đến là thực thể nào hoặc nhầm lẫn với một địa danh khác cùng tên. Cách khắc phục là xây dựng một hồ sơ thực thể thống nhất, với định danh ổn định, tên gọi chính thức và tọa độ chính xác. Nếu được tổ chức tốt, vấn đề này thường có thể giải quyết trong vài tuần.*
> • ***Thiếu dữ liệu chất lượng*** *xảy ra khi điểm đến gần như không có mô tả đáng tin cậy, thiếu nguồn chính thức hoặc thiếu các đánh giá xác thực từ người dùng. Giải pháp là công bố dữ liệu mô tả từ các nguồn có thẩm quyền và từng bước tích lũy đánh giá thật. Quá trình này thường cần vài tháng.*
> • ***Thiếu dữ liệu giao dịch*** *xuất hiện khi hệ thống không có tín hiệu về khả năng đặt chỗ, tình trạng còn chỗ hay lịch sử giao dịch. Muốn khắc phục phải kết nối dữ liệu tồn kho và giao dịch với ít nhất hai nền tảng hoặc hệ thống thông qua giao diện lập trình (API). Đây cũng là quá trình kéo dài trong vài tháng.*
> • ***Thiếu dữ liệu hậu nghiệm*** *là mức khó nhất. Hệ thống không biết khách sau khi trải nghiệm có hài lòng hay không, có quay lại hay giới thiệu cho người khác hay không. Không có giải pháp tức thời cho loại dữ liệu này, nó chỉ được hình thành thông qua quá trình vận hành và tích lũy theo thời gian thực, thường phải mất nhiều năm mới đủ dày để tạo thành lợi thế.*

1. Bộ dữ liệu MovieLens của GroupLens Research, Đại học Minnesota. Trích dẫn chuẩn: F. Maxwell Harper và Joseph A. Konstan, "The MovieLens Datasets: History and Context", ACM Transactions on Interactive Intelligent Systems, tập 5, số 4, bài 19, tháng 12 năm 2015, doi:10.1145/2827872. Phân bố đánh giá tập trung ở nhóm phim phổ biến được chính nhóm tác giả nêu như một giới hạn của bộ dữ liệu.

Trong thực tế, cả bốn dạng thiếu hụt này thường bị quy về một nhận xét đơn giản là *"ít đánh giá"*. Nhưng khi tách chúng ra, có thể thấy ba dạng đầu chủ yếu là bài toán tổ chức dữ liệu và hạ tầng, hoàn toàn có thể cải thiện trong một chu kỳ ngân sách. Chỉ riêng dữ liệu hậu nghiệm là thứ không thể mua ngay bằng tiền, bởi nó phản ánh lịch sử trải nghiệm thực sự của người dùng. Đó là lý do nguồn lực tài chính có thể rút ngắn đáng kể giai đoạn khởi đầu lạnh, nhưng không thể xóa bỏ hoàn toàn nó.

Đây là một nghịch lý cơ bản của nền kinh tế tác nhân. Muốn được AI lựa chọn phải có dữ liệu, nhưng muốn có dữ liệu lại cần được AI lựa chọn trước. Nếu không có một nguồn tín hiệu ban đầu đủ mạnh, vòng lặp này rất khó được phá vỡ. Điều đó giải thích vì sao nhiều địa phương có tài nguyên hấp dẫn nhưng vẫn gần như vô hình trong các hệ thống AI. Hạn chế không nằm ở chất lượng thực tế mà nằm ở việc chưa tích lũy được đủ bằng chứng số để vượt qua giai đoạn khởi đầu. Bài toán của địa phương nhỏ vì thế không phải chỉ là quảng bá nhiều hơn, mà là xây dựng một tập tín hiệu ban đầu đủ tin cậy để thuật toán có cơ sở đưa họ vào tập lựa chọn.

Người ta có thể lập luận rằng vấn đề khởi đầu lạnh có thể phá bằng tiền. Một chiến dịch truyền thông đủ lớn sẽ tạo ra lượt tìm kiếm, lượt ghé thăm và lượt đánh giá, tức là tạo ra chính lượng dữ liệu mà thuật toán đang thiếu. Vậy đây là bài toán ngân sách chứ không phải bài toán cấu trúc. Tuy nhiên, tiền phá được lớp ngoài của khởi đầu lạnh và không phá được lớp trong. Lớp ngoài là khối lượng dữ liệu. Đúng là một chiến dịch đủ lớn sẽ tạo ra khối lượng. Lớp trong là lịch sử xác nhận, tức là chuỗi các lần dữ liệu về địa phương được kiểm chứng đúng qua thời gian. Thứ này chỉ tích lũy được bằng thời gian thực. Một địa phương chi rất nhiều trong sáu tháng sẽ có hồ sơ của một nơi được nhắc đến

nhiều trong sáu tháng. Các hệ thống xếp hạng phân biệt được điều đó với hồ sơ của một nơi được nhắc đến đều đặn trong sáu năm. Có một ngoại lệ đáng ghi nhận. Nếu chiến dịch tạo ra lượng khách thật để lại giao dịch thật cùng đánh giá thật thì tiền đã mua được thời gian chứ không chỉ mua được tiếng ồn. Ngoại lệ ấy đòi hỏi năng lực phục vụ có sẵn, nghĩa là nó chỉ áp dụng cho địa phương thiếu tiếng chứ không áp dụng cho địa phương thiếu hạ tầng.

### 5.4.5. Khi bất bình đẳng tự khuếch đại

Cơ chế tích lũy lợi thế ấy đặt ra một câu hỏi mà mục này muốn tách riêng, vì nó quyết định chính sách nên can thiệp ở đâu. Bất bình đẳng nào giữa các địa phương là hệ quả hợp lý của khác biệt về chất lượng thực, còn bất bình đẳng nào chỉ là kết quả của lợi thế dữ liệu ban đầu. Hai loại này đòi hỏi hai cách xử lý khác nhau. Loại thứ nhất là tín hiệu thị trường đang làm đúng việc của nó. Loại thứ hai là một dạng thất bại thông tin. Can thiệp vào nó không phải trợ cấp cho kém cỏi mà là sửa một sai lệch đo lường.

Vấn đề là hai loại ấy khó tách ra khi nhìn từ ngoài, vì cả hai đều biểu hiện giống nhau là một địa phương ít được đề xuất. Có một phép thử thô nhưng dùng được. Nếu một địa phương có năng lực phục vụ, có sản phẩm và có mức hài lòng của khách đã đến tương đương với nhóm dẫn đầu, mà vẫn không xuất hiện trong tập ứng viên, thì có cơ sở để nghi ngờ rằng một phần đáng kể của khoảng cách đến từ dữ liệu chứ không từ chất lượng. Đó là loại khoảng cách mà một chính sách khả đọc có thể thu hẹp. Ngược lại, nếu mức hài lòng của khách đã đến cũng thấp hơn, thì đầu tư vào tín hiệu số chỉ đưa thêm khách tới một trải nghiệm chưa sẵn sàng. Kết quả sẽ là một chuỗi xác nhận âm.

▪ ***Nếu bạn là Chủ tịch một tỉnh:***

• *Hãy xác định trong vòng một quý ai là đầu mối duy nhất chịu trách nhiệm về dữ liệu mô tả tỉnh trên các hệ thống bên ngoài. Nếu câu trả lời là không có ai, đó là phát hiện quan trọng nhất của quý đó.*

• *Hãy tách ngân sách khả đọc ra khỏi ngân sách quảng bá. Hai khoản này có chu kỳ sinh lợi khác nhau và trộn chung sẽ khiến khoản dài hạn bị cắt trước.*

• *Hãy yêu cầu một báo cáo định kỳ về mức độ hiện diện của tỉnh trên ít nhất bốn loại hệ thống, thay vì một bảng xếp hạng duy nhất.*

## 5.5. TỪ TỐI ƯU CỤC BỘ ĐẾN TẬP TRUNG HỆ THỐNG

Mỗi hệ thống AI được thiết kế để giải quyết bài toán của từng cá nhân, như tìm khách sạn phù hợp nhất, chọn nhà hàng được đánh giá cao nhất, gợi ý điểm tham quan thuận tiện nhất hay xây dựng lịch trình hiệu quả nhất. Ở quy mô của một người dùng, cách tối ưu ấy hoàn toàn hợp lý. AI giúp giảm chi phí tìm kiếm, tiết kiệm thời gian và nâng cao xác suất có một trải nghiệm tốt.

Đối với các địa phương, đây là một thay đổi mang tính cấu trúc. Bài toán không còn đơn thuần là thu hút thêm khách, mà là hiểu cơ chế phân phối của các hệ thống AI. Nếu không có đủ tín hiệu để bước vào tập lựa chọn ban đầu, địa phương gần như không có cơ hội được thuật toán đề xuất. Ngược lại, một khi đã liên tục được lựa chọn, lợi thế sẽ tự khuếch đại qua dữ liệu, đánh giá, mức độ phổ biến và niềm tin của máy.

### 5.5.1. Tối ưu cục bộ

Phần lớn hệ thống đề xuất hướng tới một mục tiêu cục bộ tương đối rõ: đưa ra phương án phù hợp cho một người dùng trong một ngữ cảnh cụ thể. Khi một du khách hỏi mình nên ở đâu, hệ thống tìm khách sạn phù hợp với ngân sách, vị trí, đánh giá và sở thích của người đó. Khi một gia đình hỏi nên tham quan những nơi nào trong hai ngày, hệ thống chọn những địa điểm có xác suất mang lại trải nghiệm hài lòng cao nhất. Mục tiêu ấy trên thực tế thường là một hàm nhiều mục tiêu, không phải một đại lượng duy nhất.

Đây được gọi là tối ưu cục bộ (local optimization). Thuật toán không cố gắng tối ưu cho toàn bộ ngành du lịch, cho một thành phố hay cho nền kinh tế của một quốc gia. Nó chỉ tối ưu cho một

bài toán rất nhỏ: làm sao để người dùng trước mắt có quyết định tốt nhất. Để đạt mục tiêu đó, AI sẽ ưu tiên những lựa chọn có nhiều bằng chứng nhất. Khách sạn có hàng chục nghìn đánh giá sẽ được ưu tiên hơn khách sạn mới mở. Nhà hàng được nhiều người xác nhận sẽ đáng tin cậy hơn một quán ít dữ liệu. Điểm tham quan có hình ảnh, bài viết và lịch sử tương tác dày đặc sẽ có xác suất được đề xuất cao hơn những nơi gần như vô hình trên không gian số.

Ở cấp độ cá nhân, đây là một chiến lược hoàn toàn hợp lý. Người dùng nhận được lựa chọn an toàn hơn, ít rủi ro hơn và phù hợp với nhu cầu của mình. Thuật toán cũng giảm xác suất đề xuất sai, từ đó tăng mức độ hài lòng và củng cố niềm tin vào hệ thống. Đó là giới hạn căn bản của tối ưu cục bộ. Thuật toán nhìn thấy từng bài toán riêng lẻ, nhưng không nhìn thấy bức tranh tổng thể mà hàng triệu quyết định riêng lẻ đang cùng nhau tạo ra.

### 5.5.2. Thất bại ở quy mô hệ thống

Khi hàng triệu quyết định tối ưu cục bộ được cộng lại, toàn bộ hệ thống có thể đi đến một kết quả mà không một cá nhân nào mong muốn. Trong khoa học hệ thống, hiện tượng này được gọi là thất bại ở quy mô hệ thống (system-level failure), mỗi thành phần đều hoạt động đúng chức năng của mình, nhưng kết quả chung lại kém hiệu quả hoặc thậm chí gây tổn hại cho toàn bộ hệ thống.

Đây không phải lỗi lập trình hay sự thiên vị có chủ ý. Đó là điều thường xảy ra khi nhiều hệ thống cùng tối ưu theo một mục tiêu tương tự mà không có cơ chế điều phối ở cấp toàn hệ thống. Điều hợp lý với từng người dùng có thể trở thành điều phi lý với xã hội. Đó là dạng thất bại phối hợp, không phải dạng thất bại kỹ thuật.

Một chuỗi quyết định tối ưu ở cấp từng tác nhân có thể tạo ra một kết quả kém tối ưu ở cấp toàn hệ thống. Khi nhiều tác nhân cùng học từ những nguồn dữ liệu tương tự, chúng có xu hướng lựa chọn những phương án có nhiều bằng chứng và tín hiệu xác nhận nhất. Kết quả là các quyết định dần hội tụ vào một nhóm rất nhỏ các điểm đến, nhà cung cấp hoặc sản phẩm. Khi dòng khách tiếp tục đổ về nhóm này, lượng giao dịch và dữ liệu tương tác của chúng lại tăng lên nhanh hơn phần còn lại của thị trường. Điều đó tạo thêm bằng chứng cho các hệ thống xếp hạng, khiến xác suất được đề xuất của chính nhóm ấy tiếp tục tăng. Qua thời gian, lợi thế ban đầu được khuếch đại thành một vòng lặp tự củng cố. Một số rất ít điểm đến ngày càng được coi là lựa chọn "chắc chắn", trong khi nhiều nơi khác dần biến mất khỏi tập phương án mà tác nhân cân nhắc. Hệ quả là toàn hệ thống rơi vào trạng thái mất cân bằng, một phía phải đối mặt với quá tải, còn phía kia trở nên gần như vô hình, dù chất lượng thực tế có thể không chênh lệch tương ứng. Nghịch lý nằm ở chỗ không có tác nhân nào hành động sai. Mỗi quyết định riêng lẻ đều là tối ưu theo thông tin sẵn có. Nhưng khi hàng triệu quyết định giống nhau cùng vận hành trên một cơ chế xếp hạng, tối ưu cục bộ có thể tích lũy thành khủng hoảng ở quy mô hệ thống. Không có khâu nào trong chuỗi này là một lỗi. Khủng hoảng nằm ở tổng của các khâu đúng.

### 5.5.3. Hội tụ đề xuất ở cấp hệ thống

Trong kinh tế học và tâm lý học, hành vi bầy đàn (herd behavior) mô tả hiện tượng con người có xu hướng làm theo lựa chọn của số đông, ngay cả khi họ không có đầy đủ thông tin. Người ta vào một nhà hàng vì thấy đông khách, mua một cổ phiếu vì thấy nhiều người mua, hay chọn một điểm đến vì liên tục nhìn thấy

nó xuất hiện trên mạng xã hội. Niềm tin của đám đông trở thành tín hiệu thay thế cho tri thức cá nhân. Điều đáng chú ý là trong kỷ nguyên AI, không chỉ con người mà chính các hệ thống máy cũng có thể tạo ra hành vi bầy đàn.

AI không bắt chước nhau theo nghĩa trực tiếp. Mỗi mô hình được phát triển bởi một tổ chức khác nhau, sử dụng kiến trúc khác nhau và phục vụ những nhóm người dùng khác nhau. Tuy nhiên, chúng thường được huấn luyện từ những kho dữ liệu có nhiều điểm tương đồng, cùng coi trọng các tín hiệu như mức độ phổ biến, số lượng đánh giá, độ tin cậy của nguồn và lịch sử hành vi của người dùng. Khi cùng tối ưu theo những tiêu chí đó, các hệ thống khác nhau rất dễ hội tụ về những kết quả giống nhau.

Vì vậy chữ bầy đàn ở đây nên được hiểu là một phép tương tự, không phải một mô tả kỹ thuật. Sự hội tụ giữa các hệ thống có ít nhất năm nguồn. Không nguồn nào đòi hỏi hệ thống này quan sát hệ thống kia. Thứ nhất là cùng dữ liệu nguồn, vì phần lớn đều truy hồi từ những kho công khai giống nhau. Thứ hai là cùng chỉ báo phổ biến, vì lượng đánh giá và lượng nhắc đến là những đại lượng dễ thu thập nhất. Thứ ba là cùng tiêu chí giảm rủi ro, vì nhiều hệ thống thương mại bị phạt rõ rệt cho một gợi ý sai, trong khi phần thưởng cho khả năng khám phá lựa chọn mới thường khó đo hơn và đến chậm hơn. Thứ tư là cùng nhà cung cấp mô hình hoặc công cụ, vì số mô hình nền tảng ít hơn nhiều số sản phẩm dùng chúng. Thứ năm là cùng phản hồi hành vi của người dùng, vì người dùng của các hệ thống khác nhau vẫn là cùng một dân số. Gọi hiện tượng này là hội tụ đề xuất ở cấp hệ thống thì chính xác hơn là gọi nó là bắt chước.

Một điểm đến đã nổi tiếng sẽ đồng thời xuất hiện trong Google Maps, Booking, TripAdvisor, ChatGPT, Gemini hay nhiều trợ lý AI

khác. Không phải vì các hệ thống sao chép lẫn nhau, mà vì tất cả đều đi đến cùng một kết luận từ những bằng chứng tương tự. Sự đồng thuận của máy vì thế trở thành một dạng bầy đàn thuật toán.

Hiện tượng này tạo ra một lớp củng cố mới. Trước đây, con người bị ảnh hưởng bởi đám đông. Nay con người còn bị ảnh hưởng bởi các hệ thống AI, trong khi các hệ thống AI lại cùng chịu ảnh hưởng từ những tín hiệu phổ biến do chính đám đông tạo ra. Một vòng lặp khép kín hình thành: đám đông tạo dữ liệu, dữ liệu dẫn AI đến cùng một lựa chọn, AI lại dẫn thêm đám đông quay trở lại đúng lựa chọn đó. Do đó, trong nền kinh tế tác nhân, cạnh tranh không chỉ là thay đổi nhận thức của con người. Quan trọng hơn, đó là phá vỡ hành vi bầy đàn của thuật toán bằng những tín hiệu mới đủ mạnh để các hệ thống AI bắt đầu đưa một địa phương vào tập lựa chọn của mình.

Khi nhiều hệ thống độc lập cùng tối ưu một mục tiêu tương tự trên những nguồn dữ liệu tương tự, kết quả của chúng có xu hướng hội tụ về cùng một nhóm nhỏ phương án dù chúng không trao đổi thông tin với nhau. Sự hội tụ ở đây không cần đến sự bắt chước, nó chỉ cần sự giống nhau của dữ liệu, của chỉ báo và của tiêu chí giảm rủi ro.

### 5.5.4. Cân bằng tự củng cố không tối ưu xã hội

Trong lý thuyết trò chơi, cân bằng Nash mô tả một trạng thái mà mỗi tác nhân đều đang lựa chọn phương án tốt nhất có thể, xét trên hành vi của những tác nhân còn lại.[1] Điều quan trọng

1. John F. Nash, "Equilibrium Points in N-Person Games", Proceedings of the National Academy of Sciences, tập 36, số 1, 1950, trang 48 tới 49.

là cân bằng Nash không đồng nghĩa với trạng thái tốt nhất cho toàn xã hội. Cần nói ngay rằng chương này dùng khái niệm ấy như một phép tương tự chứ không phải một mô hình hình thức. Các hệ thống đề xuất không nhất thiết là tác nhân chiến lược theo nghĩa của lý thuyết trò chơi. Chúng có thể chỉ là những bộ tối ưu độc lập, không hề mô hình hóa phản ứng của đối thủ.

Hệ thống AI trong du lịch đang tiến gần tới một trạng thái như vậy. Mỗi trợ lý AI đều đề xuất những điểm đến có xác suất làm hài lòng người dùng cao nhất. Mỗi du khách đều chấp nhận đề xuất đó vì đây là lựa chọn an toàn và đáng tin cậy nhất. Mỗi doanh nghiệp nổi tiếng tiếp tục được hưởng lợi vì đã có nhiều dữ liệu và đánh giá. Không một tác nhân riêng lẻ nào có lý do để thay đổi hành vi của mình.

> ▪ ***Bài toán phối hợp phát biểu cho đúng:***
> • *Người chơi là các nền tảng vận hành hệ thống đề xuất, không phải từng tác nhân AI.*
> • *Chiến lược là mức đa dạng mà nền tảng chủ động chèn vào kết quả gợi ý.*
> • *Hàm lợi ích là tỷ lệ người dùng chấp nhận phương án được đưa ra, cùng mức giữ chân người dùng trong dài hạn.*
> • *Không nền tảng nào có lợi khi đơn phương đa dạng hóa mạnh, vì làm vậy làm giảm tỷ lệ chấp nhận của riêng người dùng mình, trong khi lợi ích về phân tán dòng khách thì chia cho cả hệ thống.*
> • *Nếu tất cả cùng đa dạng hóa, kết quả chung có thể tốt hơn cho hệ thống mà không bên nào mất lợi thế tương đối. Đó là hình dạng của một bài toán phối hợp, không phải một mô hình trò chơi đã được hình thức hóa đầy đủ.*

Nhưng khi tất cả cùng hành động như vậy, kết quả chung lại không tối ưu. Một số địa phương phải đối mặt với quá tải du lịch, giá cả leo thang và suy giảm chất lượng sống. Trong khi đó, nhiều địa phương khác có đủ năng lực tiếp nhận khách lại gần như không được lựa chọn. Nguồn lực xã hội bị phân bổ mất cân đối, không phải vì thiếu tài nguyên mà vì thuật toán liên tục hội tụ vào cùng một nhóm lựa chọn.

Đối với quản trị địa phương, đây là một nhận thức quan trọng. Nếu các hệ thống chỉ tối ưu theo xác suất thành công của từng phiên truy vấn, dòng khách có xu hướng dồn về một số ít địa điểm. Muốn có một hệ sinh thái du lịch bền vững, hàm mục tiêu của nền tảng phải tính thêm những biến rộng hơn như năng lực tiếp nhận, tính đa dạng của trải nghiệm và mức phân bổ dòng khách. Vấn đề là những biến ấy phần lớn nằm ngoài dữ liệu mà nền tảng đang có. Đó chính là chỗ địa phương có thể can thiệp.

Bạn có thể cho rằng so sánh với cân bằng Nash là khập khiễng. Cân bằng Nash mô tả những tác nhân có chiến lược, biết đối thủ tồn tại và tính toán phản ứng của đối thủ. Các hệ thống đề xuất không làm gì trong số đó. Chúng chỉ dự đoán sở thích của một người dùng.

Điều được mượn ở đây không phải cơ chế suy luận chiến lược mà là cấu trúc kết cục: một trạng thái trong đó không tác nhân nào có động cơ đơn phương thay đổi hành vi, nhưng tổng thể lại kém hơn một trạng thái khác có thể đạt được nếu phối hợp. Cấu trúc ấy xuất hiện ở đây đầy đủ. Không một hệ thống đề xuất nào có lợi khi đơn phương giảm tần suất gợi ý điểm đến an toàn nhất, vì làm vậy chỉ khiến người dùng của riêng nó nhận câu trả lời kém hơn. Để chặt chẽ, cần phát biểu lại: đây không phải cân bằng Nash theo nghĩa lý thuyết trò chơi cổ điển, mà là một trạng thái khóa cứng do trùng hợp mục tiêu. Trạng thái ấy có cùng tính chất khó thoát.

- ***Về ứng xử với sự tập trung:***

• *Không được xem sự tập trung dòng khách là bằng chứng của thành công trong quảng bá trước khi kiểm tra sức chịu tải của điểm đến đang nhận dòng khách ấy.*

• *Không được xử lý quá tải chỉ bằng công cụ ở điểm đến đang quá tải. Một phần lời giải luôn nằm ở những điểm đến chưa được đưa vào tập lựa chọn.*

• *Không được đặt mục tiêu tăng trưởng lượt khách cho toàn tỉnh mà không đồng thời đặt mục tiêu phân bố lượt khách theo không gian và theo mùa.*

### 5.5.5. Nghịch lý của nền kinh tế tác nhân: Máy càng thông minh, đám đông càng tập trung

Trước khi phát biểu nghịch lý, cần nói rõ điều kiện của nó. Máy thông minh hơn không tất yếu làm các đề xuất hội tụ. Nếu một hệ thống khai thác tốt sở thích ngách và được thưởng cho khả năng khám phá, nó hoàn toàn có thể làm dòng khách phân tán hơn, phát hiện những nơi phù hợp với từng nhóm nhỏ và cá nhân hóa sâu tới mức phá vỡ khuôn mẫu chung. Nghịch lý xuất hiện trong điều kiện ngược lại. Điều kiện ấy gồm năm thành phần đã nêu ở phần trước, nhiều hệ thống cùng học từ dữ liệu tương tự, cùng ưu tiên tín hiệu đã được kiểm chứng, cùng để chỉ báo phổ biến chiếm trọng số lớn, cùng chịu mức phạt rõ rệt cho một gợi ý sai trong khi phần thưởng cho khả năng khám phá thì khó đo hơn, cùng thiếu dữ liệu về sức chứa cùng lựa chọn thay thế. Phần còn lại của mục này nói về điều kiện ấy, không nói về trí thông minh của máy nói chung.

Trực giác thông thường cho rằng AI càng thông minh thì các quyết định sẽ càng đa dạng và hiệu quả hơn. Một hệ thống hiểu sâu về sở thích cá nhân, biết phân tích hàng triệu lựa chọn và có

khả năng lập kế hoạch phức tạp dường như sẽ giúp mỗi người khám phá những trải nghiệm khác biệt thay vì đi theo số đông.

Càng thông minh, AI càng giỏi xác định đâu là phương án có xác suất thành công cao nhất. Nếu cùng một tập dữ liệu đều chỉ ra rằng một khách sạn có dịch vụ ổn định hơn, một nhà hàng có tỷ lệ hài lòng cao hơn hay một điểm đến có nhiều bằng chứng đáng tin cậy hơn, thì những hệ thống AI khác nhau sẽ ngày càng dễ đi đến cùng một kết luận. Khả năng suy luận tốt hơn không làm các đề xuất phân tán hơn, mà nhiều khi còn khiến chúng hội tụ mạnh hơn. Đây là nghịch lý của nền kinh tế tác nhân. Mỗi AI được tạo ra để giảm bất định cho người sử dụng. Nhưng khi hàng trăm triệu AI cùng loại bỏ bất định theo cùng một logic, chúng lại vô tình làm giảm sự đa dạng của toàn hệ thống. Những lựa chọn ít dữ liệu, ít bằng chứng hoặc chưa được xác nhận rộng rãi sẽ ngày càng khó được đề xuất, mặc dù chúng có thể mang lại trải nghiệm rất tốt.

Vì vậy, mục tiêu của quản trị địa phương trong nền kinh tế tác nhân không thể chỉ là xây dựng một nơi đáng đến. Mục tiêu còn là xây dựng một nơi mà các hệ thống AI có đủ cơ sở để tự tin giới thiệu. Nếu không, nghịch lý sẽ tiếp tục lặp lại, máy càng thông minh, các đề xuất càng hội tụ. Đề xuất càng hội tụ, dòng khách càng tập trung. Và khoảng cách giữa những địa phương được chọn với những địa phương còn lại sẽ ngày càng lớn.

Đó chính là thách thức trung tâm của chương này. Trong nền kinh tế tác nhân, năng lực cạnh tranh không chỉ được quyết định bởi chất lượng trải nghiệm, mà còn bởi khả năng tham gia vào quá trình ra quyết định của máy. Chỉ khi hiểu được cơ chế đó, địa phương mới có thể vừa hiện diện trong tập lựa chọn của AI, vừa tránh rơi vào vòng xoáy tập trung quá mức đang định hình lại ngành du lịch toàn cầu. Khi nhiều máy cùng định nghĩa tốt nhất theo một cách, thế giới sẽ có ít câu trả lời khác nhau hơn.

PHẦN C

# KHI ĐỀ XUẤT TRỞ THÀNH ÁP LỰC LÊN NƠI CHỐN

## 5.6. QUÁ TẢI DU LỊCH ĐƯỢC THUẬT TOÁN KHUẾCH ĐẠI

Trong nhiều năm, quá tải du lịch thường được lý giải bằng sự phát triển của hàng không giá rẻ, mạng xã hội hay các nền tảng đặt phòng trực tuyến. Những yếu tố đó quả thực đã làm giảm chi phí di chuyển và lan truyền hình ảnh điểm đến với tốc độ chưa từng có. Tuy nhiên, trong nền kinh tế tác nhân, bên cạnh những động lực ấy xuất hiện thêm một lớp định tuyến mới. Các hệ thống AI và tác nhân số có khả năng tự động lựa chọn, đề xuất và trong một số trường hợp thực hiện giao dịch thay con người. Chúng không thay thế các nguyên nhân truyền thống, nhưng có thể làm dòng khách hình thành nhanh hơn, hội tụ vào những không gian hẹp hơn và khó điều tiết hơn bằng bộ công cụ hiện có.

Khi hàng triệu AI cùng tối ưu cho mục tiêu "đưa người dùng đến lựa chọn tốt nhất", chúng thường hội tụ vào một nhóm rất nhỏ các điểm đến có xác suất thành công cao nhất. Mỗi đề xuất tạo thêm lượt khách, mỗi lượt khách tạo thêm dữ liệu, đánh giá và bằng chứng, còn dữ liệu mới lại củng cố vị trí của điểm đến trong các lần đề xuất tiếp theo. Một vòng phản hồi dương hình thành, khiến nhu cầu tập trung ngày càng mạnh vào cùng một số địa điểm.

Quy mô của dòng chảy này cần được đặt trước mắt. Theo UN Tourism, thế giới ghi nhận khoảng 1,52 tỷ lượt khách quốc tế trong năm 2025, tăng 4% so với năm 2024 và là mức cao nhất thời kỳ hậu đại dịch.[1] Riêng châu Âu đón 793 triệu lượt, tức

1. UN Tourism, World Tourism Barometer, tháng 1 năm 2026. Số liệu sơ bộ, thường được điều chỉnh trong các bản công bố sau.

khoảng 52% toàn cầu, trên một khu vực chiếm chưa tới 7% diện tích đất liền thế giới. Sự tập trung ấy có từ trước AI. Điều đang thay đổi là tốc độ mà nó tự siết chặt thêm.

### 5.6.1. Quá tải du lịch trong điều kiện có tác nhân

Quá tải du lịch từ lâu đã được xem là hệ quả của sự gia tăng khách du lịch, chi phí di chuyển giảm và hiệu ứng lan truyền trên mạng xã hội. Trong kỷ nguyên AI, hiện tượng này có thêm một cơ chế khuếch đại. Các hệ thống ra quyết định tự động có thể đồng thời định tuyến nhiều người dùng về cùng một số lựa chọn. Có thể gọi đây là quá tải du lịch trong điều kiện có tác nhân AI.

Khác với du khách tự tìm kiếm và so sánh hàng chục lựa chọn, AI thường rút gọn không gian quyết định xuống một danh sách rất ngắn. Một tác nhân số có thể chỉ đề xuất ba khách sạn, hai nhà hàng hay một lịch trình tối ưu. Nếu hàng triệu tác nhân cùng sử dụng những nguồn dữ liệu, tiêu chí xếp hạng và mô hình suy luận tương tự, xác suất chúng hội tụ vào cùng một tập điểm đến sẽ tăng rất mạnh. Hậu quả là những địa điểm vốn đã nổi tiếng càng nhanh chóng rơi vào tình trạng quá tải, trong khi nhiều điểm đến khác vẫn gần như vô hình đối với hệ thống AI. Khoảng cách giữa "được đề xuất" và "không được đề xuất" vì thế ngày càng lớn. Quá tải du lịch không còn đơn thuần là vấn đề của tổng lượng khách, mà ngày càng bị khuếch đại bởi cách các thuật toán phân bổ sự chú ý và dòng nhu cầu. Trong nền kinh tế tác nhân, vì vậy, quản trị quá tải du lịch cũng bao gồm việc can thiệp vào dữ liệu đầu vào của hệ thống đề xuất. Nếu chỉ xử lý ở khâu khách đã đến, địa phương sẽ luôn đi sau một bước so với nơi quyết định thực sự được hình thành.

Quá tải du lịch có trước AI rất lâu. Venice đã quá tải từ thập niên 1980, khi chưa hề có hệ thống đề xuất nào. Quy trách nhiệm cho thuật toán là quy sai địa chỉ, nguyên nhân thật vẫn là hàng không giá rẻ, tàu biển du lịch và thu nhập tăng. Nhận định phản bác này thắng ở phần nguyên nhân và thua ở phần cơ chế, nên kết luận của mục này phải được thu hẹp lại cho đúng. Đúng là thuật toán không tạo ra quá tải du lịch. Quá tải là hiện tượng cũ. Các nguyên nhân được nêu đều thật, đều lớn hơn tác động của hệ thống đề xuất xét về khối lượng. Điều thuật toán làm thay đổi là ba thứ khác. Thứ nhất là tốc độ, vì một điểm đến có thể chuyển từ ít người biết sang quá tải trong vài tháng thay vì vài thập niên. Thứ hai là độ tập trung trong không gian, vì các hệ thống hội tụ vào cùng những tọa độ chứ không chỉ cùng những thành phố. Thứ ba là khả năng điều tiết, vì các công cụ cũ tác động vào giá và vào hạ tầng, trong khi quyết định đã hình thành trước đó ở một khâu mà chính quyền chưa có công cụ nào. Vậy phát biểu đúng không phải là thuật toán gây ra quá tải, mà là thuật toán làm quá tải xảy ra nhanh hơn, tập trung hơn và khó điều tiết hơn bằng bộ công cụ hiện có.

### 5.6.2. Cơ chế tác động lên cư dân

Quá tải du lịch thường được đo bằng số lượng du khách, nhưng những người chịu tác động lớn nhất lại là cư dân địa phương. Khi AI liên tục điều hướng dòng khách đến cùng một khu vực, áp lực không chỉ dừng ở ngành du lịch mà lan sang toàn bộ đời sống đô thị.

Trước hết là nhà ở. Khi một địa phương liên tục được AI đề xuất, nhu cầu lưu trú ngắn hạn tăng mạnh. Nhiều căn hộ chuyển từ cho thuê dài hạn sang phục vụ khách du lịch, làm giảm nguồn cung cho

cư dân. Giá thuê và giá bất động sản tăng nhanh hơn thu nhập của người dân, khiến nhiều người buộc phải rời khỏi chính nơi mình sinh ra. Hiện tượng này không còn là hệ quả của thị trường đơn thuần mà được khuếch đại bởi dòng cầu do thuật toán phân phối.

Tiếp theo là giao thông. Khi hàng nghìn AI cùng đề xuất một lịch trình tương tự, dòng khách có xu hướng tập trung vào cùng địa điểm và cùng khung giờ. Ùn tắc không chỉ xuất hiện trên các tuyến đường mà còn ở bãi đỗ xe, phương tiện công cộng và các không gian công cộng. Thành phố được tối ưu cho du khách nhưng lại trở nên kém thuận tiện đối với cư dân hằng ngày.

Cuối cùng là môi trường. Lượng khách vượt quá sức chịu tải làm gia tăng rác thải, tiêu thụ nước, năng lượng và phát thải. Các hệ sinh thái tự nhiên, di sản văn hóa và không gian công cộng chịu áp lực ngày càng lớn. Điều nghịch lý là chính những giá trị khiến địa phương được AI ưu tiên đề xuất lại bị suy giảm bởi dòng khách mà AI góp phần định tuyến và khuếch đại.

Ba nhóm nhà ở, giao thông và môi trường chỉ mô tả bề mặt của tác động. Dưới góc nhìn kinh tế học địa phương, có thể tách ra năm cơ chế cụ thể. Thứ nhất, nhu cầu lưu trú ngắn hạn tăng làm lệch giá thuê nhà ở, kéo theo chuyển đổi công năng từ nhà ở sang lưu trú. Thứ hai, dòng khách tập trung vào cùng những khung giờ làm quá tải một hạ tầng vốn được thiết kế theo nhịp sinh hoạt của cư dân, nên vấn đề không phải tổng lượng mà là hình dạng của đường phân bố theo giờ. Thứ ba, giá thuê mặt bằng thương mại tăng làm thay đổi cơ cấu cửa hàng. Những dịch vụ có biên lợi nhuận thấp nhưng thiết yếu với cư dân bị đẩy ra trước tiên. Thứ tư, dịch vụ thiết yếu bị thay dần bằng dịch vụ cho khách, khiến cư dân phải đi xa hơn cho những việc trước đây làm ở đầu phố. Thứ năm, phần lớn chi phí ngoại ứng của bốn cơ chế

trên không được phản ánh trong giá của chuyến đi, nên người ra quyết định không nhìn thấy chi phí mình đang tạo ra. Vì vậy, quá tải du lịch không chỉ là vấn đề của ngành du lịch. Đó là bài toán về chất lượng sống của cư dân và năng lực duy trì sự bền vững của địa phương trong kỷ nguyên thuật toán.

> ▪ ***Nếu bạn là giám đốc một Sở du lịch:***
>
> • *Hãy dựng một danh mục điểm đến thay thế có thật, nghĩa là có chỗ đỗ xe, có nhà vệ sinh, có người trực và có giờ mở cửa ổn định. Một điểm đến thay thế không đủ điều kiện phục vụ sẽ bị hệ thống loại ngay sau lần thất bại đầu tiên và rất khó quay lại.*
> • *Hãy công bố dữ liệu sức chịu tải theo khung giờ, không chỉ theo ngày. Đây là loại dữ liệu mà hệ thống đề xuất có thể dùng được, khác với thông cáo kêu gọi giãn khách.*
> • *Hãy đo thành công bằng tỷ lệ phân bố khách giữa các điểm trong tỉnh, không chỉ bằng tổng lượt khách.*

### 5.6.3. Suy giảm trải nghiệm và độ trễ của phản hồi thuật toán

Mục tiêu của AI là giúp người dùng có trải nghiệm tốt hơn. Nhưng khi hàng triệu AI cùng tối ưu theo một logic giống nhau, chính trải nghiệm ấy lại có nguy cơ suy giảm. Dấu hiệu dễ nhận thấy nhất là quá đông. Những địa điểm từng hấp dẫn vì sự yên bình, tính nguyên bản hay không gian đặc trưng trở thành nơi xếp hàng, chen chúc và quá tải. Thời gian chờ tăng lên, khả năng tiếp cận giảm xuống và phần lớn thời gian của du khách bị tiêu hao vào việc xử lý đám đông thay vì tận hưởng trải nghiệm.

Xa hơn, địa phương bắt đầu mất bản sắc. Để phục vụ lượng khách ngày càng lớn, nhiều không gian được chuẩn hóa theo

những mô hình đã chứng minh hiệu quả ở nơi khác. Nhà hàng, cửa hàng lưu niệm, quán cà phê hay dịch vụ giải trí ngày càng giống nhau. Những yếu tố độc đáo của địa phương dần bị thay thế bởi các sản phẩm dễ tiêu dùng và dễ nhân rộng. Kết quả cuối cùng là chất lượng trải nghiệm giảm sút. AI từng đề xuất địa phương vì những đánh giá tích cực trong quá khứ, nhưng khi lượng khách vượt quá sức chứa, chính trải nghiệm thực tế lại kém đi. Các đánh giá mới bắt đầu phản ánh sự đông đúc, chi phí tăng cao và dịch vụ quá tải. Một vòng phản hồi ngược xuất hiện, AI góp phần khuếch đại thành công ban đầu, nhưng cũng vô tình tạo ra điều kiện làm suy giảm chính thành công đó.

Đây là chỗ chương này khác với phần lớn tài liệu về quá tải du lịch, nên cần đặt câu hỏi cho đúng. Nếu hệ thống đề xuất một nơi vì trải nghiệm ở đó tốt, trong khi chính lượng khách do nó dẫn tới làm trải nghiệm ấy xấu đi, thì hệ thống sẽ nhận ra sự xấu đi ấy nhanh đến đâu, qua dữ liệu nào. Câu trả lời quyết định biên độ dao động của cả hệ thống. Nếu tín hiệu phản hồi đến nhanh, dòng khách được điều chỉnh sớm và biên độ nhỏ. Nếu nó đến chậm, dòng khách tiếp tục dồn về trong suốt khoảng trễ, rồi bị rút đi đột ngột khi điểm đánh giá cuối cùng cũng sụt, gây thiệt hại hai lần cho địa phương.

Có cơ sở để cho rằng độ trễ này khá dài. Điểm đánh giá tổng của một địa điểm nổi tiếng được tính trên một khối lượng đánh giá tích lũy nhiều năm, nên vài trăm đánh giá xấu mới gần như không dịch chuyển được con số. Nhận xét dạng văn bản phản ánh sự đông đúc sớm hơn nhiều so với điểm số, nhưng không phải hệ thống nào cũng đọc và không phải hệ thống nào cũng gán trọng số cho phần ấy. Trong khi đó, nhóm tín hiệu phản ánh trực tiếp nhất tình trạng quá tải gần thời gian thực, tức dữ liệu về mật độ và sức chịu tải theo khung giờ, thường chưa tồn tại

dưới dạng máy đọc được vì không ai công bố. Đó là một khoảng trống dữ liệu. Nó cũng là điểm can thiệp cụ thể nhất mà một địa phương có.

Điều này cho thấy trải nghiệm không phải là một nguồn lực vô hạn. Nếu chỉ tối ưu lượng khách mà không quản trị sức chứa và chất lượng, phần lớn hệ thống đề xuất cuối cùng sẽ tự làm suy yếu giá trị mà chúng đang cố gắng tối đa hóa.

## 5.7. ĐIỀU TIẾT Ở ĐÂU TRONG CHUỖI QUYẾT ĐỊNH

Trong nhiều thập niên, các điểm đến đối mặt với quá tải du lịch thường dùng những công cụ kinh tế quen thuộc như tăng giá vé, thu phí vào cửa, giới hạn số lượng khách hay áp dụng đặt chỗ trước. Những biện pháp này dựa trên giả định rằng du khách là người trực tiếp đưa ra quyết định và sẽ điều chỉnh hành vi khi chi phí tăng lên. Trong nền kinh tế tác nhân, giả định ấy cần được phát biểu lại cho chính xác hơn.

Khi tác nhân AI trở thành người lập kế hoạch, một phần quan trọng của quyết định được hình thành trong giai đoạn dựng tập ứng viên. Hệ thống thu hẹp không gian điểm đến, phác lịch trình, so sánh phương án và loại bỏ hàng nghìn lựa chọn khác trước khi trình bày kết quả. Giá vẫn có thể tham gia vào giai đoạn ấy, thường dưới dạng ngân sách tổng của chuyến đi. Nhưng một khoản phí vào cửa hay một mức giá vé đơn lẻ chỉ chiếm một phần nhỏ trong ngân sách ấy, nên trọng số của nó trong hàm lựa chọn thường thấp. Nếu mức tăng chưa đủ lớn để làm đổi kết quả tính toán, dòng khách vẫn tiếp tục đổ về cùng một địa điểm.

Cần nói rõ để tránh một ngộ nhận thường gặp. Tác nhân có thể đưa giá vào quá trình lựa chọn ngay từ đầu. Trên thực tế phần lớn hệ thống đặt chỗ đều dùng giá làm biến lọc sớm. Vấn đề không nằm ở chỗ giá xuất hiện muộn về mặt kỹ thuật. Vấn đề là giá chỉ điều chỉnh mức hấp dẫn tương đối của những phương án đã tồn tại trong tập ứng viên. Một khoản phí bổ sung nhỏ so với tổng chi phí chuyến đi có thể không đủ làm thay đổi thứ hạng của một điểm đến nổi tiếng, đặc biệt với nhóm khách ít nhạy giá. Quan trọng hơn, tăng giá tại một điểm đến không tự tạo ra dữ liệu, niềm tin hay khả năng giao dịch cho những địa điểm thay thế đang vắng mặt trong tập ứng viên. Nếu phương án thay thế

chưa có mặt, làm Venice đắt hơn cũng không đưa được một nơi khác vào tầm nhìn của hệ thống.

Venice và Kyoto là hai ví dụ đáng đọc kỹ về giới hạn của bộ công cụ điều tiết truyền thống. Cả hai đều đã thử nhiều cách quản lý nhu cầu, từ thu phí, kiểm soát dòng khách đến phân luồng theo thời gian. Cả hai đều thu được những kết quả thật mà phần sau sẽ ghi nhận. Điều hai trường hợp này cho thấy là bài toán đang dịch chuyển từ quản lý người đến sang can thiệp vào chuỗi quyết định trước đó.

### 5.7.1. Venice và bài toán phí vào cửa

Khái niệm Định giá tắc nghẽn (congestion pricing). Công cụ điều tiết cầu bằng cách tính thêm phí vào những thời điểm hoặc khu vực có mức sử dụng cao, nhằm dịch chuyển một phần nhu cầu sang thời điểm hoặc khu vực ít tải hơn.

Venice là một trong những thành phố đầu tiên trên thế giới áp dụng phí vào cửa đối với khách du lịch trong ngày nhằm giảm tình trạng quá tải. Về bản chất, đây là một hình thức định giá tắc nghẽn khi nhu cầu vượt quá khả năng chịu tải của hạ tầng và cộng đồng địa phương, chi phí tham gia sẽ được nâng lên để làm giảm lượng người sử dụng.

Tuy nhiên, trong nền kinh tế tác nhân, hiệu quả của công cụ này trở nên hạn chế hơn. Một tác nhân AI thường tối ưu toàn bộ hành trình dựa trên hàng chục tiêu chí, qua mức độ hấp dẫn, thời gian di chuyển, đánh giá của cộng đồng, thời tiết, lịch mở cửa, khả năng đặt chỗ và tổng chi phí chuyến đi. Trong mô hình đó, phí vào cửa chỉ là một biến số nhỏ trong hàm tối ưu.

Kết quả hai năm đầu chưa cho phép tách riêng tác động của các hệ thống đề xuất, nhưng cung cấp một phép thử thực tế về giới hạn của mức phí hiện hành. Năm 2024, Venice áp dụng phí vào cửa trong 29 ngày với mức 5 euro, thu về khoảng 2,4 triệu euro từ 485.062 lượt trả phí. Năm 2025, thành phố mở rộng lên 54 ngày, thêm cả thứ sáu. Thành phố cũng đưa vào cơ chế hai mức giá là 5 euro nếu đặt trước và 10 euro nếu đặt sát ngày. Kết quả là hơn 720.000 lượt trả phí và 5.421.425 euro thu vào, gần gấp đôi năm trước.[1] Có ngày cao điểm vượt 25.000 lượt trả phí. Năm 2026, lịch áp dụng tiếp tục được mở rộng lên 60 ngày.

Nếu Venice vẫn được đánh giá là điểm đến có giá trị vượt trội, AI hoàn toàn có thể kết luận rằng khoản phí bổ sung là hợp lý và tiếp tục giữ Venice ở vị trí ưu tiên. Khi đó, công cụ giá không thay đổi quyết định mà chỉ làm tăng chi phí của chuyến đi. Ngược lại, nếu khoản phí đủ cao để khiến AI loại bỏ Venice, lượng khách có thể giảm mạnh nhưng đồng thời tác động tiêu cực đến hoạt động kinh tế địa phương. Điều này cho thấy việc lựa chọn mức phí tối ưu là một bài toán rất khó.

Cần công bằng với những gì công cụ này đã làm được, vì phần ấy không nhỏ. Trước năm 2024, con số hai mươi triệu khách mỗi năm của Venice là một ước lượng. Sau hai mùa thử nghiệm, thành phố có số lượt đăng ký theo từng ngày, có phân bố theo giờ, có tỷ lệ đặt trước so với đặt sát ngày, cùng một cơ sở dữ liệu đủ để một trung tâm nghiên cứu kinh tế du lịch phân tích.[2] Hệ

1. Comune di Venezia, số liệu contributo di accesso. Năm 2024, 29 ngày áp dụng, 485.062 lượt nộp phí, thu 2,4 triệu euro. Năm 2025, 54 ngày áp dụng, 723.497 lượt nộp phí, thu 5.421.425 euro, ngày cao nhất là ngày 2 tháng 5 với 24.951 lượt. Số liệu năm 2026 là lịch áp dụng chứ chưa phải kết quả.

2. Comune di Venezia. Ngày cao nhất là ngày 2 tháng 5 năm 2025 với 24.951 lượt. Con số hai mươi tới ba mươi triệu khách một năm trước đó chỉ là ước lượng từ dữ liệu điện thoại di động.

thống cũng tạo ra nguồn lực quản trị, một phần được dùng để giảm phí rác cho cư dân, đồng thời tạo ra một cơ chế pháp lý để đặt ngưỡng trong tương lai nếu thành phố muốn. Đo được, đặt ngưỡng được và có nguồn thu để vận hành là ba năng lực thật. Điều công cụ này chưa làm được nằm ở chỗ khác. Nó chưa khiến một hệ thống đang lập lịch trình cho khách đổi Venice sang một nơi tương đương.

Chuỗi số liệu này không chứng minh rằng cơ chế đề xuất là nguyên nhân. Nó không cho biết bao nhiêu khách đến vì được một hệ thống gợi ý, không cho biết các tác nhân có tính khoản phí vào hàm lựa chọn hay không, cũng không cho biết lượng khách sẽ là bao nhiêu nếu không thu phí. Nó cũng chưa loại trừ được tác động của mùa vụ, của năng lực hàng không hay của tỷ giá. Điều nó cho thấy là mức phí hiện hành chưa tạo ra dấu hiệu giảm nhu cầu đủ rõ.

Quan trọng hơn, dù được đưa vào sớm hay muộn, phí vào cửa chủ yếu tác động lên mức hấp dẫn tương đối của Venice. Nó không tự tạo ra một điểm đến thay thế có dữ liệu, niềm tin và khả năng giao dịch tương đương. Hệ thống vì thế vẫn có thể tiếp tục giữ Venice trong nhóm lựa chọn ưu tiên, chỉ với tổng chi phí cao hơn. Venice đã học được cách đếm chính xác dòng người, trước khi tìm được cách làm dòng người ấy đổi hướng.

### 5.7.2. Kyoto và sự phản ứng của cư dân

Khái niệm Quá tải du lịch (overtourism). Tình trạng lượng khách vượt quá sức chịu tải của hạ tầng, môi trường và đời sống cư dân, làm suy giảm đồng thời chất lượng sống của cư dân và chất lượng trải nghiệm của du khách.

Nếu Venice cho thấy giới hạn của công cụ kinh tế, thì Kyoto cho thấy giới hạn của năng lực chịu đựng của cộng đồng. Thành phố không chỉ đối mặt với lượng khách đông mà còn phải giải quyết những hệ quả xã hội do dòng khách tập trung vào cùng một số khu vực, đặc biệt là Gion, Arashiyama, Kiyomizu-dera và Fushimi Inari. Trong nhiều năm, cư dân Kyoto liên tục phản ánh tình trạng xe buýt quá tải, đường phố chật kín khách tham quan, tiếng ồn gia tăng và việc xâm phạm không gian riêng tư. Một số du khách đuổi theo hoặc chụp ảnh geisha và maiko mà không được phép, khiến chính quyền phải tăng cường biển cảnh báo, hạn chế tiếp cận một số tuyến phố và siết chặt các quy định về ứng xử. Quy mô của áp lực ấy có thể đo được. Năm 2024, Kyoto đón 10,88 triệu lượt khách quốc tế, mức kỷ lục, trên một thành phố khoảng 1,4 triệu dân.[1]

Điều đáng chú ý là nhiều du khách không hoàn toàn tự phát lựa chọn giống nhau. Các lịch trình của họ được định hình một phần bởi cùng một hệ sinh thái số gồm công cụ tìm kiếm, bản đồ, mạng xã hội, các kênh phân phối trực tuyến và nay là tác nhân AI. Khi nhiều hệ thống cùng đánh giá một số địa điểm là không thể bỏ lỡ, dòng khách có xu hướng hội tụ vào cùng một không gian.

Cần công bằng với thiết kế này. Thuế lưu trú Kyoto lũy tiến theo giá phòng, nên phòng dưới 6.000 yên vẫn giữ mức thuế 200 yên. Nghĩa là nó không đánh vào khách ít tiền.[2] Nhưng chính

1. Nghị quyết tăng thuế lưu trú do Hội đồng thành phố Kyoto thông qua ngày 25 tháng 3 năm 2025, được Bộ Nội vụ và Truyền thông Nhật Bản chấp thuận ngày 3 tháng 10 năm 2025, hiệu lực từ ngày 1 tháng 3 năm 2026. Con số nguồn thu là dự kiến. Xem mục Kyoto City trong danh mục tài liệu tham khảo.

2. Biểu thuế lưu trú Kyoto năm bậc, hiệu lực từ ngày 1 tháng 3 năm 2026, tính trên mỗi người mỗi đêm: dưới 6.000 yên nộp 200 yên, từ 6.000 tới dưới 20.000 yên nộp 400 yên, từ 20.000 tới dưới 50.000 yên nộp 1.000 yên, từ 50.000 tới dưới 100.000 yên nộp 4.000 yên, từ 100.000 yên trở lên nộp 10.000 yên.

cấu trúc lũy tiến ấy bộc lộ giới hạn của công cụ. Người thuê phòng 100.000 yên một đêm là người ít nhạy cảm nhất với khoản thuế 10.000 yên. Còn phần lớn áp lực lên xe buýt và lên các con dốc quanh Kiyomizu không đến từ khách lưu trú hạng sang mà đến từ khách trong ngày, nhóm không nộp đồng thuế lưu trú nào. Công cụ tài chính mạnh nhất của thành phố lại không chạm trực tiếp tới một nhóm gây áp lực đáng kể, đó là khách trong ngày.

Từ góc nhìn của từng cá nhân, lựa chọn ấy hoàn toàn hợp lý. Nhưng ở quy mô xã hội, kết quả lại là một thành phố phải gánh chịu áp lực vượt quá khả năng chịu tải. Đây chính là nghịch lý trung tâm của nền kinh tế tác nhân, không có AI nào cố ý tạo ra quá tải, nhưng tổng hợp hàng triệu quyết định tối ưu giống nhau lại tạo ra hiện tượng quá tải một cách có hệ thống.

Bài học từ Kyoto cho thấy quản trị du lịch trong thời đại AI không thể chỉ tập trung vào quản lý hành vi của du khách sau khi họ đã đến. Muốn giảm áp lực lâu dài, cần làm cho các hệ thống đề xuất chủ động phân tán dòng khách sang những điểm đến, thời gian và trải nghiệm thay thế có chất lượng tương đương. Khi thuật toán biết cách đa dạng hóa lựa chọn, cộng đồng địa phương mới thực sự được bảo vệ khỏi vòng lặp của quá tải du lịch.

### 5.7.3. Giới hạn của giá trong điều tiết đề xuất

Khái niệm giá và xếp hạng là hai việc khác nhau. Hai cơ chế tác động lên hai đại lượng khác nhau. Giá tác động lên mức hấp dẫn tương đối của một phương án đã có trong tập ứng viên. Xếp hạng quyết định phương án có mặt trong tập ấy hay không. Cả hai có thể cùng tham gia từ sớm, nhưng chỉ cái thứ hai quyết định phương án có tồn tại với người dùng.

Trong kinh tế học truyền thống, giá là một trong những công cụ mạnh nhất để điều tiết cung và cầu. Khi giá tăng, nhu cầu giảm. Khi giá giảm, nhu cầu tăng. Cơ chế này hoạt động hiệu quả khi con người trực tiếp cân nhắc giữa lợi ích và chi phí trước khi đưa ra quyết định.

Trong nền kinh tế tác nhân, quan hệ giữa giá và lựa chọn phức tạp hơn. Giá có thể tham gia từ sớm vào quá trình lọc và xếp hạng, nhưng nó chỉ là một trong nhiều biến của hàm lựa chọn. Mức giá cao hơn làm một phương án kém hấp dẫn hơn. Nó không tự làm các phương án thay thế trở nên khả đọc, đáng tin hay khả hành động.

Điều này tạo ra một sự khác biệt căn bản giữa định giá (pricing) và xếp hạng (ranking). Định giá tác động lên chi phí của một lựa chọn. Xếp hạng quyết định liệu lựa chọn đó có còn xuất hiện trước mắt người dùng hay không.

Giả sử một bảo tàng tăng giá vé thêm 20%, nhưng nếu vẫn được AI đánh giá là điểm đến có giá trị vượt trội thì nó vẫn tiếp tục đứng đầu danh sách đề xuất.[1] Ngược lại, một điểm đến miễn phí nhưng thiếu dữ liệu, thiếu bằng chứng và thiếu tín hiệu tin cậy có thể rất khó được AI giới thiệu. Trong trường hợp đó, mức giá bằng không cũng không tạo ra lượng khách. Nói cách khác, giá quyết định khả năng mua, còn xếp hạng quyết định khả năng được nhìn thấy. Trong phần lớn hành trình do tác nhân trung gian, khả năng được nhìn thấy là điều kiện tiên quyết của khả năng được mua. Giá chỉ bắt đầu phát huy tác dụng sau khi phương án đã đủ điều kiện được cân nhắc.

1. Giả định của tác giả. Mức tăng giá 20% là giả định đặt ra để tách bạch hai cơ chế giá và xếp hạng, không phải trường hợp của một bảo tàng có thật.

Đó là lý do nhiều chính sách chỉ dựa trên công cụ giá ngày càng khó giải quyết tận gốc tình trạng quá tải du lịch. Chúng có thể làm thay đổi chi phí của hành vi, nhưng không thay đổi logic của hệ thống đề xuất. Muốn điều tiết dòng khách, địa phương không chỉ cần quản lý giá mà còn phải hiểu cách AI tạo ra thứ tự ưu tiên. Giá không tự tạo phương án thay thế. Công cụ giá có thể tác động đến khả năng một phương án được lựa chọn, kể cả ở bước lọc sớm theo ngân sách. Nhưng nó không tự tạo ra dữ liệu, niềm tin và khả năng giao dịch cho những phương án đang vắng mặt khỏi tập ứng viên. Vì vậy, một chính sách chỉ dựa vào giá có thể làm giảm cầu ở một nơi mà không làm nơi nào khác trở nên khả đọc hơn.

### 5.7.4. Ba tầng can thiệp vào chuỗi quyết định

Khái niệm Đa dạng hóa đề xuất (recommendation diversification). Nguyên lý chủ động đưa vào kết quả gợi ý một tỷ lệ phương án nằm ngoài nhóm phổ biến nhất, nhằm tránh dồn toàn bộ nhu cầu vào một số ít lựa chọn. Vì hệ thống đề xuất tham gia định tuyến dòng khách trước khi họ đến, bộ công cụ điều tiết cũng phải mở rộng về phía trước thời điểm giao dịch. Điều này đòi hỏi một nguyên lý bổ sung là đa dạng hóa đề xuất. Thay vì để phần lớn tác nhân cùng giới thiệu một vài địa danh quen thuộc, hệ thống cần chủ động nhận diện những lựa chọn thay thế có chất lượng tương đương và phân tán nhu cầu sang nhiều không gian khác nhau.

Phê bình công cụ giá thì dễ, đề xuất công cụ thay thế thì khó hơn, nên phần còn lại của mục này cố gắng nói cụ thể. Cách sắp xếp hữu ích nhất là theo vị trí can thiệp trong chuỗi quyết định, chia làm ba tầng. Tầng trước đề xuất tác động vào dữ liệu đầu vào,

tức thứ mà hệ thống đọc trước khi dựng tập ứng viên. Tầng trong đề xuất tác động vào chính cơ chế chọn và sắp xếp. Đây là tầng địa phương ít quyền nhất. Tầng sau đề xuất tác động vào giao dịch và vào dòng người đã hình thành. Đây là nơi chứa gần như toàn bộ bộ công cụ truyền thống. Bảng 5.1 tóm tắt ba tầng can thiệp ấy.

**Bảng 5.1.** Ba tầng can thiệp vào chuỗi quyết định

| Tầng | Công cụ | Ai kiểm soát | Địa phương làm được gì |
|---|---|---|---|
| **Trước đề xuất** | Dữ liệu sức chịu tải theo khung giờ, trạng thái đông đúc gần thời gian thực, hồ sơ thực thể của điểm đến thay thế, dữ liệu mô tả có nguồn gốc | Địa phương | Làm được gần như toàn bộ. Đây là tầng có đòn bẩy cao nhất trên mỗi đồng chi |
| **Trong đề xuất** | Đa dạng hóa kết quả, giới hạn tần suất lặp một điểm đến, ưu tiên phân tán theo thời gian và không gian | Nền tảng | Không trực tiếp. Chỉ tác động qua thỏa thuận chia sẻ dữ liệu và qua điều kiện hợp tác |
| **Sau đề xuất** | Đặt chỗ theo khung giờ, hạn ngạch, phí vào cửa, thuế lưu trú, điều tiết giao thông | Địa phương | Làm được, hiện đang làm. Nhưng đây là tầng có đòn bẩy thấp nhất với dòng khách đã được định tuyến |

Bộ công cụ hiện hành của phần lớn cơ quan quản lý du lịch nằm gần trọn ở tầng thứ ba. Đề xuất của chương này là chuyển một phần ngân sách sang tầng một, không phải bỏ tầng ba, nơi cùng một số tiền tác động được vào nhiều quyết định hơn. Câu hỏi thực tế nhất là địa phương làm được gì khi không kiểm soát thuật toán của nền tảng. Danh sách dưới đây gồm những việc nằm trong thẩm quyền của một chính quyền cấp tỉnh hoặc một cơ quan quản lý điểm đến, không cần bất kỳ sự đồng thuận nào từ phía nền tảng, trừ hai mục cuối.

- ***Việc địa phương làm được ở tầng trước đề xuất:***

*• Công bố dữ liệu sức chịu tải theo khung giờ dưới dạng máy đọc được, không phải dưới dạng thông cáo.*
*• Chuẩn hóa tín hiệu đông đúc theo một thang thống nhất trong toàn địa bàn, để tín hiệu từ các điểm khác nhau có thể so sánh được với nhau.*
*• Kết nối tồn kho và khung giờ của các điểm tham quan công lập với ít nhất hai kênh có giao diện lập trình.*
*• Xây dựng và duy trì một danh mục điểm đến thay thế có năng lực phục vụ thật, với hồ sơ dữ liệu chi tiết ngang với điểm đến nổi tiếng nhất trong địa bàn.*
*• Công bố dữ liệu mô tả từ nguồn chính thức dưới giấy phép cho phép máy sử dụng lại, để hệ thống có nguồn thẩm quyền thay cho nguồn thứ cấp.*
*• Theo dõi tỷ lệ tập trung đề xuất như một chỉ số quản lý, đo bằng phần dòng khách hoặc phần lượt đề xuất dồn vào nhóm điểm đến đứng đầu, rồi đặt mục tiêu giảm dần chỉ số này.*
*• Đưa điều kiện về dữ liệu vào hợp đồng với các đơn vị bán vé và các kênh phân phối, gồm nghĩa vụ cập nhật và nghĩa vụ trả lại dữ liệu tổng hợp cho địa phương.*
*• Đàm phán thỏa thuận chia sẻ dữ liệu hai chiều với các nền tảng lớn, lấy dữ liệu sức chịu tải và dữ liệu điểm thay thế làm phần địa phương góp vào.*

Trong thương mại điện tử, các hệ thống gợi ý từ lâu đã áp dụng nguyên lý này để tránh việc toàn bộ lưu lượng truy cập chỉ tập trung vào một vài sản phẩm. Trong nghe nhạc hay xem phim, thuật toán cũng dần bổ sung những nội dung ít phổ biến hơn nhằm tăng tính đa dạng và giảm hiện tượng "người thắng lấy tất cả". Du lịch trong tương lai cũng sẽ đối mặt với yêu cầu tương tự.

Nguyên lý này không mới trong khoa học máy tính. Cái giá của nó cũng đã được đo. Năm 2005, Cai-Nicolas Ziegler cùng Sean McNee, Joseph Konstan và Georg Lausen công bố phương pháp đa dạng hóa theo chủ đề tại hội nghị World Wide Web.[1] Kết luận của họ có hai vế. Vế thứ hai thường bị bỏ quên khi trích dẫn. Đa dạng hóa làm giảm độ chính xác trung bình của từng gợi ý. Nhưng nó làm tăng mức hài lòng của người dùng với cả danh sách. Nói cách khác, muốn có đa dạng thì phải chấp nhận trả giá bằng độ chính xác điểm. Cái giá trực tiếp trước tiên xuất hiện trong chỉ số đánh giá của nền tảng, vì độ chính xác của từng gợi ý có thể giảm. Nhưng chi phí và lợi ích cuối cùng được phân bổ giữa nền tảng, người dùng và các địa phương, nên đây không chỉ là một bài toán kỹ thuật mà còn là một bài toán phân phối. Đó là lý do vì sao rất ít nền tảng có động cơ tự nguyện đa dạng hóa mạnh nếu chỉ số đánh giá nội bộ của họ vẫn chủ yếu là độ chính xác của từng gợi ý.

Có nhận định rằng nếu đa dạng hóa làm giảm độ chính xác thì yêu cầu nền tảng đa dạng hóa chính là yêu cầu họ phục vụ người dùng kém đi. Một chính sách buộc AI gợi ý những nơi kém phù hợp hơn là chính sách lấy lợi ích của du khách bù cho lợi ích của địa phương.

Phản bác này sắc và chỉ sai ở một chỗ, đó là chữ chính xác. Cái mà đa dạng hóa làm giảm là độ chính xác đo trên từng gợi ý đơn lẻ, tính bằng khả năng người dùng chấp nhận đúng phương án được đưa ra đầu tiên. Cái mà nó làm tăng, theo chính nghiên cứu năm 2005 nói trên, là mức hài lòng của người dùng với toàn

1. Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan và Georg Lausen, "Improving Recommendation Lists Through Topic Diversification", Proceedings of the 14th International Conference on World Wide Web, ACM, New York, 2005, trang 22 tới 32.

bộ danh sách. Hai đại lượng này khác nhau và có thể đi ngược chiều nhau. Thêm nữa, phép so sánh cần tính cả trạng thái điểm đến vào thời điểm khách tới. Một gợi ý chính xác về mặt dữ liệu nhưng dẫn khách tới một quảng trường đông đến mức không đi nổi thì không phải một gợi ý tốt, nó chỉ là một gợi ý đúng theo dữ liệu của tháng trước. Điểm phải thừa nhận là ở đây có một sự đánh đổi thật, không phải một bữa trưa miễn phí. Ai chịu chi phí của sự đánh đổi ấy là một câu hỏi phân phối chứ không phải một câu hỏi kỹ thuật.

Muốn AI đề xuất đa dạng hơn, địa phương không thể chỉ yêu cầu thuật toán thay đổi. Họ phải tạo ra đủ dữ liệu để những lựa chọn thay thế thực sự tồn tại trong không gian số. Những điểm đến mới cần có định danh rõ ràng, dữ liệu phong phú, đánh giá đáng tin cậy, hình ảnh chất lượng, khả năng đặt chỗ, thông tin thời gian thực và mối liên hệ với các trải nghiệm khác. Chỉ khi đó, AI mới có cơ sở để coi chúng là những phương án đủ sức cạnh tranh với các điểm đến nổi tiếng.

Điều này đánh dấu sự chuyển dịch từ quản trị dòng khách sang quản trị hệ thống đề xuất. Chính quyền địa phương sẽ không chỉ quản lý đường phố, bến bãi hay giá vé, mà còn phải quản lý hạ tầng dữ liệu, bản thể luận, tri thức và tín hiệu số mà AI sử dụng để lập kế hoạch.

Trong nền kinh tế tác nhân, điều tiết AI không có nghĩa là kiểm soát thuật toán, mà là thiết kế hệ sinh thái dữ liệu để thuật toán có nhiều lựa chọn tốt hơn. Khi đó, sự đa dạng trong đề xuất có thể trở thành một lớp công cụ bổ sung quan trọng cho việc tăng giá hay giới hạn số lượng khách sau khi quyết định đã được đưa ra.

Các công cụ quản lý dòng khách khác nhau ở thời điểm chúng can thiệp vào quá trình ra quyết định của hệ thống:

- ***Phí vào cửa và thuế lưu trú*** chỉ tác động sau khi điểm đến đã được lựa chọn. Chúng hầu như không làm thay đổi tập phương án mà người dùng hoặc tác nhân cân nhắc, mà chủ yếu điều chỉnh chi phí của quyết định cuối cùng. Vì vậy, công cụ này thường chỉ sàng lọc theo khả năng chi trả chứ không phân bổ lại dòng khách một cách đáng kể.

- ***Hạn ngạch và cơ chế đặt chỗ bắt buộc*** cũng can thiệp sau khi lựa chọn đã hình thành. Tác động của chúng chủ yếu đến từ tín hiệu "hết chỗ", buộc một số khách phải thay đổi thời điểm khởi hành. Tuy nhiên, chúng thường chỉ dịch chuyển nhu cầu sang ngày hoặc giờ khác, thay vì dẫn khách đến những điểm đến thay thế.

- ***Các chiến dịch quảng bá điểm đến mới*** xuất hiện sớm hơn, ở giai đoạn trước khi người dùng lựa chọn. Tuy nhiên, nếu điểm đến chưa có dữ liệu nền đủ dày và chưa tích lũy được các tín hiệu xác nhận đáng tin cậy, hiệu quả của quảng bá thường bị giới hạn. Quảng cáo có thể thu hút sự chú ý, nhưng không tự tạo ra lịch sử giao dịch hay niềm tin của thuật toán.

- ***Dữ liệu mở về các điểm đến thay thế*** can thiệp ngay từ lúc hệ thống xây dựng tập ứng viên. Khi những điểm đến này được mô tả đầy đủ, chuẩn hóa và dễ truy cập bằng máy, chúng có cơ hội xuất hiện trong danh sách đề xuất ngay từ đầu. Đây là tác động trực tiếp lên tập lựa chọn, nhưng chỉ phát huy hiệu quả nếu năng lực phục vụ ngoài đời thực đủ để đáp ứng lượng khách được phân bổ.

- ***Dữ liệu sức chịu tải theo thời gian thực*** tiếp tục can thiệp ở bước xếp hạng. Khi hệ thống biết nơi nào đang quá tải, nơi nào còn dư năng lực, thuật toán có thể điều chỉnh thứ tự đề xuất để phân bổ dòng khách hợp lý hơn. Muốn làm được điều đó cần có chuẩn dữ liệu thống nhất, khả năng cập nhật liên tục và sự phối hợp giữa đơn vị cung cấp dữ liệu với các nền tảng tiếp nhận.

Ba công cụ đầu thuộc bộ công cụ quản lý du lịch truyền thống, chúng chủ yếu điều chỉnh hành vi sau khi quyết định gần như đã được hình thành. Hai công cụ sau phản ánh tư duy của nền kinh tế tác nhân, thay vì kiểm soát dòng khách ở điểm cuối, chúng can thiệp ngay vào dữ liệu và cơ chế lựa chọn của máy. Đáng chú ý, đây lại là những hạng mục vẫn còn nằm ngoài ưu tiên đầu tư và thói quen ngân sách của phần lớn cơ quan quản lý du lịch hiện nay.

Điều tiết hệ thống đề xuất là việc bất khả thi với một địa phương. Các hệ thống ấy do những tập đoàn xuyên quốc gia vận hành, không công bố tiêu chí và không chịu thẩm quyền của một chính quyền cấp tỉnh. Nói địa phương nên điều tiết AI là nói một điều nghe hay mà không làm được.

Phản bác này đúng nếu hiểu điều tiết là ra lệnh cho thuật toán. Nó sai nếu hiểu điều tiết là thay đổi dữ liệu đầu vào. Không một chính quyền địa phương nào ép được một mô hình thay đổi cách xếp hạng. Nhưng mọi chính quyền địa phương đều kiểm soát thứ mà mô hình phải đọc: hồ sơ thực thể của các điểm đến trong địa bàn, dữ liệu sức chịu tải, giờ giấc phục vụ, tình trạng còn chỗ và mức độ đầy đủ của thông tin về những nơi chưa nổi tiếng. Đó là một quyền lực có thật, tuy gián tiếp. Điểm cần thừa nhận là quyền lực ấy bất đối xứng. Địa phương tác động được vào đầu vào và không tác động được vào hàm mục tiêu, nên mọi kết

quả đều phụ thuộc vào việc hàm mục tiêu ấy có coi trọng sự đa dạng hay không. Vì thế phần còn lại của bài toán thuộc về chính sách cấp quốc gia và cấp quốc tế, một chủ đề mà chương này chỉ đặt ra chứ chưa giải quyết.

- ***Nếu bạn là người điều hành một điểm đến di sản:***
  - *Hãy công bố giới hạn sức chứa dưới dạng dữ liệu máy đọc được, kèm khung giờ. Một con số trong thông cáo báo chí không đi vào bất kỳ hệ thống nào.*
  - *Hãy giữ dữ liệu về những khu vực ít khách trong quần thể ở cùng mức chi tiết với khu vực nổi tiếng nhất. Chênh lệch chi tiết trong hồ sơ dữ liệu sẽ tái tạo đúng chênh lệch dòng khách ngoài thực địa.*
  - *Hãy coi mỗi lần một tác nhân đặt chỗ thất bại là một sự cố cần ghi nhận, vì hệ thống sẽ ghi nhận nó lâu hơn bạn tưởng.*

- ***Nếu bạn là chủ một cơ sở lưu trú nhỏ:***
  - *Hãy bảo đảm bốn thông tin luôn đúng ở mọi nơi: vị trí, giờ nhận phòng, giá công bố và tình trạng còn phòng. Đây là bốn trường cơ bản mà nhiều tác nhân giao dịch cần kiểm tra trước khi có thể cân nhắc việc đặt chỗ.*
  - *Hãy chấp nhận rằng một đánh giá tiêu cực trung thực ít nguy hiểm hơn một thông tin sai lệch, vì thứ nhất làm giảm điểm còn thứ hai làm giảm độ tin cậy của toàn bộ hồ sơ.*
  - *Hãy kết nối với ít nhất hai kênh phân phối có giao diện lập trình thay vì phụ thuộc vào một kênh duy nhất, kể cả khi kênh ấy đang mang lại phần lớn doanh thu.*

Trước khi sang phần kết, một khoảng nghỉ. Ba khuyến nghị vừa nêu đều kiểm tra được trong một buổi làm việc với dữ liệu của chính cơ sở mình và người đọc nên kiểm tra ít nhất một khuyến nghị trước khi đồng ý với phần còn lại.

## 5.8. KẾT LUẬN CHƯƠNG

Trong nền kinh tế tác nhân, vấn đề trung tâm không chỉ là lượng khách, mà còn là cách dòng khách được các hệ thống số định tuyến. Khi nhiều hệ thống cùng truy hồi từ những nguồn dữ liệu tương tự, cùng dùng những tín hiệu đánh giá gần giống nhau và cùng tối ưu theo những tiêu chí tương tự, chúng có xu hướng hội tụ vào một số ít lựa chọn.

Nếu Chương 4 đã giải thích vì sao một địa phương được AI tin tưởng, thì Chương 5 cho thấy niềm tin ấy được chuyển hóa thành quyền lực phân bổ. Quyền lực đó không chỉ quyết định dòng người sẽ đi đâu, mà còn quyết định dòng tiền sẽ chảy về đâu, sự chú ý sẽ tập trung vào đâu và cuối cùng là cơ hội phát triển sẽ thuộc về ai.

Từ góc nhìn đó, quản trị địa phương trong nền kinh tế tác nhân không dừng ở việc xây dựng dữ liệu đầy đủ hay chuẩn hóa thông tin. Mục tiêu cao hơn là thiết kế một hệ sinh thái số giúp địa phương vừa đủ khả năng được thuật toán lựa chọn, vừa đủ năng lực phân tán và điều tiết dòng khách để bảo vệ tài nguyên, cộng đồng và bản sắc. Một địa phương thành công không phải là địa phương đứng đầu mọi bảng xếp hạng, mà là địa phương chuyển hóa được sự ưu tiên của thuật toán thành giá trị kinh tế dài hạn mà không đánh đổi tính bền vững.

> ▪ ***Về mục tiêu của cạnh tranh trong nền kinh tế tác nhân:***
>
> • *Không được đặt mục tiêu duy nhất là được đề xuất nhiều hơn. Mục tiêu phải là được đề xuất đúng người, đúng thời điểm và trong giới hạn sức chịu tải.*

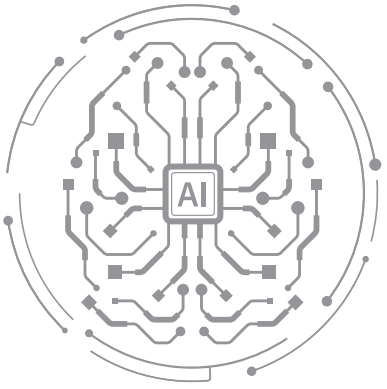
> • *Không được đầu tư vào khả đọc mà bỏ qua năng lực phục vụ. Khả đọc đưa khách đến, năng lực phục vụ quyết định họ có để lại một xác nhận tích cực hay không.*
> • *Không được coi chương này đã giải quyết xong bài toán phân phối. Nó mới chỉ mô tả cơ chế phân phối. Việc giữ lại giá trị kinh tế từ dòng khách được phân phối là chủ đề của chương tiếp theo.*

Đó cũng là điểm mở sang Chương 6. Sau khi hiểu cách AI tạo ra và phân bổ nhu cầu, câu hỏi tiếp theo không còn là "Ai được đề xuất?", mà là "Ai thực sự giữ lại được giá trị kinh tế do những đề xuất ấy tạo ra?". Bởi được AI lựa chọn mới chỉ là điểm khởi đầu. Điều quyết định sự thịnh vượng của một địa phương nằm ở khả năng chuyển dòng người thành đầu tư, việc làm, thu nhập, tài sản và năng lực phát triển lâu dài. Được máy chọn là điều kiện cần. Nó chưa bao giờ là điều kiện đủ.

## CHƯƠNG 6

# AI GIỮ LẠI GIÁ TRỊ

*Bốn bộ lọc đọc như một biểu giá. Chuỗi giá trị bị cắt ở quyền lựa chọn và giao diện giao dịch, tô thuật toán, ba lớp chi phí tham gia và nguy cơ địa phương bị OEM hóa.*


AI

Bốn bộ lọc, tới chương này, là một mô hình nhận thức và trở thành một biểu giá. Mỗi bộ lọc là một khoản chi: phân giải và chứng thực tạo ra chi phí khả đọc, xếp hạng tạo ra chi phí cạnh tranh phân phối, thực thi tạo ra chi phí khả hành động. Ai kiểm soát một bộ lọc thì thu tô ở đó. Đó cũng là cách cơ chế tái cấu trúc chuyển thành tiền: thế giới tự sửa mình theo lược đồ và người giữ lược đồ được trả công cho việc đó.

Trong nền kinh tế trải nghiệm, ai tạo giá trị thì giữ giá trị. Trong nền kinh tế tác nhân, hai vế ấy không còn cùng một chủ. Trong nền kinh tế trải nghiệm, ai tạo ra giá trị thì người đó giữ được phần lớn giá trị. Trong nền kinh tế tác nhân, hai vế ấy tách rời nhau. Giá trị vẫn sinh ra ở nơi chốn, nhưng lắng lại ở nơi kiểm soát quyền đọc, quyền xếp hạng, quyền lựa chọn và quyền giao dịch.

Trong kinh tế trải nghiệm, chuỗi giá trị của địa phương được mô tả như một quá trình tích lũy khép kín. Địa phương đầu tư vào tài sản, tài sản tạo ra trải nghiệm, trải nghiệm để lại ký ức, ký ức kết thành thương hiệu, thương hiệu hút người và tiền, rồi dòng tiền ấy quay lại làm giàu thêm tài sản ban đầu. Chuỗi này dựa trên một giả định ít khi được nói ra vì nó quá hiển nhiên, nơi tạo ra giá trị cũng là nơi giữ lại phần lớn giá trị. Trung gian có tồn tại và từ lâu đã biết giới hạn tập lựa chọn, nhưng phần lớn giao dịch vẫn cần khách hàng xác nhận từng bước.

Nền kinh tế tác nhân không xóa chuỗi ấy. Nó chen vào đoạn nối giữa trải nghiệm và chi tiêu, rồi tách đoạn nối ấy thành một chuỗi quyết định mới. Khi người dùng chuyển từ tự tìm kiếm sang giao nhiệm vụ cho một tác nhân AI, một phần ngày càng lớn của quá trình lựa chọn được ủy thác cho máy. Khi tác nhân được cấp quyền hành động, một phần quyền thực hiện giao dịch cũng dịch chuyển theo. Hai điểm cắt ấy đủ để tái phân phối toàn bộ dòng giá trị của chuỗi, dù trải nghiệm vẫn diễn ra đúng tại nơi chốn cũ và với chất lượng không đổi.

Đó là chủ đề của chương này. Địa phương vẫn đầu tư vào di sản, cảnh quan, hạ tầng và con người. Doanh nghiệp địa phương vẫn phục vụ du khách. Nhưng ngày càng nhiều giá trị được giữ lại bởi những chủ thể không tạo ra trải nghiệm, chỉ kiểm soát quá trình khiến trải nghiệm được đọc, được xếp hạng, được chọn và được thanh toán. Câu hỏi của chương không phải giá trị sinh ra ở đâu. Câu hỏi là dòng tiền lắng lại ở đâu.

Nếu Chương 4 giải thích máy hình thành tri thức về địa phương như thế nào và Chương 5 phân tích cách máy phân bổ sự chú ý cùng dòng khách, thì Chương 6 khép lại phần cơ chế bằng câu hỏi về phân phối. Chương 7 sẽ chuyển từ phân tích sang thiết kế, tức là địa phương phải tự dựng lại chính mình ra sao để giành lại vị trí trong chuỗi mới.

## 6.1. CHUỖI GIÁ TRỊ BỊ CẮT Ở QUYỀN LỰA CHỌN VÀ GIAO DỊCH

Chuỗi cũ vận hành trên một điều kiện nền: phần lớn giao dịch vẫn cần con người xác nhận quyết định cuối cùng. Chính con người tìm thông tin, so sánh phương án, cảm nhận trải nghiệm và chọn nơi tiêu tiền. Doanh nghiệp trung gian giúp kết nối cung với cầu, nhưng quyền chọn cuối cùng vẫn nằm ngoài họ. Nhờ vậy, phần lớn giá trị rốt cuộc quay về nơi đã tạo ra trải nghiệm.

Mục này không viết lại chuỗi ấy. Nó chỉ xác định hai chỗ mà nền kinh tế tác nhân cắt vào, rồi giải thích vì sao chỉ cần cắt đúng hai chỗ đó là quy luật phân bổ giá trị đã đổi. Điều đáng nói là hai vết cắt này rất khó nhìn thấy từ bên trong địa phương. Khách vẫn đến, doanh thu vẫn phát sinh, các con số thống kê vẫn đẹp. Thứ đã đổi chủ là quyền tiếp cận khách hàng và quyền quyết định giao dịch.

### 6.1.1. Chuỗi trải nghiệm truyền thống

Nguyên lý của chuỗi cũ rất đơn giản. Mọi giá trị kinh tế lâu dài đều bắt đầu từ trải nghiệm, không từ truyền thông. Tài sản của địa phương, gồm cảnh quan, di sản, ẩm thực, lối sống, tri thức bản địa và con người, chỉ trở thành giá trị khi được tổ chức thành một thứ mà khách có thể sống qua. Một ngọn núi cần đường mòn, người dẫn đường và một câu chuyện thì mới thành giá trị kinh tế.

Trải nghiệm để lại ký ức. Ký ức mới là tài sản vô hình quan trọng nhất, vì nó quyết định doanh thu tương lai chứ không chỉ doanh thu hôm nay. Khi hàng triệu ký ức cá nhân trùng nhau đủ nhiều, chúng kết thành thương hiệu. Thương hiệu kéo theo chi

tiêu, chi tiêu đủ lớn thì kéo theo đầu tư, rồi đầu tư quay lại nâng cấp chính tài sản ban đầu. Giá trị lắng lại như phù sa sau mỗi mùa nước. Toàn bộ vòng tuần hoàn ấy dựa trên một điều kiện đã nói ở trên, phần lớn lựa chọn cuối cùng vẫn do con người xác nhận. Đó chính là điều kiện sắp bị gỡ bỏ.

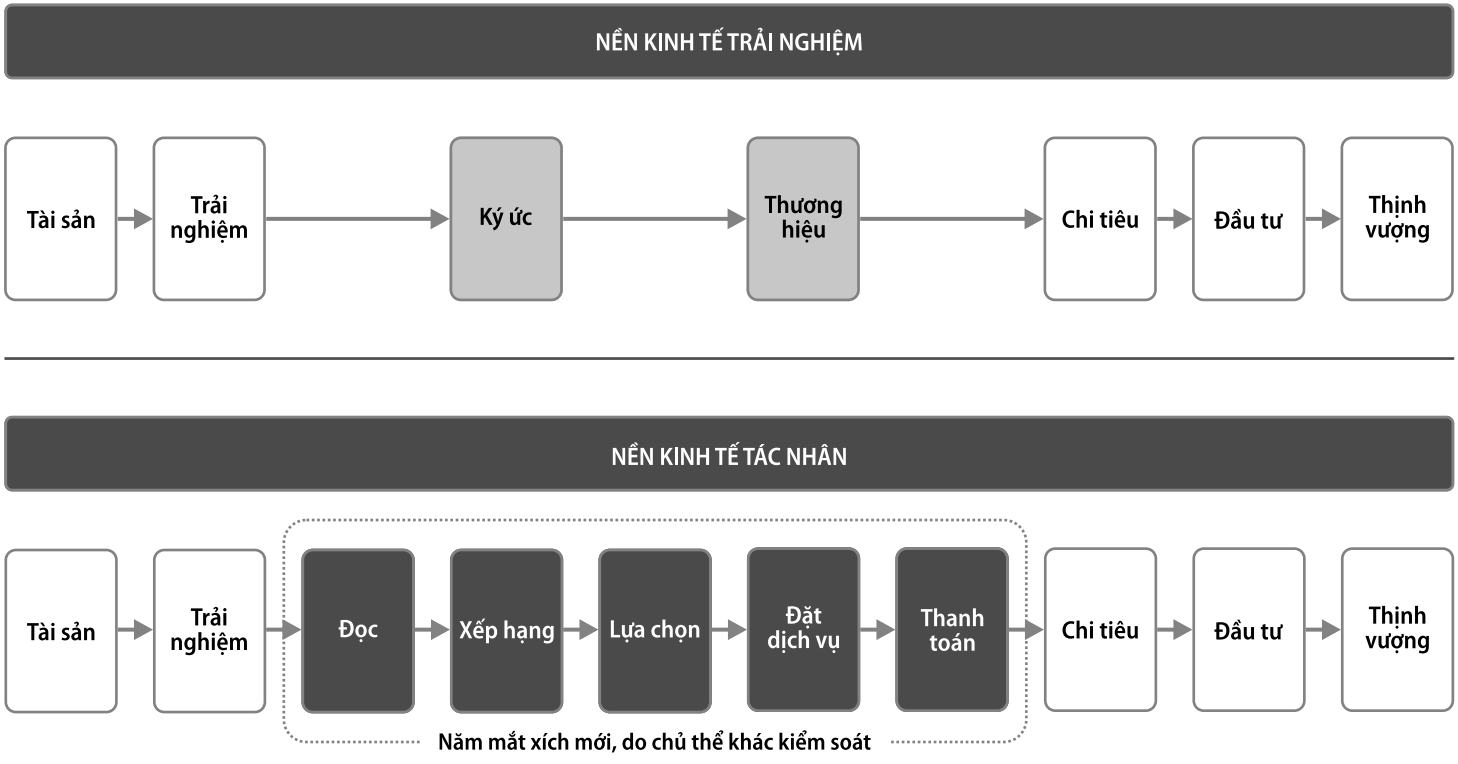
### 6.1.2. Tác nhân chen vào ở đâu

Chuỗi cũ không đứt vì trải nghiệm mất giá. Nó đổi vì giữa các mắt xích xuất hiện một chủ thể mới. Chủ thể ấy không xây khách sạn, không tổ chức lễ hội, không bảo tồn di sản. Nó không trực tiếp cung cấp trải nghiệm tại nơi chốn. Vai trò của nó là đọc dữ liệu, so sánh phương án, điều phối và ngày càng thay người dùng thực hiện quyết định.

Năm công đoạn trước đây do du khách tự làm nay lần lượt được máy tiếp quản. Trước hết là đọc, nếu địa phương không hiện diện trong bộ nhớ của máy hoặc bị mô tả sai, chuỗi dừng ngay từ bước đầu. Kế đó là xếp hạng, nơi một thay đổi nhỏ về thứ tự có thể chuyển hàng triệu lượt khách từ nơi này sang nơi khác. Rồi đến lựa chọn, khi người dùng không còn hỏi nên đi đâu mà giao thẳng nhiệm vụ lên kế hoạch chuyến đi.

Hai công đoạn cuối mới là chỗ tiền thực sự chảy qua. Ở bước đặt dịch vụ, chủ thể nắm hệ thống tồn kho và mạng lưới đối tác quyết định trải nghiệm nào được đưa vào hành trình. Ở bước thanh toán, chủ thể xử lý giao dịch giữ luôn dữ liệu chi tiêu, thông tin khách hàng và các dịch vụ tài chính đi kèm. Thanh toán trông như một thao tác kỹ thuật, nhưng đó là nơi chuỗi khép lại.

Có thể hình dung sự thay đổi này bằng hai chuỗi giá trị song song, đặt cạnh nhau ở Hình 6.1.




**Hình 6.1.** Hai chuỗi giá trị đặt cạnh nhau

*Nguồn: Tác giả xây dựng*

Năm mắt xích mới ấy không thay thế trải nghiệm mà bao bọc lấy trải nghiệm. Chúng trở thành những điểm kiểm soát (control points) của chuỗi giá trị mới. Ai kiểm soát được nhiều điểm hơn sẽ có nhiều khả năng giữ lại dữ liệu, khách hàng và biên lợi nhuận hơn. Vì vậy, điều thay đổi căn bản của nền kinh tế tác nhân không phải là địa phương tạo ra ít giá trị hơn, mà là ngày càng nhiều giá trị được giữ lại trước khi kịp quay trở về nơi đã tạo ra trải nghiệm.

Mỗi mắt xích mới trong nền kinh tế tác nhân không chỉ thực hiện một chức năng kỹ thuật, mà còn trả lời một câu hỏi quyết định, do một nhóm chủ thể khác nhau kiểm soát và tạo ra một dòng giá trị kinh tế riêng:

- ***Đọc trả lời câu hỏi:*** *Địa phương có tồn tại và được máy nhận diện hay không?* Quyền kiểm soát thường thuộc về các nhà cung cấp mô hình AI, kho tri thức, bản đồ số và các hệ thống quản lý thực thể. Giá trị được tạo ra thông qua quyền truy cập dữ liệu, lợi thế sở hữu tri thức và khả năng huấn luyện mô hình.

- ***Xếp hạng trả lời câu hỏi:*** *Ai được nhìn thấy trước?* Mắt xích này chủ yếu do các nền tảng đề xuất, công cụ tìm kiếm và hệ thống gợi ý kiểm soát. Giá trị được thu từ quảng cáo, phí hiển thị và các chương trình ưu tiên dành cho đối tác.

- ***Lựa chọn trả lời câu hỏi:*** *Phương án nào được chọn thay người dùng?* Đây là lớp quyền lực mới của các tác nhân AI và trợ lý số, nơi quyết định cuối cùng không còn do người dùng trực tiếp đưa ra. Giá trị kinh tế thường đến từ phí giới thiệu, chia sẻ doanh thu hoặc các thỏa thuận phân phối.

- ***Đặt dịch vụ trả lời câu hỏi:*** *Giao dịch sẽ được thực hiện trên kênh nào?* Quyền kiểm soát nằm trong tay các nền tảng đặt chỗ, hệ thống phân phối và hạ tầng giao dịch. Hình thức thu giá trị phổ biến là hoa hồng trên mỗi đơn đặt chỗ hoặc mỗi giao dịch thành công.

- ***Thanh toán trả lời câu hỏi:*** *Dòng tiền sẽ đi qua con đường nào?* Mắt xích này do các cổng thanh toán, tổ chức thẻ và ví điện tử kiểm soát. Giá trị không chỉ đến từ phí giao dịch mà còn từ dữ liệu chi tiêu, khả năng cung cấp các dịch vụ tài chính đi kèm và việc duy trì quyền kiểm soát dòng tiền trong toàn bộ hệ sinh thái.

Nhìn tổng thể, năm mắt xích này không chỉ quyết định liệu một địa phương có được lựa chọn hay không, mà còn quyết định ai nắm quyền kiểm soát sự chú ý, giao dịch và dòng tiền trong nền kinh tế tác nhân. Đây chính là lớp giá trị mới được chèn vào giữa trải nghiệm và chi tiêu, đồng thời cũng là nơi quyền lực kinh tế dịch chuyển từ nhà cung cấp dịch vụ sang các chủ thể kiểm soát hạ tầng số.

Một điểm cần tách bạch trước khi đi tiếp. Trong nhiều cuộc thảo luận, bốn thứ khác nhau thường bị gọi chung là AI, mô hình suy luận, tác nhân thay mặt người dùng, nền tảng phân phối tồn kho và hạ tầng thanh toán. Chúng có thể thuộc cùng một tập đoàn, cũng có thể thuộc bốn chủ thể độc lập cạnh tranh nhau. Tác nhân không phải lúc nào cũng là bên thu tiền. Bên thu tiền cũng không phải lúc nào cũng là bên suy luận.

Trong nền kinh tế tác nhân, chuỗi giá trị được hình thành từ năm lớp hạ tầng, mỗi lớp đảm nhiệm một chức năng khác nhau, do những chủ thể khác nhau kiểm soát và tạo ra một dòng doanh thu riêng:

- ***Mô hình và tác nhân*** là lớp chịu trách nhiệm suy luận, lập kế hoạch và đưa ra lựa chọn thay người dùng. Quyền kiểm soát thường thuộc về các nhà cung cấp mô hình AI hoặc các tác nhân độc lập. Giá trị kinh tế được tạo ra thông qua phí thuê bao, phí tính theo từng tác vụ hoặc cơ chế chia sẻ doanh thu từ các giao dịch mà tác nhân khởi tạo.

- ***Giao diện*** là lớp duy trì quan hệ trực tiếp với người dùng cuối. Đây có thể là trợ lý số, siêu ứng dụng hoặc công cụ tìm kiếm, nơi câu hỏi được đặt ra và câu trả lời được hiển thị. Dòng doanh thu chủ yếu đến từ quảng cáo, phí giới thiệu, phí hiển thị ưu tiên và các hình thức thương mại hóa sự chú ý của người dùng.

- ***Phân phối*** là lớp kết nối nhu cầu với năng lực cung ứng. Các nền tảng đặt dịch vụ, hệ thống phân phối toàn cầu và chợ trực tuyến kiểm soát tồn kho, xác nhận khả năng phục vụ và điều phối giao dịch. Giá trị được thu chủ yếu dưới dạng hoa hồng trên mỗi đơn đặt chỗ hoặc giao dịch thành công.

- ***Thanh toán*** là lớp hoàn tất giao dịch và ghi nhận dòng tiền. Quyền kiểm soát nằm trong tay các ví điện tử, ngân hàng và tổ chức thẻ. Ngoài phí giao dịch, lớp này còn tạo ra giá trị từ dữ liệu chi tiêu và các dịch vụ tài chính đi kèm như tín dụng, trả góp, bảo hiểm hay chương trình khách hàng thân thiết.

- ***Cung ứng*** là lớp duy nhất tạo ra trải nghiệm thực trong thế giới vật lý. Khách sạn, nhà hàng, điểm tham quan và cộng đồng địa phương trực tiếp cung cấp dịch vụ cho du khách, tạo ra doanh thu từ việc bán sản phẩm và trải nghiệm. Tuy nhiên, trong nền kinh tế tác nhân, dù vẫn là nơi tạo ra giá trị gốc, lớp cung ứng ngày càng phụ thuộc vào bốn lớp phía trên để được nhìn thấy, được lựa chọn và được giao dịch.

Điểm khác biệt căn bản so với nền kinh tế trải nghiệm là quyền lực không còn tập trung chủ yếu ở lớp cung ứng. Nó được phân tán lên các lớp hạ tầng số phía trước, từ mô hình AI, giao diện, phân phối đến thanh toán, nơi kiểm soát việc ai được nhìn thấy, ai được lựa chọn và dòng tiền sẽ đi qua con đường nào. Các lớp trên có thể tách rời hoặc tích hợp trong một chủ thể duy nhất. Mức độ tích hợp càng cao thì khả năng thu giá trị càng lớn. Đó mới là biến số cần theo dõi, không phải bản thân AI.

### 6.1.3. Điểm cắt đầu tiên là quyền lựa chọn

Trong toàn bộ chuỗi giá trị của nền kinh tế trải nghiệm, quyền lực luôn bắt đầu từ quyền lựa chọn. Ai quyết định du khách sẽ đến đâu, ở khách sạn nào, ăn tại nhà hàng nào và tham gia hoạt động gì thì người đó có ảnh hưởng lớn nhất đến dòng tiền của toàn bộ hệ thống. Trước đây, quyền lực ấy thuộc về con người. Du khách đọc nhiều nguồn thông tin, tham khảo bạn bè, xem bản đồ, so sánh giá cả rồi tự mình đưa ra quyết định cuối cùng. Các doanh nghiệp chỉ có thể cố gắng thuyết phục, chứ không thể quyết định thay khách hàng.

Nền kinh tế tác nhân làm thay đổi điểm xuất phát ấy. Khi người dùng chuyển từ việc tự tìm kiếm sang giao nhiệm vụ cho AI, quyền lựa chọn bắt đầu dịch chuyển. Thay vì đọc hàng chục trang web hay hàng trăm đánh giá, người dùng chỉ cần nói: "Hãy lên kế hoạch cho chuyến đi phù hợp nhất." Từ thời điểm đó, AI không còn đóng vai trò công cụ tra cứu mà trở thành chủ thể thực hiện lựa chọn.

Quyền lực vì thế không còn nằm chủ yếu ở nơi tạo ra trải nghiệm mà dịch chuyển về nơi kiểm soát quá trình lựa chọn.

Người kiểm soát thuật toán quyết định những lựa chọn nào xuất hiện trước, lựa chọn nào bị loại bỏ và trong trường hợp tác nhân được phép hành động, lựa chọn nào sẽ được thực hiện. Đó là quyền lực phân bổ nhu cầu ở quy mô lớn. Chỉ cần thay đổi một tiêu chí xếp hạng hoặc một tham số tối ưu, dòng khách có thể chuyển từ địa phương này sang địa phương khác mà không cần bất kỳ chiến dịch quảng bá nào.

Sự dịch chuyển này hầu như vô hình. Đối với người dùng, AI chỉ đơn giản là "giúp lựa chọn". Nhưng về mặt kinh tế học, quyền lực đã đổi chủ. Từ đây, câu hỏi quan trọng không còn là ai tạo ra trải nghiệm tốt nhất, mà ngày càng trở thành ai nắm quyền quyết định trải nghiệm nào được lựa chọn. Trong nền kinh tế tác nhân, hai câu hỏi ấy không còn có cùng một đáp án.

### 6.1.4. Điểm cắt thứ hai là giao diện giao dịch

Nếu quyền lựa chọn là điểm cắt đầu tiên của chuỗi giá trị, thì giao diện giao dịch là điểm cắt thứ hai. Đây là nơi mối quan hệ trực tiếp giữa địa phương và khách hàng dần bị thay thế bằng mối quan hệ thông qua nền tảng và tác nhân AI.

Trong nền kinh tế trải nghiệm truyền thống, dù khách có thể biết đến địa phương qua báo chí, đại lý hay mạng xã hội, họ vẫn thường tương tác trực tiếp với doanh nghiệp địa phương ở nhiều khâu quan trọng: đặt phòng với khách sạn, gọi điện cho hướng dẫn viên, mua vé tại điểm tham quan hay thanh toán trực tiếp tại nhà hàng. Mỗi lần tương tác như vậy đều giúp doanh nghiệp hiểu khách hàng hơn, xây dựng lòng trung thành và giữ lại dữ liệu cho những lần phục vụ tiếp theo.

Nền kinh tế tác nhân thay đổi cấu trúc ấy bằng cách đưa giao diện trở thành trung tâm của mọi giao dịch. Người dùng không còn tiếp xúc trực tiếp với từng nhà cung cấp mà giao tiếp với một tác nhân AI hoặc một nền tảng duy nhất. Chính giao diện này đọc yêu cầu, tổng hợp lựa chọn, thực hiện đặt dịch vụ, xử lý thanh toán và theo dõi toàn bộ hành trình của khách hàng. Đối với người dùng, mọi trải nghiệm đều diễn ra trên cùng một cửa sổ hội thoại hoặc một ứng dụng duy nhất.

Điều này có nghĩa là quan hệ giữa địa phương và du khách ngày càng trở thành quan hệ gián tiếp. Địa phương cung cấp trải nghiệm, nhưng không còn sở hữu giao diện nơi quyết định mua bán diễn ra. Họ biết có khách đến nhưng không biết khách đã cân nhắc những phương án nào, vì sao họ được lựa chọn, hay khách sẽ tiếp tục đi đâu sau đó. Những dữ liệu có giá trị nhất về hành vi khách hàng được giữ lại ở lớp giao diện chứ không còn ở nơi tạo ra dịch vụ.

Vì vậy, điểm cắt thứ hai không chỉ là việc xuất hiện thêm một ứng dụng hay một công cụ đặt dịch vụ. Nó đánh dấu sự chuyển dịch của quan hệ khách hàng khỏi địa phương sang lớp giao diện. Từ đây, doanh nghiệp địa phương có nguy cơ trở thành người cung cấp dịch vụ ở phía sau, còn người nắm giao diện mới là người sở hữu khách hàng và có khả năng giữ lại phần giá trị lớn nhất của chuỗi giao dịch.

### 6.1.5. Từ tạo giá trị đến giữ giá trị

Trong cách hiểu thông thường, chuỗi giá trị mô tả một loạt hoạt động nối tiếp nhau cùng làm tăng giá trị của sản phẩm hoặc dịch vụ. Nhưng chuỗi tạo giá trị không tự động quyết định chuỗi

giữ giá trị. Hai cấu trúc này có thể trùng nhau, cũng có thể tách rời. Trong nền kinh tế trải nghiệm chúng gần như trùng nhau. Chính sự trùng khớp ấy đã khiến người ta quên rằng đó là một khả năng chứ không phải một quy luật.

Nhưng nền kinh tế tác nhân làm xuất hiện một logic khác. Giá trị vẫn được tạo ra ở địa phương, song quyền quyết định nơi giá trị sẽ dừng lại không còn thuộc về địa phương nữa. Giữa nơi tạo ra giá trị và nơi nhận được giá trị xuất hiện nhiều lớp trung gian mới: hệ thống AI, nền tảng giao dịch, cổng thanh toán, hạ tầng dữ liệu và các dịch vụ điều phối. Mỗi lớp đều có khả năng giữ lại một phần doanh thu, dữ liệu hoặc quyền lực thị trường trước khi phần còn lại đến được với người trực tiếp tạo ra trải nghiệm. Trọng tâm cạnh tranh vì thế chuyển từ việc tạo thêm giá trị sang năng lực giữ lại giá trị đã tạo ra.

Ở cấp độ cao hơn nữa, nền kinh tế tác nhân dẫn tới hiện tượng có thể gọi là chuỗi trích giá trị (value extraction). Khác với "giữ giá trị", vốn hàm ý mỗi chủ thể cố gắng giữ lại phần giá trị mình tạo ra, "trích giá trị" mô tả quá trình một chủ thể rút giá trị khỏi chuỗi mà không trực tiếp tạo ra trải nghiệm tương ứng. Người kiểm soát điểm truy cập, thuật toán đề xuất, giao diện giao dịch hay hệ thống thanh toán có thể thu được doanh thu, dữ liệu và lợi thế cạnh tranh chủ yếu nhờ vị trí trung gian của mình. Giá trị không còn chỉ được tạo ra mà còn được khai thác từ dòng chảy của giá trị.

Khi một chủ thể trung gian giành được quyền ra quyết định thay người mua, năng lực tạo ra giá trị và năng lực giữ lại giá trị không còn hội tụ ở cùng một chủ thể. Mức độ tách rời tăng theo tỷ trọng quyết định được ủy thác cho máy và theo mức độ trung gian kiểm soát giao diện, dữ liệu cùng giao dịch. Chất lượng trải nghiệm vẫn cần thiết nhưng không còn đủ để quyết định phần giá trị được giữ lại.

Sự chuyển dịch này cũng giải thích vì sao nhiều địa phương ngày càng đón nhiều khách hơn nhưng không tương xứng với mức tăng của thu nhập địa phương. Doanh thu vẫn tăng, song biên lợi nhuận ngày càng mỏng hơn vì một phần giá trị đã được giữ lại ở các lớp trung gian trước khi dòng tiền đến nơi tạo ra trải nghiệm. Chuỗi giá trị của nền kinh tế trải nghiệm không biến mất. Nó bị phủ lên bởi một chuỗi mới, trong đó năng lực giữ và trích xuất giá trị trở thành nguồn quyền lực kinh tế quan trọng không kém năng lực tạo ra giá trị.

Các chuyên gia phản biện sẽ cho rằng: Trung gian nào cũng thu phí. Hãng lữ hành, đại lý vé và OTA đã tồn tại hàng chục năm. Tác nhân AI chỉ là một kênh phân phối mới, không có gì thay đổi về bản chất.

Phản bác này đúng ở một điểm quan trọng và cần được thừa nhận, trung gian không phải phát minh của kỷ nguyên AI. Trung gian cũng tạo ra giá trị thật, bao gồm giảm chi phí tìm kiếm, gộp cầu phân tán và bảo lãnh giao dịch. Điểm khác biệt không nằm ở việc thu phí mà ở vị trí thu phí. Đại lý truyền thống, hãng lữ hành và hệ thống phân phối toàn cầu đều có thể giới hạn và định hướng tập phương án, nhưng phần lớn vẫn cần khách hàng xác nhận từng giao dịch. Tác nhân AI có thể được ủy quyền để kết hợp nhiều quyết định và tự hoàn tất cả chuỗi. Cái mới không phải việc trung gian lần đầu có ảnh hưởng tới lựa chọn, mà là mức độ cá nhân hóa, khả năng tích hợp nhiều quyết định vào một lần ủy quyền, cùng độ vô hình của những phương án đã bị loại. Một trung gian được ủy quyền quyết định làm giảm đáng kể khả năng khách hàng tự mở rộng lại tập lựa chọn, đặc biệt khi giao dịch được hoàn tất trong cùng một giao diện. Sự dịch chuyển từ hiển thị sang quyết định là điều mới, còn việc thu phí thì không. Chuỗi giá trị không bị phá vỡ. Nó bị đặt thêm một trạm thu phí ở khúc hẹp nhất.

## 6.2. TÔ THUẬT TOÁN VÀ QUYỀN LỰC ĐIỂM NGHẼN

Trong kinh tế học, tô (economic rent) là phần thu vượt quá mức cần thiết để duy trì một nguồn lực trong cách sử dụng hiện tại, thường xuất hiện khi một chủ thể kiểm soát tài sản khan hiếm, điểm nghẽn hoặc một quyền mà người khác khó thay thế. Chủ đất thu tô nhờ quyền sở hữu đất đai. Chủ cảng thu tô nhờ kiểm soát cửa ngõ giao thương. Người sở hữu hạ tầng thiết yếu thu tô nhờ nắm giữ điểm mà mọi dòng lưu chuyển đều phải đi qua.

Nền kinh tế tác nhân tạo ra một dạng điểm nghẽn mới. Đó không còn là cửa khẩu, tuyến đường hay tài nguyên thiên nhiên, mà là các cổng lựa chọn của máy. Mỗi khi một tác nhân AI truy hồi thông tin, xếp hạng phương án, đề xuất dịch vụ hay thực hiện giao dịch, nó thường phải đi qua một số nền tảng, giao diện, mô hình hoặc hạ tầng dữ liệu nhất định. Ai kiểm soát những điểm trung chuyển ấy sẽ có khả năng thu một phần giá trị của mọi giao dịch, ngay cả khi họ không trực tiếp tạo ra sản phẩm hay trải nghiệm.

Đặc điểm của loại tô này là không đến từ quyền sở hữu tài sản vật chất mà từ quyền kiểm soát luồng quyết định. Giá trị không bị lấy đi ở khâu sản xuất mà ở khâu lựa chọn. Khi AI trở thành người quyết định địa điểm nào được ghé thăm, dịch vụ nào được đặt và doanh nghiệp nào được ưu tiên, quyền lực kinh tế dịch chuyển từ người tạo ra giá trị sang người kiểm soát quá trình phân bổ giá trị.

Trước khi đi tiếp cần một cảnh báo về đo lường, vì đây là chỗ dễ sai nhất trong mọi phân tích về nền tảng. Doanh thu của một trung gian không phải là tô. Tỷ lệ hoa hồng cũng không phải là tô. Cả hai đại lượng ấy bao gồm chi phí thu hút khách, hạ tầng

công nghệ, phòng chống gian lận, chăm sóc khách hàng, bảo lãnh giao dịch, xử lý thanh toán và một mức lợi nhuận bình thường mà bất kỳ doanh nghiệp nào cũng cần để tồn tại. Tô chỉ là phần vượt lên trên tất cả những khoản đó.

Cuộc cạnh tranh quan trọng không chỉ là tạo ra trải nghiệm tốt hơn mà còn là giành quyền đứng tại các điểm nghẽn của chuỗi quyết định. Phần còn lại của mục này sẽ phân tích cách loại tô thuật toán hình thành, vì sao nó có khả năng tích lũy nhanh hơn nhiều hình thức tô truyền thống và những hệ quả mà nó tạo ra đối với địa phương, doanh nghiệp cũng như toàn bộ chuỗi giá trị của nền kinh tế trải nghiệm.

### 6.2.1. Điểm nghẽn tạo ra khả năng thu tô

Trong lịch sử kinh tế, những vị trí tạo ra dòng lưu chuyển luôn có khả năng tạo ra tô. Một cây cầu bắc qua con sông giúp rút ngắn quãng đường nên người qua lại sẵn sàng trả phí. Một cảng biển kết nối tuyến hàng hải quốc tế có thể thu phí bốc dỡ, lưu kho và dịch vụ hậu cần. Một cửa khẩu kiểm soát dòng thương mại có thể tạo ra nguồn thu từ mọi hàng hóa đi qua. Điểm chung của các trường hợp này là giá trị không nằm ở bản thân cây cầu hay cánh cổng mà nằm ở việc hầu hết dòng lưu chuyển đều phải đi qua chúng.

Trong kinh tế học, đây được gọi là quyền lực của điểm nghẽn (bottleneck). Ai kiểm soát điểm nghẽn không nhất thiết phải sản xuất nhiều hơn người khác nhưng vẫn có thể thu được một phần giá trị do toàn bộ hệ thống tạo ra. Càng nhiều người sử dụng tuyến đường đó, giá trị của điểm nghẽn càng tăng. Khi lưu lượng đủ lớn, vị trí ấy ngày càng khó bị thay thế.

Nền kinh tế số đã mở rộng khái niệm điểm nghẽn từ hạ tầng vật lý sang hạ tầng số. Các cổng thanh toán, kho ứng dụng, chợ trực tuyến hay mạng xã hội đều trở thành những cây cầu của thế giới số. Họ không trực tiếp sản xuất hàng hóa, nhưng kết nối người bán với người mua, nhà cung cấp với khách hàng và người tạo nội dung với người xem. Phần lớn giao dịch đi qua đều để lại một khoản phí, một phần dữ liệu hoặc một nguồn doanh thu.

Đến lượt mình, nền kinh tế tác nhân đẩy điểm nghẽn dịch chuyển lên một tầng cao hơn. Nếu trước đây người dùng tự đi qua cây cầu số để tìm kiếm và lựa chọn, thì nay chính các tác nhân AI là bên thực hiện hành trình ấy. Vì vậy, giá trị lớn nhất không còn nằm ở nơi giao dịch diễn ra mà ở nơi quyết định giao dịch sẽ diễn ra ở đâu. Điểm nghẽn của thời đại AI không còn là đường đi của con người mà là đường đi của quyết định.

Tô kinh tế có xu hướng tích tụ ở khúc hẹp nhất của chuỗi quyết định, không ở khâu tạo ra nhiều giá trị nhất. Độ hẹp được đo bằng chi phí để người mua đi vòng qua chủ thể ấy. Khi chi phí đi vòng đủ cao để cạnh tranh không còn tạo sức ép hiệu quả, phần thu của chủ thể kiểm soát điểm nghẽn chủ yếu bị giới hạn bởi pháp luật, phản ứng chính trị và khả năng xuất hiện công nghệ thay thế.

### 6.2.2. Thuật toán trở thành hạ tầng thu phí

Khi AI trở thành người đại diện cho người dùng, thuật toán không còn chỉ là công cụ tính toán mà trở thành một dạng hạ tầng kinh tế. Mỗi đề xuất được tạo ra là một lần thuật toán phân phối sự chú ý. Mỗi bảng xếp hạng là một lần thuật toán phân phối cơ hội kinh doanh. Mỗi lộ trình được lựa chọn là một lần

thuật toán phân phối dòng khách và dòng tiền. Chính quá trình phân phối ấy tạo ra khả năng thu phí. Khi người mua hoặc người bán khó đi vòng qua hệ thống, một phần khoản thu đó có thể chuyển thành tô.

Ở bước đề xuất, hệ thống xếp hạng và phân phối quyết định doanh nghiệp nào xuất hiện trước người dùng. Chỉ cần được đưa vào nhóm lựa chọn đầu tiên, xác suất được chọn đã tăng lên đáng kể. Khoản giá trị tăng thêm này không nhất thiết đến từ việc doanh nghiệp tốt hơn mà đến từ việc thuật toán đã trao cho họ lợi thế về khả năng được nhìn thấy.

Ở bước xếp hạng, tổ hợp của thuật toán, dữ liệu đầu vào, hợp đồng thương mại và quy tắc nền tảng quyết định thứ tự giữa các lựa chọn. Nghiên cứu về hành vi số cho thấy phần lớn người dùng chỉ xem vài kết quả đầu tiên. Vì vậy, việc đứng đầu danh sách mang lại lợi ích kinh tế rất lớn. Người kiểm soát cơ chế xếp hạng không cần sở hữu khách sạn, nhà hàng hay điểm tham quan, nhưng lại có thể ảnh hưởng trực tiếp đến doanh thu của tất cả các đơn vị đó.

Ở bước định tuyến hành trình, hệ thống có thể phối hợp nhiều quyết định trong cùng một kế hoạch. AI không chỉ chọn một khách sạn mà còn chọn chuyến bay, phương tiện di chuyển, nhà hàng, điểm tham quan và hình thức thanh toán. Khi nhiều quyết định được tích hợp trong cùng một hệ sinh thái, mỗi mắt xích đều có thể trở thành nơi tạo doanh thu thông qua hoa hồng, phí nền tảng, quảng cáo, dịch vụ tài chính hoặc khai thác dữ liệu.

Phần lớn các khoản thu này diễn ra gần như vô hình đối với người dùng. Họ chỉ nhìn thấy một câu trả lời được cho là tối ưu, trong khi phía sau là cả một chuỗi cơ chế lựa chọn, ưu tiên

và phân bổ giá trị. Thuật toán vì thế không chỉ là bộ não của hệ thống mà còn trở thành hạ tầng thu phí của nền kinh tế tác nhân. Bảng 6.1 đặt cạnh nhau mức thu phổ biến trên mỗi giao dịch tại một số điểm nghẽn.

**Bảng 6.1.** Mức thu phổ biến trên mỗi giao dịch tại một số điểm nghẽn

| Nền tảng | Mức thu phổ biến trên mỗi giao dịch |
|---|---|
| **Đặt phòng trực tuyến** | Hoa hồng 15 đến 25% giá trị đặt phòng.[1] *(Ghi chú: Có thể lên 25 đến 30% khi cơ sở tham gia chương trình hiển thị ưu tiên).* |
| **Booking Holdings** (tính gộp toàn tập đoàn) | Khoảng 14,5% năm 2025.[2] *(Ghi chú: Doanh thu 26,92 tỷ trên 186,1 tỷ đô la Mỹ giá trị đặt chỗ).* |
| **Kho ứng dụng di động** | 15 đến 30% giá trị giao dịch số.[3] *(Ghi chú: Có bậc giảm cho nhà phát triển nhỏ và cho năm thứ hai của thuê bao).* |
| **Tổ chức thẻ và cổng thanh toán** | Khoảng 2 đến 5%.[4] *(Ghi chú: Mức tham chiếu thường dùng khi so sánh với phí kho ứng dụng).* |
| **Kênh đặt trực tiếp của cơ sở lưu trú** | Chi phí biến đổi trên mỗi giao dịch thường vào khoảng 4–5% giá trị đặt phòng đối với khách sạn đã có trang web và hệ thống đặt phòng vận hành ổn định; chi phí chủ yếu gồm quảng cáo theo hiệu quả (performance marketing) và phí xử lý thanh toán. |

*Nguồn: xem danh mục cuối sách. Không con số nào trong bảng này là số đo của tô. Chúng là mức thu gộp, trong đó có phí dịch vụ, chi phí vận hành và lợi nhuận. Các mức cũng không so sánh trực tiếp được với nhau vì dịch vụ kèm theo khác nhau. Bảng chỉ nhằm cho thấy khoảng cách thứ tự độ lớn giữa phí hạ tầng và phí phân phối.*

1. Khoảng hoa hồng phổ biến của các đại lý du lịch trực tuyến, tổng hợp từ tài liệu công bố của Booking Holdings và Expedia Group cùng báo cáo ngành của Cloudbeds, Mews, Sojern và Hospitality Net. Mức thực tế thay đổi theo hợp đồng và mức độ hiển thị.

2. Báo cáo thường niên năm 2025 của Booking Holdings. Doanh thu 26,92 tỷ đô la Mỹ trên 186,1 tỷ đô la Mỹ giá trị đặt chỗ, tức khoảng 14,5% tính gộp toàn tập đoàn.

3. Điều khoản công bố của App Store và Google Play, cùng phân tích chính sách cạnh tranh của Brookings Institution, 2024.

4. Mức tham chiếu trong phân tích chính sách cạnh tranh của Brookings Institution, 2024, về phí của tổ chức thẻ và cổng thanh toán.

Khác với cây cầu hay cảng biển, hạ tầng này mở rộng dễ hơn nhiều. Suy luận của máy vẫn tiêu tốn điện, chip và tiền thật, đôi khi rất đáng kể, nên nói chi phí phục vụ thêm một triệu truy vấn bằng không là sai. Nhưng chi phí biên ấy thấp hơn hẳn so với việc mở thêm một cảng, đồng thời có xu hướng giảm theo quy mô. Chính khoảng cách giữa tốc độ tăng của doanh thu và tốc độ tăng của chi phí, chứ không phải con số tuyệt đối, là thứ khiến khoản thu ở lớp này tích lũy nhanh hơn các hình thức tô truyền thống.

Mức 15 đến 25% không phải là tô mà là giá của một dịch vụ thật.[1] Nền tảng bỏ tiền quảng cáo, xây hệ thống, bảo lãnh thanh toán và xử lý khiếu nại. Đúng. Một phần đáng kể khoản thu ấy là chi phí thật. Booking Holdings chi 8,19 tỷ đô la Mỹ cho tiếp thị trong năm 2025, tương đương khoảng 4,4% giá trị đặt chỗ. Nói cách khác, gần một phần ba doanh thu của nền tảng quay lại thị trường dưới dạng chi phí thu hút khách. Đó là dịch vụ, không phải tô. Phần doanh thu còn lại sau chi phí tiếp thị vẫn phải trang trải công nghệ, nhân sự, vận hành và nhiều khoản khác. Chỉ sau khi tính đủ những khoản ấy mới xuất hiện câu hỏi về lợi nhuận vượt chuẩn và về tô. Biên lợi nhuận trước lãi vay, thuế và khấu hao 36,9% ở một ngành mà khách sạn độc lập thường vận hành với biên thấp hơn nhiều là một chênh lệch cần giải thích. Cần nói rõ rằng, đây không phải phép so sánh đồng nhất về kế toán, vì một bên là tập đoàn toàn cầu theo mô hình phi tài sản còn bên kia là doanh nghiệp độc lập gánh tòa nhà và nhân sự. Con số chỉ cho thấy độ chênh của mô hình kinh doanh. Có hai

1. Booking Holdings, thông cáo kết quả kinh doanh quý IV và cả năm 2025, ngày 18 tháng 2 năm 2026, hồ sơ Form 8-K nộp Ủy ban Chứng khoán và Giao dịch Hoa Kỳ, https://www.sec.gov/Archives/edgar/data/1075531/000107553126000008/q4-25bkngearningsrelease.htm. Biên lợi nhuận trước lãi vay, thuế và khấu hao đã điều chỉnh năm 2025 là 36,9%, so với 35,0% năm 2024. Đây là chỉ tiêu phi chuẩn mực kế toán do doanh nghiệp tự định nghĩa.

cách giải thích và cả hai đều có căn cứ, hiệu quả vượt trội của mô hình phi tài sản, hoặc phần tô phát sinh từ vị trí. Cuốn sách này không có dữ liệu để tách hai phần ấy. Điều có thể nói là có đủ dấu hiệu để không thể giả định phần tô bằng không.

### 6.2.3. Phí dịch vụ, lợi nhuận và tô là ba thứ khác nhau

Một trong những thay đổi lớn nhất của nền kinh tế tác nhân là sự tách rời giữa người tạo ra giá trị và người thu được giá trị. Khách sạn tạo ra nơi lưu trú, nhà hàng tạo ra bữa ăn, cộng đồng tạo ra bầu không khí, nhưng một phần đáng kể doanh thu lại dừng ở lớp điều phối. Vấn đề là gọi toàn bộ phần dừng lại ấy là tô thì vừa sai về kinh tế học vừa làm luận điểm dễ bị bác bỏ. Cần tách làm ba tầng.

***Tầng thứ nhất*** là phí dịch vụ trung gian. Đây là khoản thu tương ứng với giá trị mà trung gian thật sự tạo ra: giảm chi phí tìm kiếm cho hàng triệu người, gộp cầu phân tán, bảo lãnh thanh toán, xử lý khiếu nại, chống gian lận. Không có lớp này thì thị trường du lịch quốc tế vận hành tệ hơn nhiều. Khoản thu ấy chính đáng.

***Tầng thứ hai*** là lợi nhuận nền tảng, tức phần còn lại sau khi trừ chi phí. Một mô hình phi tài sản, phục vụ toàn cầu từ một nền mã nguồn duy nhất, hoàn toàn có thể đạt biên lợi nhuận cao hơn một khách sạn phải nuôi tòa nhà và đội ngũ. Đó là hiệu quả của mô hình, không phải bằng chứng của quyền lực thị trường.

***Tầng thứ ba*** mới là tô thuật toán, phần lợi nhuận vượt trên mức mà cạnh tranh cho phép, phát sinh từ quyền kiểm soát điểm nghẽn, từ chi phí chuyển đổi mà người bán phải chịu, từ dữ

liệu độc quyền và từ khả năng loại một nhà cung cấp ra khỏi tầm nhìn của người mua. Dấu hiệu của nó là mức thu duy trì được lâu dài dù đối thủ có thể sao chép công nghệ, chứ không phải bản thân mức thu cao.

Các con số của Booking Holdings không đo trực tiếp tầng thứ ba. Chúng cho thấy quy mô giá trị đang được giữ lại ở lớp trung gian và đặt ra câu hỏi phần nào trong đó là giá của dịch vụ, phần nào là hiệu quả của mô hình, phần nào là tô phát sinh từ vị trí. Cuốn sách này không có dữ liệu để tách ba phần ấy. Có đủ dấu hiệu để không thể giả định phần thứ ba bằng không, đặc biệt khi cơ quan cạnh tranh ở nhiều nước đã can thiệp vào chính những điều khoản làm tăng chi phí chuyển đổi và hạn chế khả năng đi vòng. Việc can thiệp ấy chứng minh có lo ngại về quyền lực thị trường, nó không tự định lượng được phần tô.

Với địa phương, ý nghĩa thực tiễn của việc tách ba tầng nằm ở chỗ nó chỉ ra can thiệp nào có tác dụng. Không thể thương lượng để bỏ tầng một, vì đó là dịch vụ có thật. Khó tác động vào tầng hai, vì đó là hiệu quả. Tầng ba là đối tượng trực tiếp nhất của luật cạnh tranh. Tiêu chuẩn mở và quyền chuyển đổi dữ liệu còn có thể làm giảm chi phí đi vòng, tăng cạnh tranh và hạn chế khả năng phần lợi nhuận bình thường chuyển thành tô.

### 6.2.4. Dữ liệu và lãi kép của quyền lực

Tô thuật toán không chỉ hình thành từ quyền kiểm soát điểm nghẽn mà còn được khuếch đại bởi dữ liệu. Mỗi lần người dùng tìm kiếm, đặt dịch vụ, thanh toán hay đánh giá, hệ thống lại thu được thêm thông tin về hành vi và sở thích. Dữ liệu mới giúp thuật toán dự đoán chính xác hơn, từ đó tạo ra các đề xuất hiệu

quả hơn và thu hút thêm nhiều người dùng. Một vòng lặp tích lũy bắt đầu hình thành.

Ba cơ chế thường bị gộp làm một ở đây, nhưng chúng khác nhau. Hiệu ứng mạng lưới trực tiếp xảy ra khi giá trị với mỗi người dùng tăng lên vì có thêm người dùng khác, như trong một mạng xã hội. Hiệu ứng dữ liệu xảy ra khi thêm người dùng làm chất lượng dự đoán tốt lên, kể cả khi những người dùng ấy chẳng bao giờ gặp nhau. Lợi thế quy mô thuần túy chỉ là chuyện chi phí cố định được chia cho nhiều đơn vị hơn. Trong nền kinh tế tác nhân, cơ chế mạnh nhất là cơ chế thứ hai. Người dùng của một trợ lý du lịch không nhất thiết hưởng lợi trực tiếp chỉ vì có thêm người dùng khác, nhưng hệ thống có thể học từ lượng quyết định lớn hơn và thu hút thêm nhà cung cấp. Chất lượng đề xuất tăng, thêm người dùng, thêm dữ liệu. Đó là lãi kép.

Quan trọng hơn, dữ liệu không chỉ cải thiện chất lượng của thuật toán mà còn củng cố chính quyền lực của nó. Khi hệ thống liên tục được lựa chọn, nó thu thêm dữ liệu. Dữ liệu mới giúp hệ thống tiếp tục được lựa chọn. Vòng phản hồi tích cực (feedback loop) này có thể chuyển lợi thế ban đầu thành vị thế dẫn đầu khó bị thu hẹp.

Đó là lý do các nền tảng AI lớn thường có xu hướng ngày càng mạnh hơn theo thời gian. Họ không chỉ tích lũy doanh thu mà còn tích lũy dữ liệu, cải thiện thuật toán, mở rộng mạng lưới và củng cố khả năng kiểm soát các điểm nghẽn của nền kinh tế tác nhân. Đây là một dạng lãi kép của quyền lực số mà các hình thức tô truyền thống rất khó đạt được.

Lợi thế của chủ thể điều phối tăng mạnh theo lượng quyết định nó quan sát được, không chỉ theo lượng tài sản vật chất

nó sở hữu. Vì mỗi quyết định vừa tạo doanh thu vừa tạo dữ liệu để cải thiện quyết định sau, lợi thế tích lũy theo cơ chế lãi kép. Mệnh đề này có điều kiện dừng: nó chỉ đúng khi dữ liệu mới còn mang thông tin mới. Ở những thị trường mà sở thích thay đổi chậm và tập lựa chọn nhỏ, lợi ích biên của dữ liệu giảm nhanh và lợi thế bão hòa.

Nhà phản biện sẽ cho rằng, nếu tô của nền tảng cao đến vậy, thị trường sẽ tự sửa. Sẽ có nhiều tác nhân AI cạnh tranh nhau, phí giảm xuống và địa phương hưởng lợi: Đây là phản bác mạnh nhất trong chương và nó chưa được bác bỏ hoàn toàn. Cạnh tranh giữa các tác nhân là có thật, chi phí suy luận đang giảm nhanh. Một tác nhân thu phí quá cao sẽ mất người dùng. Nhưng có hai lý do khiến cơ chế tự sửa này chậm hơn kỳ vọng. Thứ nhất, cạnh tranh diễn ra ở phía người dùng cuối, nơi dịch vụ thường miễn phí, chứ không ở phía nhà cung cấp, nơi phí được thu. Người dùng chọn tác nhân theo chất lượng câu trả lời, không theo mức phí mà khách sạn phải trả. Thứ hai, nhiều tác nhân cùng lấy dữ liệu và tồn kho từ một số ít nhà tổng hợp, nên cạnh tranh ở lớp giao diện không nhất thiết tạo ra cạnh tranh ở lớp phân phối. Bằng chứng đối chiều cũng tồn tại, khi Đạo luật Thị trường Số của Liên minh châu Âu buộc Apple hạ mức thu ở châu Âu, một nghiên cứu do chính Apple công bố tháng 11 năm 2025, vì vậy cần đọc với thận trọng, cho thấy hơn 90% nhà phát triển chuyển sang điều khoản mới đã không chuyển phần tiết kiệm ấy sang giá bán.[1] Phần tô giảm ở một mắt xích không tự động chảy về mắt xích khác. Tô không mọc lên ở nơi giá trị được tạo ra. Nó mọc lên ở nơi giá trị buộc phải đi qua.

1. Apple, nghiên cứu công bố tháng 11 năm 2025 về tác động của Đạo luật Thị trường Số tới giá ứng dụng tại châu Âu. Nghiên cứu do một bên có lợi ích trực tiếp công bố trong khi đang tranh tụng về chính đạo luật được nghiên cứu, nên chỉ dùng làm bằng chứng đối chiều.

### 6.2.5. Tô thuật toán khác gì tô đất đai

Tô đất đai và tô thuật toán đều là khoản thu phát sinh từ việc kiểm soát một nguồn lực khan hiếm. Tuy nhiên, bản chất của nguồn lực đó hoàn toàn khác nhau. Nếu tô đất đai dựa trên quyền sở hữu một vị trí vật lý thì tô thuật toán dựa trên quyền kiểm soát quá trình phân phối trong không gian số.

Đất đai là tài nguyên hữu hạn. Một khu đất chỉ có thể thuộc về một chủ sở hữu tại một thời điểm và giá trị của nó gắn với vị trí địa lý cụ thể. Người sở hữu khu đất có thể thu tô vì mọi hoạt động muốn diễn ra tại vị trí đó đều phải trả chi phí sử dụng.

Tô thuật toán không cần chiếm hữu không gian vật lý. Thứ được kiểm soát là quyền phân phối. Khi một hệ thống AI quyết định khách sạn nào được giới thiệu, nhà hàng nào được đặt bàn hay tuyến du lịch nào được lựa chọn, nó đang kiểm soát quyền tiếp cận khách hàng. Giá trị kinh tế phát sinh không phải vì sở hữu địa điểm mà vì sở hữu quyền quyết định ai sẽ được nhìn thấy trước.

Một mảnh đất chỉ có thể phục vụ số lượng người nhất định trong cùng một thời điểm. Ngược lại, một thuật toán có thể đồng thời ảnh hưởng đến hàng triệu quyết định trên phạm vi toàn cầu. Chi phí phục vụ thêm một truy vấn vẫn tồn tại, nhưng thấp hơn nhiều so với việc mở rộng một hạ tầng vật chất tương đương, đồng thời có khả năng giảm theo quy mô. Chính vì vậy, khả năng mở rộng của tô thuật toán lớn hơn rất nhiều so với tô đất đai.

Khác biệt thứ hai nằm ở tốc độ tích lũy quyền lực. Giá trị của đất đai thường tăng theo sự phát triển của khu vực xung quanh

và chịu nhiều giới hạn về diện tích, hạ tầng và quy hoạch. Trong khi đó, tô thuật toán tăng theo lượng dữ liệu quan sát được, theo hiệu ứng dữ liệu và theo khả năng học của hệ thống. Càng nhiều người sử dụng hệ thống, thuật toán càng chính xác. Càng chính xác, hệ thống càng được sử dụng nhiều hơn. Quyền lực được tích lũy theo cơ chế lãi kép của dữ liệu chứ không phải theo sự khan hiếm của tài nguyên vật chất.

Điều quan trọng hơn là tô đất đai thường gắn với một lãnh thổ cụ thể, còn tô thuật toán có thể vượt qua mọi biên giới. Một nền tảng đặt tại một quốc gia vẫn có thể kiểm soát việc lựa chọn khách sạn, nhà hàng và điểm tham quan ở hàng trăm quốc gia khác. Giá trị được tạo ra tại địa phương nhưng một phần lợi ích kinh tế có thể được chuyển đến nơi đặt nền tảng hoặc nơi kiểm soát thuật toán. Vì vậy, nếu nhiều hình thức tô của thời đại công nghiệp phát sinh từ quyền kiểm soát đất đai, cảng, tuyến vận tải hoặc hạ tầng vật chất khan hiếm, thì tô của nền kinh tế tác nhân là phần thu từ việc kiểm soát dòng quyết định, dòng dữ liệu và dòng phân phối. Không cần sở hữu nơi chốn, người kiểm soát thuật toán vẫn có thể sở hữu con đường dẫn đến nơi chốn ấy. Và trong nhiều trường hợp, quyền kiểm soát con đường còn tạo ra giá trị kinh tế lớn hơn cả việc sở hữu chính điểm đến.

**Khách sạn M**

Để các cơ chế của chương không dừng ở tầng khái niệm, phần còn lại sẽ theo một trường hợp duy nhất. Khách sạn M là một cơ sở lưu trú hai mươi sáu phòng ở một thị xã ven biển miền Trung. Đây là trường hợp dựng lại từ nhiều cơ sở có thật, không phải một doanh nghiệp cụ thể. Các con số là giả định để tính toán chứ không phải số liệu khảo sát. Giá phòng bình quân 1,2 triệu đồng một đêm, công suất 62%, doanh thu phòng khoảng

7 tỷ đồng một năm.[1] Bảy mươi phần trăm lượt đặt đến qua nền tảng, với mức thu thực khoảng 18% trên phần doanh thu đó, đã gồm khuyến mại bắt buộc nhưng chưa gồm khoản chi riêng cho chương trình hiển thị ưu tiên. Như vậy mỗi năm khoảng 889 triệu đồng doanh thu phòng được giữ lại ở lớp phân phối trước khi phần còn lại đi vào tiền lương, thuế và tái đầu tư tại địa phương. Một phần khoản này là giá của dịch vụ có thật. Câu hỏi của các mục tiếp theo là phần nào phản ánh chi phí dịch vụ, phần nào phản ánh hiệu quả của mô hình, phần nào có thể phát sinh từ vị trí điểm nghẽn.

1. Giả định của tác giả. Trường hợp trong tiết này do tác giả dựng lại, không phải doanh nghiệp có thật. Các con số dưới đây không phải dữ liệu của một cơ sở hay một địa phương cụ thể và không nên đọc như khuyến nghị định mức.

## 6.3. CHI PHÍ THAM GIA NỀN KINH TẾ TÁC NHÂN

Trong nền kinh tế truyền thống, doanh nghiệp và địa phương đầu tư để tạo ra giá trị mới: xây dựng hạ tầng, nâng cao chất lượng dịch vụ, phát triển sản phẩm hay cải thiện trải nghiệm của người dân và du khách. Chi phí ấy, dù lớn đến đâu, vẫn trực tiếp làm tăng năng lực cạnh tranh hoặc chất lượng sống. Trong nền kinh tế tác nhân, xuất hiện thêm một nhóm chi phí hoàn toàn khác: những khoản chi không làm trải nghiệm tốt hơn, nhưng nếu không chi thì địa phương có thể biến mất khỏi tầm nhìn của máy.

Đó là chi phí tham gia nền kinh tế tác nhân. Địa phương phải chuẩn hóa thực thể, duy trì nội dung và bằng chứng, đồng bộ dữ liệu vận hành, thiết lập kết nối giao dịch, giám sát sai lệch và bảo đảm hệ thống được cập nhật liên tục. Một phần các khoản đầu tư này giúp máy đọc đúng. Một phần khác giúp máy kiểm tra tình trạng và hành động được. Chúng không nhất thiết tạo ra một trải nghiệm mới cho du khách, nhưng quyết định địa phương có còn nằm trong tập phương án mà máy có thể xem xét và thực hiện hay không.

Có thể gọi đây là chi phí tồn tại số (cost of digital existence). Cũng giống như doanh nghiệp phải có mặt bằng và hệ thống vận hành trước khi bán sản phẩm đầu tiên, địa phương phải vượt qua một ngưỡng năng lực số trước khi có cơ hội được tác nhân AI đề xuất và giao dịch.[1] Khoản chi này không bảo đảm chiến thắng, nhưng thiếu nó có thể khiến địa phương bị loại trước khi chất lượng trải nghiệm được đem ra so sánh.

1. Đề xuất của tác giả. Các ngưỡng trong bảng chỉ số giữ giá trị do tác giả đề xuất, chưa được kiểm định thực nghiệm.

Vì vậy, cạnh tranh giữa các địa phương không còn chỉ là cuộc đua tạo ra trải nghiệm hấp dẫn hơn. Đó còn là cuộc đua duy trì khả năng được máy nhìn thấy. Trong nhiều trường hợp, khoản đầu tư đầu tiên không phải để được chọn, mà chỉ để không bị loại khỏi tập lựa chọn của các tác nhân AI. Đây chính là loại chi phí mới mà quản trị địa phương trong kỷ nguyên AI buộc phải tính đến. Bảng 6.2 tách khoản chi ấy thành các cấu phần.

**Bảng 6.2.** Cấu phần chi phí tham gia nền kinh tế tác nhân

| Lớp | Nhóm chi phí | Nội dung | Tính chất |
|---|---|---|---|
| Khả đọc | Định danh và lược đồ | Mã định danh thực thể, chuẩn mô tả, phân loại, ranh giới hành chính | Đầu tư một lần, duy trì thấp |
| Khả đọc | Nội dung và bằng chứng | Mô tả đa ngữ, ảnh, tư liệu di sản, chứng nhận, dữ liệu mở | Đầu tư lớn ban đầu, cập nhật theo chu kỳ |
| Khả đọc | Giám sát và sửa sai | Theo dõi cách máy mô tả địa phương, phát hiện và yêu cầu đính chính sai lệch | Tăng theo số nền tảng |
| Khả hành động | Dữ liệu vận hành | Tồn kho, giá, lịch, tình trạng theo thời gian thực | Chi phí thường xuyên, không thể ngắt |
| Khả hành động | Kết nối và giao dịch | Giao diện lập trình ứng dụng, đặt chỗ, xác nhận, thanh toán, hoàn và hủy | Chi phí thường xuyên, cộng phí nền tảng |
| Cạnh tranh phân phối | Hiển thị ưu tiên | Phí xếp hạng, chương trình đối tác, khuyến mại để giữ thứ hạng | Tùy chọn nhưng khó rút lui |
| Nền | Năng lực con người | Nhân sự dữ liệu, quy trình, thẩm quyền phê duyệt | Chi phí cố định, khó thuê ngoài bền vững |

*Nguồn: Tác giả xây dựng.*

### 6.3.1. Ba lớp chi phí tham gia

Chi phí tham gia nền kinh tế tác nhân gồm ba lớp khác nhau. Lớp thứ nhất là chi phí khả đọc, để máy nhận diện và hiểu đúng địa phương. Lớp thứ hai là chi phí khả hành động, để máy kiểm tra tình trạng và hoàn tất giao dịch. Lớp thứ ba là chi phí cạnh tranh phân phối, để một phương án được ưu tiên trong quá trình lựa chọn. Ba lớp có quan hệ với nhau nhưng không nên bị gộp thành một khái niệm duy nhất.

Lớp khả đọc gồm những khoản phải trả để máy nhận diện địa phương đúng là nó. Ở đây có lược đồ dữ liệu (schema), tức chuẩn mô tả để máy hiểu một địa điểm là gì và có thuộc tính nào. Có đồ thị tri thức (knowledge graph), tức mạng lưới nối con người, nơi chốn, sự kiện và tài sản thành một cấu trúc có thể suy luận. Có định danh thực thể để tên một ngôi làng không bị nhầm với tên một khách sạn cùng tên. Thiếu lớp này, vấn đề không phải địa phương bị đánh giá thấp, mà là nó có thể không được nhận diện như một phương án độc lập trong tập ứng viên.

Lớp thứ hai là chi phí khả hành động, tức khoản phải trả để máy kiểm tra được tình trạng thật và hoàn tất được giao dịch. Đây là một năng lực khác với khả đọc và sẽ là khái niệm trung tâm của Chương 7. Đây là giao diện lập trình ứng dụng (API), dữ liệu tồn kho theo thời gian thực, khả năng xác nhận và thanh toán. Một tác nhân được yêu cầu tự hoàn tất đặt chỗ có xu hướng loại bỏ hoặc hạ ưu tiên những phương án mà nó không thể xác nhận và giao dịch trực tiếp, dù phương án ấy có thể hấp dẫn hơn.

Lớp thứ ba là chi phí cạnh tranh phân phối, tức khoản trả để được ưu tiên hiển thị, có thứ hạng tốt hơn và tỷ lệ chuyển đổi cao hơn. Lớp này về bản chất gần với quảng cáo. Gộp nó vào hai lớp

trên sẽ làm khái niệm chi phí khả đọc phình ra và mất sức thuyết phục, vì khi đó mọi khoản chi tiếp thị đều có thể được gọi là chi phí tồn tại.

Ranh giới quan trọng nằm giữa lớp hai và lớp ba. Hai lớp đầu là điều kiện để được xét. Lớp thứ ba là nỗ lực để thắng trong lần xét ấy. Hai lớp đầu có tính chất hạ tầng và là đối tượng ưu tiên của đầu tư công. Lớp thứ ba gần với marketing và cạnh tranh thương mại hơn, nên về nguyên tắc phải do chủ thể hưởng lợi trực tiếp chi trả.

**Khách sạn M: ba lớp chi phí**

Với khách sạn M, ba lớp có độ lớn rất khác nhau. Chi phí khả đọc, gồm chuẩn hóa thông tin cơ sở, ảnh, mô tả song ngữ, định danh trên bản đồ số và công theo dõi xem máy đang mô tả mình ra sao, vào khoảng 40 triệu đồng mỗi năm. Chi phí khả hành động, gồm phần mềm quản lý phòng có kết nối, đồng bộ tồn kho theo thời gian thực và cổng thanh toán, khoảng 120 triệu đồng mỗi năm. Chi phí cạnh tranh phân phối, tức khoản trả riêng cho chương trình hiển thị ưu tiên, khoảng một 150 triệu đồng. Khoản này nằm ngoài mức hoa hồng 18% chứ không nằm trong đó, nên không bị tính hai lần.[1]

Giả sử một khách sạn có 26 phòng, đạt công suất khai thác bình quân 62% và giá phòng bình quân (ADR) 1,2 triệu đồng/đêm. Với mức vận hành này, doanh thu phòng ước đạt khoảng 7,06 tỷ đồng mỗi năm, được xem là 100% doanh thu phòng để làm cơ sở phân tích các khoản chi phí và dòng giá trị trong nền

1. Tính toán của tác giả trên một kịch bản dựng để minh họa cơ chế, không phải số liệu quan sát của một doanh nghiệp hay địa phương có thật.

kinh tế tác nhân. Bảng 6.3 phân tách ba lớp chi phí ấy theo tỷ trọng trên doanh thu phòng.[1]

**Bảng 6.3.** Khách sạn M, ba lớp chi phí tính trên doanh thu phòng

| Khoản mục | Giá trị | Tỷ lệ doanh thu phòng |
|---|---|---|
| Doanh thu phòng | 7,06 tỷ đồng | 100,0% |
| Qua nền tảng (70%) | 4,94 tỷ đồng | 70,0% |
| Qua kênh trực tiếp (30%) | 2,12 tỷ đồng | 30,0% |
| Hoa hồng nền tảng (18% doanh thu OTA) | 889 triệu đồng | 12,6% |
| Chi phí giao dịch trực tiếp (quảng cáo, thanh toán...) | 106 triệu đồng | 1,5% |
| Chi phí khả đọc (dữ liệu, nội dung, cập nhật...) | 40 triệu đồng | 0,6% |
| Chi phí khả hành động (API, booking engine, kết nối...) | 120 triệu đồng | 1,7% |
| Chi phí cạnh tranh phân phối (khuyến mại, ưu tiên hiển thị...) | 150 triệu đồng | 2,1% |
| Tổng chi phí phân phối và tham gia | 1,31 tỷ đồng | 18,5% |

*Nguồn của toàn bảng: kịch bản và tính toán minh họa của tác giả, dựng lại từ nhiều cơ sở có thật. Không phải dữ liệu của một doanh nghiệp cụ thể và không nên đọc như khuyến nghị định mức.*

1. Giả định của tác giả. Toàn bộ cân đối tài chính của khách sạn M trong tiết này là mô hình do tác giả dựng, không phải số liệu kế toán của một cơ sở có thật. Các giả định được nêu đầy đủ ở đầu tiết.

Kịch bản do tác giả dựng, không phải số liệu của một doanh nghiệp có thật. Ba dòng cuối là chi phí độc lập với hoa hồng, đã kiểm tra để không tính trùng. Con số đáng chú ý là tỷ trọng của một kênh duy nhất, không phải tổng 18,5%.[1]

Điều đáng làm với một cân đối như vậy là thử một kịch bản. Giả sử khách sạn M kéo tỷ trọng nền tảng từ 70% xuống 55% mà vẫn giữ nguyên công suất. Hoa hồng giảm còn 699 triệu, tiết kiệm 190 triệu. Chi phí biến đổi của kênh trực tiếp tăng từ 106 lên 159 triệu, tức thêm 53 triệu.[2] Nhưng để giữ được lượng đặt trực tiếp cao hơn, M phải bước lên một bậc chi phí cố định tăng thêm, khoảng 80 triệu đồng mỗi năm cho nâng cấp hệ thống đặt phòng trực tiếp, hệ thống quản lý quan hệ khách hàng, nhân sự chăm sóc khách và sản xuất nội dung. Khoản này nằm ngoài mức 5% biến đổi và không trùng với nó. Tiết kiệm ròng vì thế còn khoảng 57 triệu, tương đương 0,8% doanh thu.

Con số nhỏ ấy mới là bài học. Bỏ nền tảng gần như không bao giờ là câu trả lời, vì nền tảng thật sự mang lại khách mà một cơ sở hai mươi sáu phòng không tự tìm được. Thứ thay đổi trong kịch bản không nằm ở 57 triệu đồng tiết kiệm được, mà ở chỗ doanh thu trực tiếp tăng từ 2,12 lên 3,18 tỷ đồng. Nếu M thiết kế đúng quy trình xin đồng ý và chăm sóc sau giao dịch, phần doanh thu tăng thêm ấy còn có thể đi kèm dữ liệu khách hàng và quyền liên hệ lại hợp pháp. Vấn đề chưa bao giờ là rời khỏi kênh phân phối. Vấn đề là không để một kênh duy nhất sở hữu toàn bộ khách hàng và dữ liệu.

1. Tính toán của tác giả, không phải số liệu quan sát của một doanh nghiệp hay địa phương có thật.

2. Ước lượng ngành về chi phí biến đổi trên mỗi giao dịch của kênh đặt trực tiếp, gồm quảng cáo theo chuyển đổi và phí thanh toán. Mức này được nêu trong tài liệu của bên bán giải pháp phân phối, nên cần đọc thận trọng vì bên công bố có lợi ích liên quan.

**Khách sạn M: câu hỏi này trả lời được đến đâu**

Trường hợp khách sạn M không cho phép tính phần tô. Muốn tách được tầng thứ ba khỏi hai tầng đầu thì cần biết cấu trúc chi phí thật của nền tảng, mức lợi nhuận cạnh tranh của ngành và chi phí đi vòng thật sự của người mua, cả ba đều nằm ngoài tầm với của một cơ sở lưu trú và của cuốn sách này. Cái mà trường hợp này làm được là chỉ ra bốn thứ cần đo và có thể đo: chi phí phân phối thực tính đủ, khả năng đi vòng sang kênh khác, quyền giữ dữ liệu sau mỗi giao dịch, cuối cùng là mức phụ thuộc vào một nền tảng duy nhất.

### 6.3.2. Không đầu tư cũng mất thị trường

Một trong những đặc điểm nguy hiểm nhất của chi phí tham gia là nó hiếm khi xuất hiện trên báo cáo tài chính. Địa phương không nhận được một hóa đơn ghi rằng mình đã mất khách vì thiếu dữ liệu hay vì API không hoạt động. Thị trường đơn giản là chuyển sang nơi khác, trong khi nguyên nhân thực sự gần như vô hình. Đây chính là chi phí vô hình của nền kinh tế tác nhân.

Khi một hệ thống AI không thể xác minh thông tin, không truy cập được dữ liệu thời gian thực hoặc không thực hiện được giao dịch, hệ thống có xu hướng ưu tiên một phương án khác có dữ liệu đầy đủ hơn hoặc có khả năng hoàn tất giao dịch đáng tin cậy hơn. Việc loại bỏ ấy diễn ra tự động và hiếm khi được giải thích cho người quản lý địa phương. Thường không có khiếu nại, không có cảnh báo, cũng không có cơ hội phản biện. Chỉ có dòng khách, dòng tiền và sự chú ý âm thầm chuyển hướng.

Hệ quả lớn hơn là chi phí cơ hội. Mỗi lần AI không đề xuất một địa phương là một chuỗi giá trị tiềm năng không bao giờ được

hình thành: du khách không đến, doanh nghiệp không bán được hàng, dữ liệu mới không được tạo ra, đánh giá không xuất hiện, danh tiếng không tích lũy và những lần đề xuất sau lại càng khó xảy ra. Một quyết định không được chọn hôm nay có thể làm mất nhiều quyết định khác trong tương lai thông qua các vòng phản hồi của thuật toán.

Khoản mất mát này không chỉ là một giao dịch. Một lần không được chọn cũng đồng nghĩa không có đánh giá mới, không có dữ liệu mới và không có bằng chứng mới để được chọn ở vòng sau. Cần phân biệt ba thứ ở đây: chi phí kế toán là khoản đã chi và ghi được vào sổ, chi phí cơ hội là doanh thu của lần không được chọn, còn tổn thất động là phần mất đi qua vòng phản hồi của thuật toán. Chỉ loại thứ nhất xuất hiện trên báo cáo. Hai loại sau lớn dần theo thời gian mà không nằm trên bất kỳ dòng ngân sách nào.

### 6.3.3. Chi phí này mang tính lũy thoái

Chi phí tham gia không phân bổ đồng đều giữa các địa phương. Về hình thức, mọi nơi đều phải thực hiện những công việc tương tự: chuẩn hóa dữ liệu, định danh thực thể, duy trì nội dung, kết nối nền tảng, xây dựng giao diện lập trình ứng dụng và cập nhật thông tin theo thời gian thực. Nhưng cùng một khoản đầu tư lại tạo ra gánh nặng rất khác nhau tùy theo quy mô kinh tế, năng lực ngân sách và số lượng giao dịch của từng địa phương.

Một thành phố lớn có thể phân bổ chi phí xây dựng hạ tầng dữ liệu cho hàng triệu cư dân, hàng chục nghìn doanh nghiệp và một lượng lớn khách hàng. Chi phí trên mỗi giao dịch vì thế giảm nhanh khi quy mô tăng lên. Ngược lại, một thị trấn nhỏ,

một làng nghề hay một điểm đến mới nổi vẫn phải xây dựng những lớp hạ tầng tối thiểu tương tự, nhưng chỉ có một lượng nhỏ người dùng và giao dịch để phân bổ chi phí. Tính theo tỷ lệ ngân sách hoặc giá trị kinh tế tạo ra, địa phương nhỏ phải trả đắt hơn.

Đây là tính lũy thoái của chi phí tham gia: chủ thể có ít nguồn lực nhất lại phải dành tỷ lệ nguồn lực lớn nhất chỉ để đạt điều kiện tham gia tối thiểu.[1] Một địa phương lớn có thể xem việc xây dựng kho dữ liệu, bản đồ số hay hệ thống đặt dịch vụ là một hạng mục trong ngân sách chuyển đổi số. Với một địa phương nhỏ, cùng những hạng mục ấy có thể chiếm phần đáng kể ngân sách phát triển.

Bất lợi không chỉ nằm ở tiền. Địa phương nhỏ còn thiếu nhân lực dữ liệu, năng lực kỹ thuật, kinh nghiệm quản trị tiêu chuẩn và khả năng đàm phán với nền tảng. Họ thường phải thuê ngoài với chi phí cao, phụ thuộc vào nhà cung cấp và khó duy trì hệ thống sau khi dự án kết thúc. Trong khi đó, các đô thị lớn có thể xây dựng đội ngũ nội bộ, dùng chung hạ tầng giữa nhiều ngành và liên tục cải thiện dữ liệu theo thời gian.

Vì vậy, chi phí tham gia có nguy cơ làm sâu thêm khoảng cách giữa trung tâm và ngoại vi. Nơi đã có nhiều dữ liệu, nhiều giao dịch và nhiều năng lực kỹ thuật sẽ càng dễ trở nên khả đọc. Nơi ít nguồn lực lại phải trả nhiều hơn để đạt cùng một mức hiện diện. Nếu không có hạ tầng dùng chung, tiêu chuẩn mở và cơ chế hỗ trợ, nền kinh tế tác nhân có thể tạo ra một nghịch lý: những địa phương cần được nhìn thấy nhất lại là những nơi khó trả nổi chi phí để được máy nhìn thấy.

1. Phép tính minh hoạ do tác giả dựng để làm rõ cơ chế, không phải số liệu quan sát.

Chi phí tham gia nền kinh tế tác nhân có phần cố định lớn và ít co giãn theo quy mô. Vì vậy, tính theo tỷ lệ ngân sách hoặc theo giá trị kinh tế tạo ra, gánh nặng rơi nặng nhất lên chủ thể nhỏ nhất. Nếu không có hạ tầng dùng chung, khả đọc sẽ khuếch đại chênh lệch sẵn có giữa trung tâm và ngoại vi thay vì thu hẹp nó.

Điều đáng chú ý là nhóm năng lực tối thiểu mà mọi địa phương cần có gần như không thay đổi, bất kể quy mô lớn hay nhỏ. Một đô thị lớn, một huyện lỵ trung bình hay một xã có nghề truyền thống đều phải xây dựng được bốn năng lực nền tảng: định danh thực thể, tổ chức dữ liệu, cập nhật thường xuyên và quản trị dữ liệu.

Khác biệt nằm ở quy mô và chi phí. Đô thị lớn có thể phải chuẩn hóa hàng chục nghìn thực thể, trong khi huyện lỵ chỉ khoảng vài trăm và một xã có nghề truyền thống chỉ vài chục. Tuy nhiên, nhờ quy mô lớn, chi phí bình quân trên mỗi thực thể của đô thị lại thấp hơn nhiều. Ngược lại, ở các địa phương nhỏ, chi phí trên mỗi thực thể rất cao, khiến khoản đầu tư này có thể chiếm một tỷ trọng đáng kể, thậm chí vượt quá khả năng ngân sách.

Sự khác biệt còn nằm ở năng lực duy trì sau khi dự án kết thúc. Các đô thị lớn thường có đội ngũ chuyên trách để tiếp tục cập nhật và quản trị dữ liệu. Huyện lỵ vẫn phụ thuộc khá nhiều vào nhà thầu hoặc đơn vị hỗ trợ. Còn ở nhiều xã, hệ thống thường dừng lại sau khi nghiệm thu vì thiếu nhân lực, kinh phí và quy trình vận hành lâu dài. Điều này cho thấy thách thức của các địa phương nhỏ nằm ở việc duy trì dữ liệu luôn chính xác và cập nhật theo thời gian, không ở việc xây dựng dữ liệu ban đầu.

Chuyên gia phản biện sẽ cho rằng, địa phương nhỏ không cần lo. Khách đến làng nghề chủ yếu bằng truyền miệng và mạng xã hội, không qua trợ lý AI: Điều này đúng ở hiện tại và

có thể sai dần theo thời gian. Tính lũy thoái không đòi hỏi tác nhân AI chiếm đa số lượt khách. Chỉ cần nhóm khách phụ thuộc mạnh vào các kênh số có mức chi tiêu cao hơn bình quân, tác động doanh thu đã có thể lớn hơn tỷ trọng số lượt. Đây là một khả năng cần đo, không phải kết luận rút ra từ số liệu hiện có. Sáu tháng đầu năm 2026, Việt Nam đón 12,3 triệu lượt khách quốc tế và khoảng 81 triệu lượt khách nội địa, với tổng thu du lịch khoảng 569 nghìn tỷ đồng. Trong một kịch bản minh hoạ, không phải một suy luận rút ra từ hai con số trên, giả sử nhóm khách đến qua kênh phụ thuộc khả đọc chỉ chiếm một phần năm số lượt nhưng chi tiêu gấp đôi mức bình quân. Khi đó phần doanh thu phụ thuộc khả đọc đã là một phần ba. Ngưỡng nguy hiểm đến trước khi tỷ lệ lượt khách kịp báo động.[1] Địa phương không trả tiền để được chọn. Địa phương trả tiền để không bị loại trước khi được xem xét.

### 6.3.4. Từ marketing sang hạ tầng dữ liệu

Trong nhiều thập niên, địa phương tiếp cận thị trường chủ yếu bằng marketing. Họ xây dựng khẩu hiệu, tổ chức sự kiện, mua quảng cáo, mời báo chí, sản xuất video và triển khai các chiến dịch truyền thông. Logic rất rõ ràng: muốn có thêm khách, phải làm cho nhiều người biết đến hơn. Nhưng khi quyền tìm kiếm và lựa chọn dần được chuyển giao cho các tác nhân AI, mức độ nhận biết của con người không còn đủ để bảo đảm một địa phương được chọn. Một điểm đến có thể nổi tiếng nhưng vẫn bị bỏ qua nếu dữ liệu phân mảnh, thông tin mâu thuẫn, khả năng đặt dịch vụ không rõ ràng hoặc hệ thống không thể xác minh tình trạng

1. Ngưỡng do tác giả đề xuất, chưa được kiểm định thực nghiệm. Con số chỉ nhằm nêu bậc độ lớn và cần được hiệu chỉnh theo dữ liệu của từng địa phương.

thực tế. Ngược lại, một nơi ít nổi tiếng hơn nhưng có dữ liệu đầy đủ, cấu trúc rõ và giao dịch thuận tiện có thể được AI ưu tiên.

Vì thế, ngân sách cạnh tranh phải dịch chuyển từ truyền thông thuần túy sang hạ tầng dữ liệu. Địa phương không chỉ cần kể câu chuyện hấp dẫn, mà còn phải chuyển câu chuyện ấy thành những thực thể, thuộc tính, quan hệ và dịch vụ mà máy có thể đọc. Không chỉ giới thiệu rằng một lễ hội đáng trải nghiệm, mà phải cung cấp thời gian, địa điểm, lịch trình, giá vé, năng lực phục vụ, phương thức di chuyển và khả năng đặt chỗ. Không chỉ nói một làng nghề có bản sắc, mà phải định danh các nghệ nhân, sản phẩm, không gian trải nghiệm, giờ mở cửa và điểm giao dịch.

Marketing tạo ra sự chú ý. Hạ tầng dữ liệu tạo ra khả năng được lựa chọn. Hai chức năng này có liên hệ nhưng không thể thay thế nhau. Quảng cáo có thể khiến người dùng hỏi về một địa phương, nhưng dữ liệu mới quyết định liệu AI có thể đưa địa phương ấy vào câu trả lời hay không. Truyền thông có thể tạo nhu cầu, nhưng API, dữ liệu thời gian thực và quy trình giao dịch mới giúp nhu cầu được chuyển thành hành động.

Trong nền kinh tế tác nhân, địa phương không thể chỉ mua quyền xuất hiện trước mắt con người. Họ phải xây dựng năng lực xuất hiện trong quá trình suy luận của máy. Marketing giúp địa phương được biết đến. Hạ tầng dữ liệu giúp địa phương có thể được chọn.

> ▪ ***Về ngân sách:***
>
> • ***Mục tiêu của hai loại ngân sách hoàn toàn khác nhau:*** *Ngân sách truyền thông nhằm giúp địa phương được con người biết đến, còn ngân sách khả đọc nhằm giúp máy nhận diện đúng, suy luận đúng và có thể thực hiện giao dịch.*

- ***Chu kỳ đầu tư cũng khác nhau:*** *Ngân sách truyền thông thường vận hành theo từng chiến dịch, có thể tăng giảm hoặc tạm dừng. Ngược lại, ngân sách khả đọc là khoản đầu tư liên tục. Nếu ngừng cập nhật, dữ liệu nhanh chóng mất giá trị và hệ thống sẽ suy giảm khả năng hoạt động.*
- ***Chỉ số đánh giá không giống nhau:*** *Truyền thông đo bằng lượt xem, độ phủ hay mức độ nhận biết thương hiệu. Khả đọc được đo bằng số thực thể đã chuẩn hóa, độ tươi của dữ liệu, tỷ lệ giao dịch thành công và tần suất được đưa vào tập lựa chọn của máy.*
- ***Tài sản tạo ra cũng khác biệt:*** *Chi cho truyền thông chủ yếu tạo hiệu ứng ngắn hạn và giá trị giảm nhanh sau khi chiến dịch kết thúc. Chi cho khả đọc tạo ra hạ tầng dữ liệu có thể tích lũy, cập nhật và tái sử dụng trên mọi nền tảng.*

Rủi ro lớn nhất của truyền thông là chi nhiều nhưng không chuyển thành lựa chọn hay giao dịch. Rủi ro của khả đọc lại nằm ở chỗ xây dựng được hạ tầng ban đầu nhưng không có nguồn lực để duy trì, khiến dữ liệu nhanh chóng lỗi thời và mất hiệu quả.

- ***Chi phí tham gia nền kinh tế tác nhân:***
- *Không phê duyệt một dự án dữ liệu địa phương nếu hồ sơ không kèm dự toán vận hành cho ít nhất ba năm sau nghiệm thu.*
- *Không đo hiệu quả của lớp dữ liệu bằng chỉ số truyền thông.*
- *Không để mỗi cơ quan tự xây một hệ định danh riêng cho cùng một tập thực thể của địa phương.*
- *Không giao toàn bộ việc mô tả địa phương cho một nhà thầu duy nhất mà không giữ lại quyền xuất dữ liệu ở định dạng mở.*

### 6.3.5. Ngưỡng chi phí cố định mới

Chi phí tham gia nền kinh tế tác nhân gồm cả phần cố định và phần biến đổi. Phí thanh toán tăng theo số giao dịch, dữ liệu vận hành tăng theo số cơ sở, công giám sát tăng theo số nền tảng. Tuy nhiên, trước khi giao dịch đầu tiên diễn ra, địa phương đã phải vượt qua một ngưỡng đầu tư tối thiểu gồm định danh thực thể, chuẩn hóa dữ liệu, quy trình cập nhật và thiết lập kết nối. Chính ngưỡng đầu vào ấy là chi phí cố định mới. Nhà máy phải được xây trước khi sản phẩm đầu tiên được bán. Cửa hàng phải trả tiền thuê mặt bằng trước khi khách hàng bước vào. Trong nền kinh tế tác nhân, khả đọc đang trở thành một loại chi phí cố định mới: khoản đầu tư phải có trước khi địa phương có cơ hội được máy lựa chọn.

Địa phương phải xây dựng và duy trì một mức hạ tầng tối thiểu, bất kể hôm nay có bao nhiêu khách được AI giới thiệu. Dữ liệu phải được chuẩn hóa. Thực thể phải được định danh. Thông tin phải được cập nhật. Hệ thống phải kết nối được với các nền tảng. Dịch vụ phải có khả năng truy vấn, đặt trước, xác nhận và thanh toán. Những khoản chi này tồn tại ngay cả khi số giao dịch còn rất thấp. Khác với quảng cáo, chi phí tham gia không thể bật lên khi cần và tắt đi khi hết chiến dịch. Một dữ liệu không được duy trì sẽ nhanh chóng lỗi thời. Một API không ổn định làm tăng tỷ lệ lỗi, khiến nền tảng và tác nhân có xu hướng ưu tiên những nguồn có khả năng hoàn tất giao dịch đáng tin cậy hơn. Một kho tri thức không cập nhật sẽ tạo ra mâu thuẫn với các nguồn mới. Vì thế, đây không phải một dự án hoàn thành một lần, mà là một năng lực phải được vận hành liên tục.

Có thể gọi đây là chi phí để được lựa chọn. Tuy nhiên, tên gọi này dễ gây hiểu nhầm, bởi việc trả chi phí không bảo đảm địa

phương sẽ được chọn. Nó chỉ mua quyền tham gia vào quá trình xét chọn. Cũng giống như trả tiền thuê mặt bằng không bảo đảm bán được hàng, xây dựng hạ tầng khả đọc không bảo đảm AI sẽ đề xuất địa phương. Nhưng trong những quyết định mà địa phương chưa được người dùng chỉ định trước và tác nhân phải tự dựng tập ứng viên, việc không trả khoản chi phí ấy làm giảm mạnh xác suất được đưa vào so sánh.

Điều này tạo ra một tầng cạnh tranh mới. Trước khi cạnh tranh về chất lượng trải nghiệm, giá cả hay bản sắc, các địa phương phải vượt qua ngưỡng khả đọc tối thiểu. Những nơi không đạt ngưỡng bị loại từ trước khi các lợi thế thật sự của họ được so sánh. Do đó, thị trường không chỉ chọn trải nghiệm tốt nhất, mà trước hết chọn trong số những trải nghiệm đã được máy đọc được. Vì vậy, quản trị địa phương phải nhìn khả đọc như một loại hạ tầng thiết yếu, không phải một khoản chi truyền thông tùy chọn. Nó cần ngân sách ổn định, nhân lực chuyên trách, tiêu chuẩn vận hành và cơ chế bảo trì dài hạn. Trong nền kinh tế tác nhân, chi phí đầu tiên của cạnh tranh không còn là làm sao để thắng, mà là trả đủ để còn được đứng trong cuộc chơi.

## 6.4. KHI GIAO DIỆN NÉN NƠI CHỐN THÀNH THUỘC TÍNH

Trong nền kinh tế trải nghiệm, mỗi địa phương vốn là một tổ hợp độc nhất của lịch sử, văn hóa, con người và ký ức. Không có hai nơi chốn nào hoàn toàn giống nhau. Tuy nhiên, khi quá trình lựa chọn ngày càng được thực hiện bởi các hệ thống số, sự phong phú ấy phải đi qua một lớp giao diện chuẩn hóa trước khi đến được với người dùng. Chính lớp giao diện này bắt đầu định nghĩa cách địa phương được nhìn nhận.

Máy có thể tiếp nhận một câu chuyện dài, nhưng để so sánh và hành động, hệ thống phải chuyển câu chuyện ấy thành một tập tín hiệu, thuộc tính hoặc biểu diễn có thể xếp hạng. Tên địa điểm, số sao, mức giá, khoảng cách, thời gian di chuyển, đánh giá của người dùng, mức độ phổ biến hay khả năng đặt dịch vụ ngay lập tức trở thành những tín hiệu nổi bật nhất. Những yếu tố khó lượng hóa như bầu không khí, ký ức cộng đồng, nhịp sống địa phương hay chiều sâu văn hóa bị đẩy xuống phía sau vì chúng không dễ biểu diễn trong cùng một giao diện.

Cần nói ngay một điều để tránh hiểu nhầm về sau. Không phải bản thân giao diện làm mất bản sắc. Một giao diện được thiết kế tốt hoàn toàn có thể làm lộ ra những khác biệt mà trước đây du khách không có cách nào tìm thấy, chẳng hạn một xưởng gốm chỉ mở cửa ba ngày mỗi tuần hay một quán ăn gia đình không có biển hiệu. Thứ làm phẳng khác biệt là lược đồ so sánh nghèo nàn cộng với mục tiêu tối ưu hẹp. Giao diện chỉ làm phẳng những khác biệt mà kiến trúc dữ liệu và hàm mục tiêu không cho phép biểu diễn. Phân biệt này quan trọng, vì nó chuyển vấn đề từ chỗ không thể làm gì được sang chỗ có thể thiết kế lại.

### 6.4.1. Giao diện làm phẳng khác biệt

Các nền tảng số đều phải giải quyết cùng một bài toán: làm thế nào để người dùng hoặc tác nhân AI có thể so sánh hàng nghìn lựa chọn chỉ trong vài giây. Vì vậy, chúng gần như đều sử dụng những cấu trúc giao diện giống nhau: thẻ thông tin, danh sách và tập kết quả được rút gọn còn vài phương án đứng đầu. Mỗi điểm đến được trình bày bằng một bức ảnh đại diện, một tên gọi, vài dòng mô tả ngắn, số sao đánh giá, mức giá, khoảng cách, thời gian di chuyển và một vài nhãn nổi bật như "Phổ biến", "Được đề xuất" hay "Đặt nhiều".

Những giao diện này giúp việc lựa chọn trở nên nhanh chóng nhưng đồng thời cũng làm giảm chiều sâu của thông tin. Một làng nghề có lịch sử hàng trăm năm và một khu mua sắm mới xây có thể xuất hiện dưới cùng một định dạng hiển thị. Một lễ hội truyền thống chỉ còn là một mục "Sự kiện", trong khi một ngọn núi linh thiêng chỉ còn là một "Điểm tham quan". Tất cả đều được biểu diễn bằng những trường dữ liệu giống nhau để hệ thống có thể so sánh.

Với tác nhân AI, bước nén này còn quyết liệt hơn. Máy đọc được cả văn bản dài, nhưng để xếp hạng thì nó phải quy mọi thứ về các tín hiệu đã chuẩn hóa, rồi chỉ xem xét sâu một số rất ít ứng viên đứng đầu. Những khác biệt không được biểu diễn trong lớp giao diện gần như không tham gia vòng so sánh ấy. Hiện tượng này có thể gọi là làm phẳng giao diện (interface flattening). Nó không xóa bỏ bản sắc của địa phương trong thực tế, nhưng làm giảm khả năng bản sắc ấy được nhìn thấy trong quá trình lựa chọn. Càng nhiều quyết định được thực hiện qua giao diện số, giá trị cạnh tranh của địa phương càng phụ thuộc vào những gì có thể hiển thị và so sánh, thay vì toàn bộ chiều sâu của trải nghiệm thực tế.

Thuộc tính nào không được biểu diễn trong lược đồ mà hệ thống dùng để so sánh thì không tham gia cuộc cạnh tranh. Nó không thua. Nó vắng mặt. Hệ quả là cạnh tranh dịch chuyển dần về những chiều đo được, ngay cả khi các chiều đo được không phải là chiều tạo ra giá trị lớn nhất.

Trong nền kinh tế tác nhân, những yếu tố máy đọc tốt nhất chưa chắc là những yếu tố quan trọng nhất đối với trải nghiệm. Máy có thể dễ dàng xử lý các thông tin như giá và tổng chi phí chuyến đi, khoảng cách và thời gian di chuyển, giờ mở cửa, tình trạng còn chỗ, hay khả năng đặt và thanh toán ngay. Đây đều là các dữ liệu có cấu trúc, dễ chuẩn hóa và dễ đưa vào quá trình xếp hạng. Ở lớp tiếp theo là những yếu tố có thể đo được nhưng dữ liệu còn yếu hoặc chưa được sử dụng rộng rãi, như tính xác thực của sản phẩm thủ công, mức độ đông đúc theo từng khung giờ, chất lượng người kể chuyện, tác động của khách đến cư dân hay tính bền vững của chuỗi cung ứng địa phương. Những dữ liệu này đã có thể thu thập nhưng chưa đủ phổ biến hoặc chưa đủ chuẩn để ảnh hưởng mạnh đến quyết định của máy. Khó nhất là các giá trị không thể thu gọn thành một chỉ số duy nhất, như bầu không khí và nhịp sống của nơi chốn, ký ức cộng đồng và ý nghĩa nghi lễ, quan hệ giữa chủ và khách, sức chịu tải của cộng đồng hay tri thức bản địa chưa được ghi chép. Đây thường là những yếu tố tạo nên bản sắc sâu nhất của một địa phương, nhưng cũng là phần khó biểu diễn nhất trong ngôn ngữ của máy.

### 6.4.2. Từ nơi chốn thành một sản phẩm

Trong thương mại điện tử, mỗi sản phẩm đều có một đơn vị quản lý riêng (SKU, tức stock keeping unit), bao gồm các thuộc tính chuẩn hóa như tên, danh mục, giá, tình trạng còn hàng và các

thông số kỹ thuật. Mục tiêu của SKU là giúp hệ thống dễ dàng lưu trữ, tìm kiếm, so sánh và giao dịch.

Các nền tảng du lịch và tác nhân AI đang áp dụng một logic rất gần với đó cho các điểm đến. Một địa phương được gắn định danh, phân loại theo danh mục, gắn tọa độ, mức giá, thời gian di chuyển, điểm đánh giá, khả năng đặt dịch vụ, lịch mở cửa và hàng loạt thuộc tính chuẩn hóa khác. Đối với hệ thống, điểm đến trở thành một thực thể có thể được truy vấn, lọc, xếp hạng và đưa vào giao dịch gần giống như bất kỳ sản phẩm số nào.

Quá trình này tạo ra hiện tượng hàng hóa hóa (commoditization). Khi nhiều địa phương cùng đáp ứng những tiêu chuẩn hiển thị giống nhau, cạnh tranh dần chuyển từ việc kể một câu chuyện độc đáo sang việc tối ưu các thuộc tính mà giao diện có thể đo lường. Chỉ cần hai điểm đến có khoảng cách tương đương, mức giá tương đương và điểm đánh giá tương đương, chúng đã có thể trở thành những lựa chọn thay thế cho nhau trong mắt thuật toán, mặc dù giá trị văn hóa và trải nghiệm thực tế có thể rất khác biệt.

Hàng hóa hóa ở đây diễn ra chủ yếu tại lớp giao diện, chứ không phải ở bản chất của địa phương. Bên dưới giao diện, mỗi nơi vẫn sở hữu bản sắc riêng, lịch sử riêng và cộng đồng riêng. Nhưng trên màn hình của máy và của người dùng, tất cả đều xuất hiện như những mục dữ liệu có thể so sánh trực tiếp. Quyền lực vì thế dần chuyển từ người kể câu chuyện sang người thiết kế giao diện và các tiêu chí lựa chọn.

### 6.4.3. Thuộc tính nào không được biểu diễn sẽ biến mất

Để máy xử lý được, bản sắc của một nơi chốn phải nén thành một tập thuộc tính hữu hạn. Những thuộc tính sống sót qua bước nén thường là thứ dễ đo và dễ so sánh: giá, điểm đánh giá, khoảng cách, thời gian di chuyển, số lượng đánh giá, tỷ lệ còn chỗ, thời gian mở cửa, khả năng đặt trực tuyến hay mức độ phổ biến. Đây đều là các biến số mà thuật toán có thể đưa vào mô hình xếp hạng và tối ưu hóa. Ngược lại, những giá trị như cảm giác thuộc về nơi chốn, chiều sâu văn hóa, sự chân thực của cộng đồng hay ký ức tập thể rất khó được biểu diễn bằng các trường dữ liệu chuẩn hóa nên thường có trọng số thấp hoặc hoàn toàn biến mất trong quá trình lựa chọn.

Khi một thuộc tính không được biểu diễn, nó không biến mất khỏi thực tại mà biến mất khỏi quá trình so sánh và lựa chọn của máy. Sự nén ấy còn tạo ra một hiệu ứng chuẩn hóa hành vi. Khi các địa phương nhận thấy máy ưu tiên giá, điểm đánh giá hay khả năng đặt dịch vụ tức thời, họ sẽ có xu hướng đầu tư nhiều hơn vào việc cải thiện các chỉ số đó. Dần dần, nguồn lực quản trị chuyển từ việc làm giàu bản sắc sang việc tối ưu các thuộc tính hiển thị. Những gì có thể đo được sẽ được đầu tư nhiều hơn những gì khó đo, mặc dù chính các giá trị khó đo ấy mới tạo nên sự khác biệt lâu dài của một nơi chốn.

Nghịch lý nằm ở đây. Muốn được máy lựa chọn, địa phương buộc phải chuyển bản sắc thành dữ liệu. Nhưng càng chuyển thành dữ liệu, bản sắc càng có nguy cơ bị rút gọn thành một vài thuộc tính có thể so sánh. Vì vậy, bài toán của quản trị địa phương không chỉ là tăng điểm đánh giá, giảm khoảng cách cảm nhận hay tối ưu giá, mà còn là tìm cách biểu diễn những giá trị vô hình dưới các dạng dữ liệu mà máy có thể đọc được, thay vì để

chúng hoàn toàn biến mất khỏi quá trình lựa chọn. Khi đó, mục tiêu không phải là biến bản sắc thành một con số, mà là biến bản sắc thành bằng chứng số đủ mạnh để vẫn được nhìn thấy trong một thế giới mà mọi quyết định đều bắt đầu từ dữ liệu.

### 6.4.4. Khi AI tối ưu theo tiêu chí giống nhau

Một người có thể lựa chọn điểm đến vì ký ức tuổi thơ, vì một cuốn tiểu thuyết yêu thích hay đơn giản vì muốn khám phá điều chưa ai biết. Con người sử dụng nhiều tiêu chí khác nhau, trong đó có cả những yếu tố cảm xúc không thể đo lường. Chính sự đa dạng trong cách lựa chọn giúp nhiều loại điểm đến khác nhau cùng có cơ hội được phát hiện.

Ngược lại, các tác nhân AI có xu hướng tối ưu theo một tập tiêu chí tương đối giống nhau. Dù được phát triển bởi những công ty khác nhau, phần lớn hệ thống đều phải trả lời một câu hỏi gần giống nhau: lựa chọn nào mang lại kết quả tốt nhất cho người dùng với mức rủi ro thấp nhất? Vì vậy, chúng thường ưu tiên những tín hiệu tương tự như độ tin cậy của nguồn dữ liệu, mức đánh giá, giá cả, khoảng cách, khả năng đặt dịch vụ ngay lập tức, tỷ lệ thành công của giao dịch và mức độ hài lòng của người dùng trước đó.

Sự hội tụ này làm cho các hệ thống AI ngày càng đưa ra những đề xuất giống nhau. Một điểm đến có điểm đánh giá cao, dữ liệu đầy đủ và khả năng giao dịch thuận tiện sẽ đồng thời được nhiều hệ thống lựa chọn. Ngược lại, một địa phương có bản sắc đặc sắc nhưng dữ liệu nghèo nàn hoặc thiếu khả năng giao dịch sẽ cùng lúc bị nhiều hệ thống loại bỏ. Không cần có sự thông đồng giữa các nền tảng, kết quả vẫn dần hội tụ vì tất cả đều đang tối ưu

trên cùng một không gian tín hiệu. Khi con người lựa chọn, sự đa dạng trong sở thích tạo ra nhiều thị trường ngách. Khi hàng triệu tác nhân AI cùng lựa chọn theo những tiêu chí tương tự, không gian cạnh tranh có xu hướng đồng nhất hơn và quá trình tập trung hóa diễn ra mạnh hơn.

Nhà phản biện cho rằng hàng hóa hóa là chuyện của các nền tảng đặt phòng có từ hai mươi năm trước, không phải của AI. Gán nó cho tác nhân AI là quy kết sai nguyên nhân: Chính xác. Điều này cần nói rõ để tránh một sai lầm phổ biến trong các phân tích về AI. Việc trình bày điểm đến bằng thẻ thông tin chuẩn hóa đã có từ lâu trước khi có tác nhân. Tác nhân AI không tạo ra hiện tượng này. Vai trò của nó là khuếch đại và định tuyến. Điểm mới nằm ở hai chỗ. Thứ nhất, con người nhìn thấy hai mươi kết quả rồi tự lọc, còn tác nhân trả về ba phương án đã lọc sẵn, nên hiệu ứng nén mạnh hơn nhiều lần. Thứ hai, nhiều tác nhân cùng tối ưu trên một không gian tín hiệu gần giống nhau, nên sự đa dạng giữa các nền tảng, vốn từng là chỗ trú cho những điểm đến khác lạ, thu hẹp lại. AI là bộ khuếch đại và bộ định tuyến, không phải nguyên nhân gốc.

Cầu được trung gian bởi tác nhân có độ co giãn theo giá cao hơn cầu do con người trực tiếp quyết định. Cơ chế dự kiến gồm ba phần: tác nhân so sánh giá của nhiều phương án cùng lúc thay vì tuần tự, ma sát chuyển đổi giữa các phương án gần bằng không, cộng thêm việc phần lớn tác nhân được giao một ngân sách để tối ưu.

Nếu giả thuyết đúng, thị phần do tác nhân phân phối càng lớn thì áp lực giảm giá càng mạnh, còn phần giá trị cao cấp sẽ co lại trước ở phân khúc giữa. Giả thuyết sẽ sai nếu quan sát thấy một trong các điều sau: lòng trung thành thương hiệu vẫn giữ được

khách qua kênh tác nhân, trải nghiệm đủ độc đáo để không có phương án thay thế trong tập so sánh, tác nhân tối ưu đa mục tiêu chứ không chủ yếu theo giá, hoặc sản phẩm được đóng gói thành combo khó so sánh trực tiếp. Giả thuyết này chưa có dữ liệu đủ dài để kiểm định và được nêu ở đây như một mệnh đề có thể bác bỏ. Thứ máy không đo được thì không thua. Nó chỉ đơn giản là không dự thi.

### 6.4.5. Hàng hóa hóa dẫn tới cạnh tranh giá

Khách sạn M (cái gì hiện lên và cái gì không): Trên giao diện, khách sạn M hiện ra bằng bảy trường dữ liệu: giá, điểm đánh giá 8,4, khoảng cách tới bãi biển, giờ nhận phòng, còn ba phòng trống, có bãi đỗ xe, hủy miễn phí. Thứ thực sự làm khách quay lại thì không có trường nào: bữa sáng do một người trong làng nấu theo mùa, cùng với việc chủ nhà biết chỉ đường tới một bãi đá mà bản đồ số chưa ghi tên. Hai điều đó không thua trong cuộc so sánh. Chúng không tham dự cuộc so sánh. Đây chính là chỗ mà một lược đồ giàu hơn, chẳng hạn có trường cho nguồn gốc nguyên liệu hoặc cho hoạt động do chủ nhà dẫn, sẽ thay đổi kết quả mà không cần khách sạn M hạ giá thêm đồng nào.

Trong kinh tế học, khi các sản phẩm ngày càng khó phân biệt, cạnh tranh thường chuyển sang yếu tố dễ so sánh nhất, đó là giá. Đây chính là hệ quả điển hình của quá trình hàng hóa hóa. Khi người mua không còn nhận thấy sự khác biệt đủ lớn về chất lượng hay giá trị, quyết định cuối cùng thường được đưa ra dựa trên mức giá thấp hơn.

Hệ quả là các địa phương có thể bị cuốn vào cuộc đua giảm giá thay vì cuộc đua gia tăng giá trị. Khuyến mại nhiều hơn, giảm

giá phòng nhiều hơn, trợ giá dịch vụ nhiều hơn trở thành cách nhanh nhất để cải thiện khả năng được lựa chọn. Tuy nhiên, đây thường là lợi thế ngắn hạn. Khi một địa phương giảm giá, những địa phương khác cũng có thể phản ứng tương tự, khiến toàn bộ thị trường cùng giảm biên lợi nhuận mà không tạo ra khác biệt bền vững.

Nguy hiểm hơn, giá trị cao cấp dần biến mất. Trong nền kinh tế trải nghiệm, khả năng định giá cao không đến từ chi phí sản xuất mà đến từ tính độc đáo của trải nghiệm, từ bản sắc văn hóa và từ giá trị biểu tượng mà nơi chốn mang lại. Nếu các yếu tố đó không được biểu diễn đầy đủ trong lớp giao diện, thuật toán sẽ khó nhận diện lý do vì sao một địa phương xứng đáng với mức giá cao hơn. Phần giá trị cao cấp không biến mất vì trải nghiệm kém đi, mà vì hệ thống lựa chọn không còn nhìn thấy đầy đủ những yếu tố tạo nên giá trị cao cấp.

Đó là lý do chiến lược của địa phương trong nền kinh tế tác nhân không thể chỉ là tối ưu các chỉ số cạnh tranh hay liên tục hạ giá. Mục tiêu quan trọng hơn là chuyển hóa bản sắc thành các tín hiệu số có thể được máy nhận diện, để sự khác biệt vẫn hiện diện ngay trong lớp giao diện. Chỉ khi thuật toán hiểu được điều gì làm cho một nơi chốn thực sự độc đáo, địa phương mới có thể duy trì năng lực định giá cao và tránh bị cuốn vào vòng xoáy cạnh tranh bằng giá của một thị trường đã bị hàng hóa hóa.

Nếu hình dung toàn ngành như một đường cong giá trị, đỉnh lợi nhuận không còn nằm hoàn toàn ở hai đầu truyền thống là sáng tạo trải nghiệm và tiêu dùng trải nghiệm. Một đỉnh mới xuất hiện ở lớp hạ tầng số, nơi các hệ thống AI, nền tảng dữ liệu, giao diện tác nhân và mạng lưới thanh toán quyết định ai được nhìn thấy, ai được lựa chọn và ai được giao dịch. Đây là sự dịch

chuyển tương tự quá trình internet từng chuyển giá trị từ nhà sản xuất sang các nền tảng tìm kiếm và thương mại điện tử, nhưng diễn ra sâu hơn vì AI không chỉ kết nối mà còn thay con người đưa ra quyết định.

Do đó, câu hỏi trung tâm của nền kinh tế tác nhân không còn là địa phương tạo ra bao nhiêu giá trị, mà là địa phương giữ lại được bao nhiêu giá trị sau khi toàn bộ hệ thống trung gian mới đã hoàn thành quá trình phân phối và điều phối. Hiểu được đường cong giá trị mới này là điều kiện để các địa phương xác định đúng vị trí của mình trong chuỗi giá trị đang tái cấu trúc và xây dựng chiến lược nhằm không chỉ tạo ra trải nghiệm hấp dẫn mà còn bảo vệ năng lực nắm giữ giá trị trong kỷ nguyên AI.

## 6.5. ĐƯỜNG CONG LỢI NHUẬN MỚI

Xem Hình 6.2. Sự dịch chuyển của biên lợi nhuận

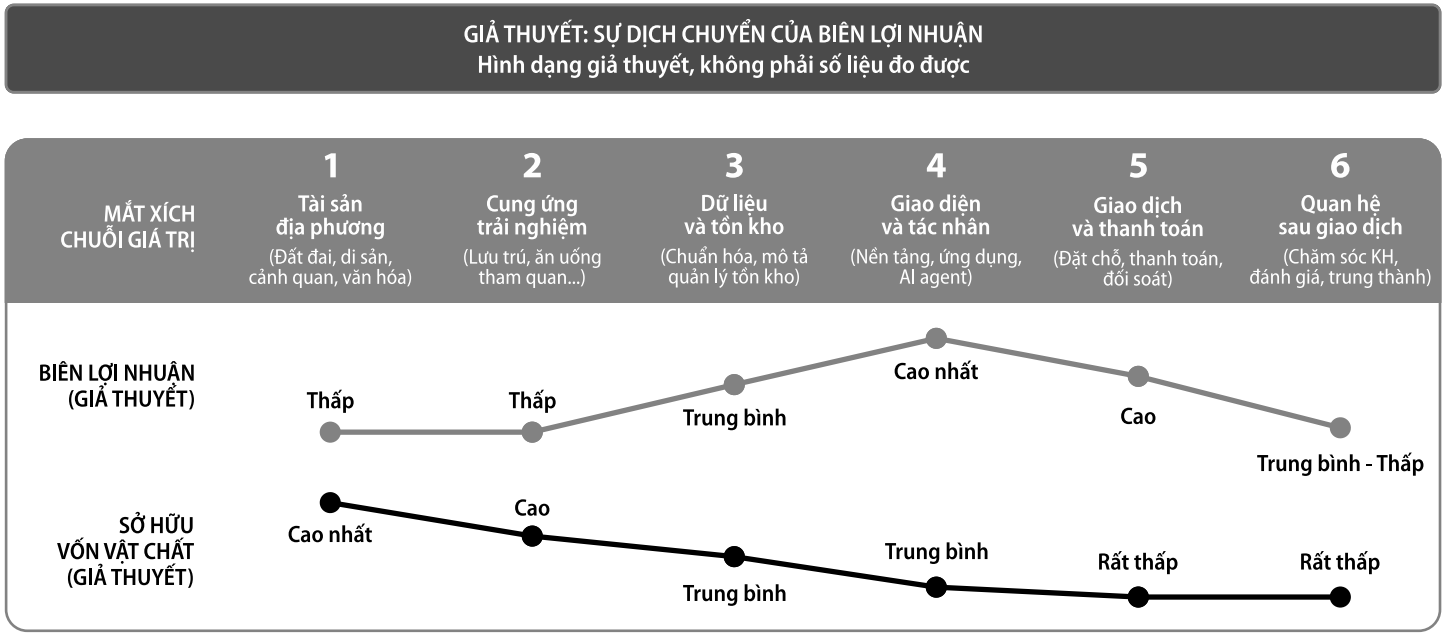
### 6.5.1. Giá trị dịch chuyển khỏi nơi tạo trải nghiệm

Trải nghiệm vẫn là điều kiện cần, nhưng giao diện mới là nơi quyết định trải nghiệm nào được nhìn thấy. Một khách sạn xuất sắc không xuất hiện trong câu trả lời của tác nhân thì mất cơ hội tiếp cận khách, trong khi một chủ thể không tạo ra trải nghiệm nào vẫn có thể giữ phần lớn giá trị nhờ kiểm soát điểm chạm đầu tiên giữa nhu cầu và nguồn cung.

Đây là bước tiếp theo của một quy luật đã từng xuất hiện trong nền kinh tế internet, khi giá trị chuyển từ cửa hàng vật lý sang công cụ tìm kiếm rồi sang sàn thương mại điện tử. Điểm khác là lần này quyền lực không dừng ở nơi trưng bày lựa chọn mà đi thêm một chặng, tới nơi lựa chọn được thực hiện.

Một đặc điểm nổi bật của nền kinh tế tác nhân là giá trị ngày càng tích tụ ở những lớp hạ tầng mà người dùng hầu như không nhìn thấy. Nếu trước đây doanh thu tập trung ở các doanh nghiệp trực tiếp phục vụ khách, thì nay một phần ngày càng lớn được chuyển sang các hệ thống cung cấp năng lực cho AI hoạt động.

***Lớp đầu tiên là giao diện lập trình ứng dụng.*** Mỗi lần tác nhân AI truy vấn tình trạng phòng, giá vé, lịch mở cửa hay khả năng đặt chỗ đều thông qua các giao diện lập trình. Chủ thể kiểm soát API hoặc tiêu chuẩn kết nối có thể thu phí trực tiếp, hoặc thu giá trị gián tiếp bằng quyền truy cập dữ liệu và quyền điều phối giao dịch. Khi hàng triệu tác nhân hoạt động đồng thời, lớp kết nối này trở thành một loại đường cao tốc của nền kinh tế số.




Ý NGHĨA:
Vốn vật chất và lao động tập trung ở hai mắt xích đầu (1-2)
Biên lợi nhuận và quyền sở hữu dữ liệu dịch chuyển lên các mắt xích giữa và cuối (3-6)
Đạt đỉnh ở mắt xích 4 (giao diện và tác nhân)

Biên lợi nhuận (giả thuyết)
Sở hữu vốn vật chất (giả thuyết)

**Hình 6.2.** Sự dịch chuyển của biên lợi nhuận

*Nguồn: Tác giả xây dựng*

***Lớp thứ hai là hạ tầng điện toán.*** Toàn bộ quá trình truy hồi dữ liệu, suy luận, lập kế hoạch và thực thi của AI đều tiêu thụ tài nguyên tính toán. Mỗi cuộc hội thoại, mỗi lần đề xuất hay mỗi giao dịch đều cần năng lực lưu trữ, xử lý và truyền dữ liệu. Vì vậy, các nhà cung cấp hạ tầng điện toán có thể thu lợi từ sự tăng trưởng của toàn hệ sinh thái mà không cần sở hữu bất kỳ điểm đến hay dịch vụ du lịch nào.

***Lớp thứ ba là hạ tầng thanh toán.*** Khi AI không chỉ đề xuất mà còn tự động đặt phòng, mua vé và thanh toán, gần như mọi giao dịch đều đi qua hệ thống thanh toán số. Khoản phí trên từng giao dịch tuy nhỏ nhưng được lặp lại với quy mô rất lớn, tạo nên nguồn doanh thu ổn định và có khả năng mở rộng rất lớn theo quy mô giao dịch. Trong nhiều trường hợp, doanh nghiệp thanh toán có thể đạt tỷ suất lợi nhuận cao hơn chính đơn vị trực tiếp tạo ra trải nghiệm.

Đối với địa phương, điều này mang một hàm ý chiến lược quan trọng. Nếu chỉ đầu tư vào tài sản và trải nghiệm mà không xây dựng năng lực kết nối số, địa phương sẽ tạo ra giá trị nhưng ngày càng giữ lại được ít giá trị hơn. Trong nền kinh tế tác nhân, cạnh tranh không còn diễn ra riêng ở lớp dịch vụ, mà còn ở lớp hạ tầng dữ liệu, giao thức kết nối và hệ thống giao dịch, nơi phần lớn giá trị mới đang được tích lũy.

Trong nền kinh tế tác nhân, dữ liệu trở thành một loại tài sản có khả năng tự sinh giá trị. Khác với tài nguyên vật chất, dữ liệu không bị hao mòn khi sử dụng. Mỗi lần được khai thác, dữ liệu lại có thể tạo ra thêm dữ liệu mới. Một lượt tìm kiếm tạo ra lịch sử truy vấn. Một giao dịch tạo ra dữ liệu hành vi. Một đánh giá tạo ra tín hiệu chất lượng. Một lần AI đề xuất thành công lại tạo thêm bằng chứng rằng đề xuất đó là đúng. Chính cơ chế tự tích

lũy này tạo nên bánh đà dữ liệu (data flywheel). Mỗi vòng quay đều làm tăng giá trị của toàn bộ hệ thống. Khi dữ liệu đủ lớn, AI không chỉ hiểu khách hàng tốt hơn mà còn dự báo nhu cầu, tối ưu giá, cá nhân hóa hành trình và cải thiện hiệu quả vận hành. Giá trị kinh tế vì thế không chỉ đến từ giao dịch hiện tại mà còn đến từ khả năng tạo ra những giao dịch trong tương lai.

Đối với các địa phương, đây là một thay đổi mang tính chiến lược. Nếu trải nghiệm chỉ tạo ra doanh thu mà không tạo ra dữ liệu có cấu trúc, mỗi lượt khách sẽ kết thúc khi chuyến đi kết thúc. Ngược lại, nếu mỗi tương tác đều được chuyển hóa thành tri thức có thể đọc, có thể học và có thể tái sử dụng, thì mỗi du khách sẽ tiếp tục tạo giá trị cho địa phương ngay cả sau khi đã rời đi.

### 6.5.2. Hình dạng mới của đường cong lợi nhuận

Một hệ quả nữa của bản đồ này cần được nêu ở đây rồi để mục sau khai thác. Khi khách hàng giao quyền lựa chọn cho một tác nhân, quan hệ với người dùng cuối cũng dịch lên lớp giao diện. Bên giữ quan hệ ấy là bên quyết định giao dịch được chuyển tới ai. Hệ quả của việc này đối với địa phương sâu đến mức cần một mục riêng.

Đường cong ở Hình 6.2 gói lại toàn bộ lập luận của mục này. Trước kia phần lớn giá trị tập trung ở nơi trực tiếp tạo ra trải nghiệm, còn nền tảng chỉ giữ một phần dưới dạng phí môi giới. Nay lợi nhuận dịch lên hai lớp cao hơn là hạ tầng và điều phối, trong khi chi phí vật chất cùng lao động vẫn nằm nguyên ở lớp cung ứng. Ở cuối chuỗi là điểm đến, nơi vẫn phải đầu tư vào cảnh quan, văn hóa, hạ tầng vật chất, con người và chất lượng dịch vụ. Đây là lớp chịu phần lớn chi phí tạo ra trải nghiệm nhưng không còn giữ toàn bộ giá trị do trải nghiệm đó sinh ra. Một phần lợi nhuận đã được phân phối cho các lớp hạ tầng và điều phối trước khi doanh thu đến được với địa phương.

Đường cong lợi nhuận mới vì thế có thể được tóm tắt bằng một mệnh đề. Trong nền kinh tế trải nghiệm, giá trị tập trung ở nơi tạo ra trải nghiệm. Trong nền kinh tế tác nhân, giá trị tập trung ở nơi điều phối trải nghiệm. Chiến lược của địa phương không còn chỉ là tạo ra nhiều giá trị hơn, mà còn là giành lại quyền nắm giữ nhiều giá trị hơn trong chuỗi phân phối mới.

Mỗi lần một công nghệ hứa hẹn loại bỏ trung gian, nó thường tạo ra một trung gian mới ở tầng cao hơn, gần với điểm ra quyết định hơn và khó thay thế hơn trung gian cũ. Phi trung gian hóa ở lớp giao dịch đi kèm tái trung gian hóa ở lớp quyết định. Ngoại lệ tồn tại khi tiêu chuẩn mở và khả năng chuyển đổi được thiết lập trước khi thị trường kịp tập trung, điều rất hiếm xảy ra một cách tự nhiên.

Trung Quốc cho thấy mệnh đề này ở dạng rõ nhất. Trip.com Group, tập đoàn sở hữu Ctrip và Qunar, không chờ một tác nhân bên ngoài chen vào giữa mình với khách. Nó tự dựng tác nhân. Trợ lý TripGenie ra mắt năm 2023, tới đầu năm 2026 ghi nhận số đơn có AI hỗ trợ tăng bốn lần so với cùng kỳ, mức dùng các công cụ lõi tăng khoảng ba lần và gần sáu phần mười lượt tương tác gắn trực tiếp với một quyết định đặt chỗ.[1] Nói cách khác, lớp điều phối mà mục này vừa mô tả không dịch ra khỏi nền tảng, nó dịch lên trên ngay bên trong nền tảng. Với địa phương, hệ quả không nhẹ đi mà nặng thêm. Trước đây một điểm đến còn có thể hy vọng rằng lớp tác nhân mới sẽ phá thế độc quyền phân phối cũ. Trường hợp Trip.com cho thấy khả năng ngược lại, bên đang giữ tồn kho, dữ liệu giao dịch và quan hệ khách hàng cũng là bên có điều kiện tốt nhất để dựng tác nhân và do đó có điều

1. Trip.com Group, số liệu công bố tại hội chợ ITB Berlin tháng 3 năm 2026, nhân ba năm vận hành trợ lý TripGenie. Đây là dữ liệu doanh nghiệp tự công bố, chưa có kiểm toán độc lập, tính trên nền so sánh năm trước và không kèm số tuyệt đối.

kiện tốt nhất để giữ luôn lớp quyết định. Ở Việt Nam đây không phải chuyện xa. Trong năm 2025, lượt tải mới ứng dụng Trip.com đạt 1,8 triệu, tăng 63%, còn lượt đặt phòng nội địa qua nền tảng này tăng 75%.[1]

Nhà phản biện sẽ lập luận rằng, nếu nền tảng thu quá nhiều, khách sạn sẽ rời đi và bán trực tiếp. Thị trường có lối thoát. Thực tế, lối thoát ấy tồn tại nhưng hẹp hơn vẻ ngoài. Điều thú vị là bằng chứng đến từ cả hai phía. Về phía cho thấy lối thoát hẹp: kênh trực tiếp rẻ hơn nhiều, khoảng 4 đến 5% so với 15 đến 25%, vậy mà phần lớn cơ sở lưu trú độc lập vẫn ở lại nền tảng, vì rời đi đồng nghĩa với biến mất khỏi nơi khách hàng đang tìm kiếm.[2] Tỷ lệ khách tìm thấy khách sạn trên nền tảng rồi quay sang đặt trực tiếp được ước tính khoảng 9 đến 15% năm 2025, giảm so với mức 20 đến 25% của năm năm trước.

Về phía cho thấy lối thoát vẫn mở, sau khi Liên minh châu Âu buộc bỏ điều khoản ngang giá từ tháng 7 năm 2024, các cơ sở trong khu vực kinh tế châu Âu được quyền chào giá thấp hơn trên trang của chính mình, trong khi tỷ trọng kênh trực tiếp của Booking.com vẫn ở mức giữa năm mươi phần trăm số đêm phòng.[3] Kết luận trung thực là lối thoát tồn tại nhưng phải được luật mở ra, chứ hiếm khi tự mở. Người tạo ra trải nghiệm và người giữ lại lợi nhuận từng là một người. Đó là điều đang chấm dứt.

1. Trip.com Group, số liệu thị trường Việt Nam công bố tháng 7 năm 2026 nhân dịp mở rộng văn phòng tại Thành phố Hồ Chí Minh. Cũng là số liệu doanh nghiệp tự công bố. Cùng nguồn cho biết lượt đặt vé các chặng nối Việt Nam với Đông Bắc Á tăng 103%.

2. Ước lượng ngành, tổng hợp từ Sojern và Hospitality Net năm 2026. Đây là chi phí biến đổi trên mỗi giao dịch của kênh trực tiếp, mức tin cậy trung bình.

3. Booking Holdings, báo cáo tuân thủ Đạo luật Thị trường Số tháng 11 năm 2024: từ ngày 1 tháng 7 năm 2024, Booking.com bỏ điều khoản ngang giá cho toàn khu vực kinh tế châu Âu, hoàn tất từ ngày 2 tháng 12 năm 2024. Tỷ trọng kênh trực tiếp theo thông cáo kết quả cả năm 2025, ngày 18 tháng 2 năm 2026. Cả hai đều do doanh nghiệp công bố nhưng nằm trong hồ sơ có ràng buộc pháp lý.

## 6.6. NGUY CƠ OEM HÓA ĐỊA PHƯƠNG

Trong mục này, cụm nhà điều hành địa phương chỉ các tổ chức trực tiếp tổ chức, bán hoặc điều phối trải nghiệm tại nơi chốn, gồm doanh nghiệp lưu trú và lữ hành, đơn vị quản lý điểm đến và cơ quan xúc tiến du lịch. Khi cần nói về chính quyền với tư cách chủ thể ban hành chính sách, chương dùng từ chính quyền địa phương. Phân biệt này cần thiết vì bốn mất mát sắp phân tích không rơi đều lên các nhóm ấy. Cụm OEM hóa dùng trong mục này chỉ quá trình địa phương vẫn sản xuất trải nghiệm thật nhưng dần mất thương hiệu, quan hệ khách hàng và quyền định giá, giống một nhà sản xuất làm hàng cho nhãn hiệu của người khác.

Trong nền kinh tế trải nghiệm truyền thống, nhà điều hành địa phương vừa tạo ra trải nghiệm vừa trực tiếp xây quan hệ với khách. Mỗi chuyến đi vừa mang lại doanh thu vừa để lại một mối quan hệ có thể khai thác lại. Nền kinh tế tác nhân cắt đúng vế thứ hai. Hệ quả là nhiều địa phương có nguy cơ trở thành những nhà cung cấp phía sau (backend provider) của nền kinh tế tác nhân. Họ vận hành điểm đến, nhưng không kiểm soát được luồng khách. Họ tạo ra trải nghiệm, nhưng không sở hữu dữ liệu. Họ đầu tư dài hạn, nhưng chỉ nhận được phần nhỏ trong chuỗi giá trị. Đây không chỉ là sự thay đổi về công nghệ mà là sự tái phân bổ quyền lực trong toàn bộ ngành du lịch. Những mục tiếp theo sẽ phân tích quá trình địa phương dần mất quyền tiếp cận khách hàng, mất quyền sở hữu dữ liệu và những chiến lược cần thiết để giành lại vị trí trong chuỗi giá trị mới.

Trong nền kinh tế tác nhân, một địa phương có thể vẫn giữ được điểm đến, doanh nghiệp và dòng khách, nhưng dần đánh mất những quyền lực kinh tế quan trọng hơn. Quyền tiếp cận

khách hàng thường là quyền mất đi đầu tiên. Dấu hiệu nhận biết là địa phương hoặc doanh nghiệp không còn sở hữu danh sách khách quay lại của riêng mình, mà phụ thuộc vào nền tảng để kết nối với khách hàng. Hệ quả là mỗi lượt khách mới đều phải mua lại từ đầu, làm chi phí thu hút khách ngày càng tăng. Tiếp theo là quyền sở hữu dữ liệu hành vi. Nhiều địa phương chỉ biết ai đã đến, nhưng không biết bao nhiêu người đã tìm kiếm, cân nhắc rồi từ bỏ. Khi không hiểu lý do khách không lựa chọn, việc cải thiện sản phẩm hay tối ưu trải nghiệm chỉ còn dựa trên phỏng đoán. Quyền định giá cũng có thể bị suy yếu. Dấu hiệu dễ thấy là giá niêm yết của doanh nghiệp ngày càng bị chi phối bởi các chương trình khuyến mại hoặc chính sách của nền tảng. Về lâu dài, khả năng định vị ở phân khúc cao cấp bị bào mòn và cạnh tranh dần chuyển sang cuộc đua giảm giá. Cuối cùng là quyền điều phối chuỗi giá trị. Khi mỗi doanh nghiệp địa phương phải tự đàm phán với từng nền tảng, sức mạnh thương lượng của cả hệ sinh thái bị phân tán. Hệ quả không chỉ là điều kiện hợp tác kém thuận lợi hơn, mà còn là việc địa phương mất khả năng phối hợp và bảo vệ lợi ích chung trong chuỗi phân phối số.

### 6.6.1. Mất giao diện, mất dần dữ liệu khách hàng

Trong nền kinh tế trải nghiệm, mỗi lần một du khách đến địa phương đều là cơ hội để hình thành một mối quan hệ dài hạn. Địa phương có thể thu thập thông tin, duy trì liên lạc, mời quay trở lại và dần xây dựng một cộng đồng những người yêu mến nơi chốn. Quan hệ khách hàng vì thế là một tài sản chiến lược, không kém gì cơ sở hạ tầng hay tài nguyên du lịch. Quan hệ ấy nay bắt đầu bị cắt đứt. Người dùng không còn tìm kiếm trực tiếp trên trang web của địa phương, cũng không nhất thiết đặt phòng qua hệ thống của doanh nghiệp địa phương. Họ trò chuyện với

trợ lý AI, giao toàn bộ quá trình lựa chọn cho tác nhân, rồi hoàn tất giao dịch trên nền tảng trung gian. Địa phương chỉ xuất hiện ở điểm cuối của chuỗi cung ứng.

Khi mất quyền tiếp cận khách hàng, địa phương cũng mất khả năng xây dựng lòng trung thành. Mỗi lần khách quay lại đều phải cạnh tranh lại từ đầu trong thuật toán lựa chọn. Thay vì sở hữu một cộng đồng khách hàng của riêng mình, địa phương chỉ trở thành một lựa chọn có thể được thay thế bất cứ lúc nào nếu thuật toán đánh giá có một phương án khác phù hợp hơn. Mất quan hệ khách hàng và mất dữ liệu hành vi là hai mặt của cùng một sự việc. Những tín hiệu có giá trị nhất không phải việc khách đã đến, mà là khách đã cân nhắc những nơi nào khác, vì sao từ chối một phương án, chấp nhận mức giá nào và quyết định vào thời điểm nào. Toàn bộ nhóm tín hiệu ấy được ghi nhận ở lớp giao diện và lớp điều phối. Địa phương chỉ nhìn thấy kết quả cuối cùng.

Điều này khiến địa phương rơi vào một bất lợi ngày càng lớn. Không có dữ liệu hành vi, họ không thể hiểu đầy đủ nhu cầu của khách, khó cải thiện sản phẩm và càng khó cá nhân hóa trải nghiệm. Trong khi đó, các nền tảng sở hữu dữ liệu của hàng triệu người dùng liên tục huấn luyện thuật toán để đưa ra những dự đoán chính xác hơn, qua đó củng cố vị thế trung gian của mình. Quan trọng hơn, dữ liệu hành vi còn tạo ra một vòng lặp tích lũy. Càng nhiều người sử dụng nền tảng, nền tảng càng hiểu khách hàng. Càng hiểu khách hàng, đề xuất càng chính xác. Đề xuất càng chính xác, càng có nhiều người tiếp tục sử dụng. Trong vòng lặp đó, địa phương không chỉ mất dữ liệu của hiện tại mà còn mất khả năng tạo ra lợi thế cạnh tranh trong tương lai. Dữ liệu vì thế không còn là sản phẩm phụ của hoạt động du lịch mà trở thành tài sản cốt lõi quyết định quyền lực trong toàn bộ nền kinh tế tác nhân.

### 6.6.2. Quyền định giá bị thu hẹp

Trong nền kinh tế truyền thống, địa phương và doanh nghiệp du lịch có quyền quyết định giá trị của sản phẩm mình cung cấp. Giá phản ánh chất lượng dịch vụ, mùa vụ, thương hiệu và chiến lược kinh doanh của từng đơn vị. Người tạo ra trải nghiệm cũng là người quyết định cách trải nghiệm ấy được định giá.

Nền kinh tế tác nhân làm thay đổi cơ chế này. Khi AI trở thành người đại diện cho khách hàng, việc lựa chọn không còn dựa trên một mức giá cố định mà dựa trên bài toán tối ưu tổng thể. Thuật toán liên tục so sánh hàng nghìn lựa chọn theo thời gian thực, cân đối giữa giá, đánh giá, khoảng cách, khả năng phục vụ và sở thích cá nhân để tìm ra phương án được xem là tối ưu nhất.

Điều đó khiến quyền định giá dần dịch chuyển sang các nền tảng sở hữu dữ liệu và thuật toán. Hệ thống có thể gợi ý giảm giá để tăng xác suất được lựa chọn, tự động điều chỉnh mức giá theo nhu cầu từng thời điểm hoặc ưu tiên những đối tác chấp nhận mô hình chia sẻ doanh thu tốt hơn. Định giá động (dynamic pricing) không còn chỉ là công cụ của doanh nghiệp mà trở thành công cụ của người điều phối thị trường.

Doanh nghiệp vẫn nhập con số giá, nhưng nền tảng ngày càng quyết định mức giá nào có khả năng được nhìn thấy, được so sánh thuận lợi và được thị trường chấp nhận. Khi quyền ấy bị thu hẹp, địa phương rất khó duy trì giá trị thương hiệu dài hạn. Nếu mọi quyết định đều bị kéo về tối ưu ngắn hạn của thuật toán, điểm đến sẽ ngày càng cạnh tranh bằng giá thay vì cạnh tranh bằng bản sắc. Phần giá trị cao cấp dần bị bào mòn, trong khi quyền quyết định mức giá tối ưu nằm ở lớp điều phối chứ không còn ở nơi tạo ra trải nghiệm.

### 6.6.3. Từ nhà điều hành thành nhà cung cấp phía sau

Trong nhiều thập kỷ, địa phương giữ vai trò là người điều hành toàn bộ hệ sinh thái du lịch. Họ quy hoạch tài nguyên, đầu tư hạ tầng, xây dựng thương hiệu, điều phối doanh nghiệp và trực tiếp tạo lập môi trường để trải nghiệm diễn ra. Địa phương là trung tâm của chuỗi giá trị. Đây chính là quá trình chuyển từ người điều hành (operator) sang nhà cung ứng (supplier). Địa phương vẫn quản lý di sản, bảo vệ môi trường, duy trì hạ tầng và tổ chức trải nghiệm, nhưng những hoạt động tạo ra doanh thu có giá trị cao như thu hút khách, phân phối nhu cầu, giữ dữ liệu và tái bán hàng lại nằm ở các lớp phía trên.

Ở góc độ kiến trúc số, địa phương dần trở thành một nhà cung cấp phía sau. Họ cung cấp dữ liệu, năng lực phục vụ và khả năng thực thi cho các tác nhân AI, trong khi giao diện, quan hệ khách hàng và dòng doanh thu tiếp tục được kiểm soát bởi những nền tảng ở phía trước.

### 6.6.4. Từ khách sạn gia công đến địa phương gia công

Trong ngành sản xuất, nhiều doanh nghiệp sở hữu nhà máy hiện đại nhưng chỉ gia công cho thương hiệu khác. Họ tạo ra sản phẩm nhưng người tiêu dùng không biết tên họ. Giá trị lớn nhất thuộc về doanh nghiệp sở hữu thương hiệu, khách hàng và kênh phân phối. Đây là mô hình sản xuất theo đặt hàng của thương hiệu khác, quen gọi là OEM.

Một nguy cơ tương tự đang xuất hiện đối với các địa phương. Họ đầu tư vào cảnh quan, bảo tồn di sản, đào tạo nhân lực, phát triển dịch vụ và tạo nên những trải nghiệm độc đáo. Tuy nhiên,

khách hàng lại tiếp cận tất cả thông qua trợ lý AI hoặc nền tảng trung gian. Trong nhận thức của người dùng, nền tảng trở thành người "đưa họ đến", còn địa phương chỉ là nơi thực hiện dịch vụ.

Khi đó, địa phương có thể rơi vào trạng thái OEM hóa, làm ra trải nghiệm nhưng người khác sở hữu khách hàng, đầu tư tài sản nhưng người khác nắm dữ liệu, chịu chi phí duy trì nhưng người khác kiểm soát lợi nhuận có biên cao. Giá trị thương hiệu, dữ liệu và khả năng tái khai thác khách hàng đều dịch chuyển lên phía trên chuỗi giá trị.

Nếu xu hướng này tiếp diễn, nhiều địa phương sẽ chỉ còn cạnh tranh bằng chi phí vận hành và mức giá cung ứng, giống như các nhà máy gia công trong chuỗi sản xuất toàn cầu. Muốn tránh nguy cơ đó, địa phương không thể chỉ đầu tư vào tài sản vật chất mà còn phải xây dựng năng lực số: sở hữu dữ liệu, duy trì quan hệ trực tiếp với khách hàng, phát triển hạ tầng API và tạo ra những tài sản số mà các tác nhân AI không thể dễ dàng thay thế. Đây sẽ là điều kiện để địa phương không chỉ là nơi tạo ra trải nghiệm, mà còn giữ được một phần quyền lực trong chuỗi giá trị mới của nền kinh tế tác nhân.

***Khách sạn M (cái không có trong sổ sách):*** Khách sạn M biết chính xác bao nhiêu khách đã ở. Nó không biết bao nhiêu người đã cân nhắc rồi bỏ qua, cũng không biết vì lý do gì. Nếu tỷ lệ chuyển đổi trên nền tảng là 4%, thì với mỗi khách đến có khoảng 24 người đã nhìn thấy và không chọn.[1] Hai mươi bốn quan sát ấy là thứ quý nhất trong toàn bộ giao dịch. Chúng không thuộc về khách sạn M. Đó là lý do vì sao một cơ sở có thể phục vụ tốt

1. Phép tính minh hoạ của tác giả. Số 4% và con số 24 người dựa trên một tỷ lệ chuyển đổi giả định, không phải số liệu của một cơ sở lưu trú có thật.

suốt mười năm mà năng lực cạnh tranh không tăng lên tương ứng: nó cải thiện được cái nó thấy, không cải thiện được cái nó không thấy. Đón mười triệu lượt khách mà không có quyền hợp pháp để liên hệ lại với một ai, đó không phải là sở hữu một thị trường. Đó là một hợp đồng gia công.

## 6.7. AI TRẢ TIỀN CHO HẠ TẦNG KHẢ ĐỌC VÀ KHẢ HÀNH ĐỘNG

Nếu trong nền kinh tế công nghiệp, doanh nghiệp phải đầu tư nhà máy để sản xuất, còn trong nền kinh tế số, họ phải đầu tư hạ tầng công nghệ để hiện diện trên internet, thì trong nền kinh tế tác nhân, một lớp đầu tư mới xuất hiện là hạ tầng giúp địa phương trở nên khả đọc và khả hành động đối với máy (machine readability và actionability). Dữ liệu có cấu trúc, hệ thống khái niệm, đồ thị tri thức, giao diện lập trình ứng dụng, định danh thực thể, thông tin thời gian thực và các tín hiệu xác thực đều đòi hỏi nguồn lực để xây dựng, duy trì và cập nhật. Đây không còn là chi phí truyền thông mà đã trở thành một phần của hạ tầng kinh tế số.

Lợi ích từ lớp hạ tầng này không phân bổ đồng đều. Địa phương bỏ kinh phí để chuẩn hóa dữ liệu. Doanh nghiệp cập nhật thông tin dịch vụ. Cộng đồng tạo đánh giá và nội dung. Các nền tảng tổng hợp, xử lý rồi phân phối. Còn các tác nhân AI sử dụng toàn bộ lớp dữ liệu ấy để đưa ra quyết định. Giá trị được tạo ra bởi nhiều chủ thể nhưng không phải ai cũng thu được phần lợi ích tương xứng.

Vì vậy, câu hỏi quan trọng không còn là có nên đầu tư cho khả đọc hay không, bởi đó gần như đã trở thành điều kiện bắt buộc để tồn tại trong tập lựa chọn của máy. Câu hỏi lớn hơn là ai sẽ chi trả cho lớp hạ tầng này, ai sẽ được hưởng lợi từ nó, cùng với việc cơ chế phân bổ chi phí và lợi ích cần được thiết kế như thế nào để vừa khuyến khích đầu tư, vừa tránh việc toàn bộ giá trị tiếp tục tập trung vào các nền tảng và tác nhân điều phối. Đây sẽ là một trong những vấn đề kinh tế và chính sách quan trọng nhất của nền kinh tế tác nhân trong những năm tới.

### 6.7.1. Phần nào mang tính hạ tầng dùng chung, phần nào là đầu tư tư

Một trong những câu hỏi nền tảng của nền kinh tế tác nhân là lớp hạ tầng này nên được coi là hàng hóa công hay hàng hóa tư. Nếu đây là hàng hóa công, ngân sách nhà nước phải đầu tư. Nếu đây là hàng hóa tư, doanh nghiệp phải tự chi trả. Câu trả lời quyết định không chỉ nguồn vốn mà còn ảnh hưởng đến tốc độ chuyển đổi số của cả một địa phương.

Về bản chất, lớp hạ tầng ấy gồm nhiều tầng khác nhau. Những thành phần như bản đồ địa giới, địa danh chuẩn hóa, dữ liệu di sản, giao thông công cộng, quy hoạch, điểm tham quan, bảo tàng, sự kiện công hay hệ thống định danh địa phương phục vụ cho mọi tổ chức và mọi công dân. Nhiều thành phần trong số đó có đặc tính gần với hàng hóa công (public goods), việc một chủ thể sử dụng không làm giảm khả năng sử dụng của chủ thể khác, trong khi lợi ích xã hội thường lớn hơn khoản mà một nhà đầu tư riêng lẻ có thể thu lại. Về kỹ thuật vẫn có thể loại trừ người dùng bằng khóa truy cập hay hợp đồng, nên đặc tính đáng nói ở đây là tính phi cạnh tranh trong sử dụng và ngoại tác tích cực, không phải sự bất khả loại trừ. Nếu Nhà nước không đầu tư xây dựng lớp dữ liệu nền này, rất ít doanh nghiệp có động cơ thực hiện vì họ không thể độc quyền khai thác lợi ích thu được.

Ngược lại, dữ liệu về phòng còn trống của khách sạn, thực đơn nhà hàng, giá vé, tình trạng bàn, lịch tour, chính sách khuyến mại, API đặt dịch vụ hay hệ thống thanh toán lại tạo lợi thế cạnh tranh trực tiếp cho từng doanh nghiệp. Đây là hàng hóa tư (private goods). Chi phí xây dựng và duy trì những dữ liệu này hợp lý hơn khi do chính doanh nghiệp chi trả, bởi họ cũng là bên hưởng lợi trực tiếp từ việc được AI lựa chọn và giới thiệu.

Tuy nhiên, trên thực tế, ranh giới giữa công và tư ngày càng mờ. Một khách sạn cập nhật dữ liệu đầy đủ không chỉ giúp tăng doanh thu của chính mình mà còn làm tăng chất lượng của toàn bộ hệ sinh thái du lịch địa phương. Ngược lại, dữ liệu nền do Nhà nước công bố lại giúp doanh nghiệp giảm đáng kể chi phí tích hợp và phát triển sản phẩm. Giá trị vì thế luôn lan tỏa giữa khu vực công và khu vực tư. Vì vậy, mô hình hiệu quả là phân tầng trách nhiệm và đồng đầu tư theo từng tầng, không phải Nhà nước lo tất hay doanh nghiệp tự lo tất.

Điều cần nhấn mạnh là tầng dữ liệu nền có đặc tính của tài sản số dùng chung (digital commons): nhiều hệ thống dùng cùng lúc mà không làm hao mòn. Giá trị còn tăng lên khi có nhiều bên cùng khai thác vì mỗi lần dùng lại sinh thêm phản hồi. Nhưng dùng chung không có nghĩa là miễn phí. Nó vẫn cần vốn, cần người quản trị và cần bảo trì. Khác biệt nằm ở chỗ giá trị của nó không đến từ độc quyền khai thác mà từ việc làm giảm chi phí giao dịch cho hàng nghìn chủ thể cùng lúc.

Không phải mọi loại dữ liệu đều nên được tài trợ và quản trị theo cùng một cách. Mỗi lớp dữ liệu có tính chất khác nhau nên cũng cần một cơ chế đầu tư và chia sẻ khác nhau:

- ***Lớp nền*** bao gồm các dữ liệu như định danh thực thể, ranh giới hành chính, di sản, giao thông công cộng và sự kiện công. Đây là hạ tầng dữ liệu cơ bản nên phù hợp với nguồn ngân sách nhà nước, đồng thời cần được mở, miễn phí và có giấy phép sử dụng rõ ràng.

- ***Lớp chuẩn*** gồm lược đồ mô tả, hệ thống khái niệm và các giao thức kết nối giữa các hệ thống. Nhà nước nên giữ vai trò chủ trì xây dựng, doanh nghiệp tham gia góp ý và hoàn thiện. Các chuẩn này cần được công khai và trở thành yêu cầu bắt buộc đối với các dự án sử dụng ngân sách công.

- ***Lớp tổng hợp*** bao gồm kho dữ liệu ngành, thống kê dòng khách và các chỉ số điểm đến. Đây là lớp dữ liệu phù hợp với mô hình đồng đầu tư giữa nhà nước và doanh nghiệp. Dữ liệu tổng quan có thể mở cho cộng đồng, trong khi các tập dữ liệu chi tiết hơn có thể cung cấp theo cơ chế trả phí.

- ***Lớp vận hành*** chứa các dữ liệu thời gian thực như tồn kho, giá, tình trạng còn chỗ và khả năng phục vụ. Đây là tài sản trực tiếp của doanh nghiệp nên do doanh nghiệp đầu tư, với cơ chế chia sẻ theo hợp đồng hoặc giao diện kết nối (API) thay vì mở hoàn toàn.

- ***Lớp hành vi*** là lớp dữ liệu nhạy cảm nhất, bao gồm lịch sử tìm kiếm, hồ sơ khách hàng và các mô hình đề xuất. Doanh nghiệp là chủ thể đầu tư và khai thác, nhưng việc sử dụng phải tuân thủ chặt chẽ các quy định về bảo vệ dữ liệu cá nhân, không thể coi là dữ liệu mở.

### 6.7.2. Vì sao thị trường sẽ đầu tư thiếu

Nếu thiếu cơ chế phối hợp và khả năng thu hồi chi phí, thị trường sẽ đầu tư dưới mức xã hội mong muốn. Đặc điểm quyết định của đầu tư vào lớp dữ liệu nền là người bỏ tiền thường không phải người hưởng lợi nhiều nhất. Khi một địa phương chuẩn hóa dữ liệu, xây đồ thị tri thức hay công bố giao diện lập trình ứng dụng, lợi ích lan sang doanh nghiệp, nền tảng, nhà phát triển phần mềm, tác nhân AI và cả du khách. Đây là ngoại tác tích cực (externalities).

Chuỗi lan tỏa rất dễ thấy. Dữ liệu chính xác giúp máy trả lời tốt hơn. Máy trả lời tốt hơn giúp khách tìm được nơi phù hợp.

Khách tăng thì doanh thu của khách sạn, nhà hàng, vận tải và bán lẻ tăng theo, ngân sách địa phương cũng tăng. Giá trị được tạo ra cho rất nhiều chủ thể, trong khi chỉ một số ít bỏ vốn.

Chính vì thế thị trường sẽ đầu tư dưới mức xã hội mong muốn. Mỗi doanh nghiệp chỉ tính phần lợi ích của riêng mình khi quyết định chi, còn phần lớn giá trị lại rơi sang người khác. Kết quả là ai cũng chờ người khác làm trước, để rồi lớp hạ tầng chung không bao giờ đạt tới ngưỡng có ích. Đây là dạng thất bại thị trường quen thuộc, cùng họ với giáo dục, nghiên cứu cơ bản và hạ tầng giao thông. Cách xử lý cũng tương tự: khu vực công phải khởi động phần nền.

Ở chiều ngược lại còn có hiệu ứng tương thích của tiêu chuẩn chung. Một khách sạn dùng chuẩn dữ liệu chung thì lợi ích hạn chế. Khi hàng nghìn cơ sở cùng dùng một chuẩn, tác nhân có thể ghép chúng thành một hành trình hoàn chỉnh, khi đó giá trị của mỗi thành viên tăng theo quy mô mạng lưới. Điều này giải thích vì sao trợ cấp cho những người tham gia đầu tiên thường rẻ hơn nhiều so với chi phí sửa một hệ sinh thái đã phân mảnh.

### 6.7.3. Ba mô hình chia sẻ chi phí

Nếu khả đọc là một loại hạ tầng mới của nền kinh tế tác nhân thì câu hỏi tiếp theo là làm thế nào để chia sẻ chi phí xây dựng và vận hành hạ tầng ấy. Không một chủ thể nào có đủ động cơ và nguồn lực để đầu tư toàn bộ. Nhà nước thiếu dữ liệu vận hành chi tiết. Doanh nghiệp khó tự xây dựng các tiêu chuẩn chung. Cộng đồng tạo ra lượng lớn dữ liệu nhưng không có khả năng điều phối. Vì vậy, mô hình bền vững nhất là chia sẻ trách nhiệm giữa nhiều bên.

Mô hình quen thuộc nhất là hợp tác công tư, hiểu theo nghĩa rộng chứ không nhất thiết là một dự án PPP theo quy định pháp luật về đầu tư. Trong mô hình này, Nhà nước đầu tư các lớp hạ tầng nền như định danh địa phương, hệ thống khái niệm, tiêu chuẩn dữ liệu, giao diện lập trình ứng dụng công cộng, bản đồ số và cơ sở dữ liệu mở. Doanh nghiệp đầu tư các lớp dữ liệu kinh doanh, giao dịch, vận hành thời gian thực và các dịch vụ số. AI và các nền tảng hoạt động hiệu quả hơn nhờ hai lớp dữ liệu được kết nối, trong khi toàn bộ hệ sinh thái cùng hưởng lợi.

Tuy nhiên, nền kinh tế tác nhân còn mở ra một mô hình mới là hợp tác xã dữ liệu (data cooperative). Cụm này chỉ một mô hình quản trị và chia sẻ quyền lợi, không nhất thiết là một pháp nhân hợp tác xã theo quy định của pháp luật. Thay vì mỗi doanh nghiệp tự xây dựng một kho dữ liệu riêng, nhiều khách sạn, nhà hàng, điểm tham quan, đơn vị vận chuyển và tổ chức văn hóa có thể cùng đóng góp dữ liệu vào một nền tảng chung theo các tiêu chuẩn thống nhất. Mỗi thành viên vẫn giữ quyền sở hữu dữ liệu gốc nhưng đồng thời được hưởng lợi từ giá trị của toàn bộ mạng lưới.

Đối với địa phương, hợp tác xã dữ liệu đặc biệt phù hợp với các doanh nghiệp nhỏ và vừa. Một homestay riêng lẻ khó có khả năng xây dựng giao diện lập trình ứng dụng chuẩn, triển khai lược đồ dữ liệu hay tích hợp với hàng chục nền tảng AI. Nhưng nếu hàng trăm cơ sở cùng tham gia một hạ tầng dữ liệu chung, chi phí trên mỗi thành viên sẽ giảm mạnh, trong khi khả năng được máy đọc và lựa chọn lại tăng lên đáng kể. Quy mô tạo ra hiệu quả mà từng doanh nghiệp đơn lẻ không thể đạt được.

Về dài hạn, mô hình chia sẻ chi phí sẽ không chỉ dừng ở việc góp vốn xây dựng hạ tầng mà còn mở rộng sang chia sẻ lợi ích.

Những doanh nghiệp đóng góp nhiều dữ liệu chất lượng hơn có thể nhận được nhiều quyền truy cập hơn, nhiều dịch vụ hơn hoặc mức phí thấp hơn. Khi đó, đầu tư vào khả đọc không còn là một khoản chi bắt buộc mà trở thành một hình thức đầu tư tạo lợi nhuận cho cả hệ sinh thái.

Hợp tác ở hạ tầng và cạnh tranh ở dịch vụ sẽ trở thành nguyên tắc quan trọng. Các địa phương không nên cạnh tranh bằng những hệ thống dữ liệu khép kín, mà bằng chất lượng trải nghiệm được xây dựng trên một nền tảng dữ liệu chung đủ mạnh để AI có thể hiểu và khai thác.

Nếu ai đó nói rằng đây là lập luận chống công nghệ, tác nhân AI mang lại lợi ích rõ ràng cho du khách và cho cả địa phương, thì cần biết rằng chương này không tranh luận về việc tác nhân AI có ích hay không. Nó có ích. Phần lớn lợi ích ấy thuộc về người dùng cuối dưới dạng thời gian tiết kiệm được và quyết định tốt hơn. Đối tượng phân tích ở đây là phân phối, không phải tổng phúc lợi. Một công nghệ có thể vừa làm tăng tổng giá trị vừa chuyển phần lớn phần tăng thêm ấy về một nhóm chủ thể nhất định. Container hàng hải làm giàu cho thương mại toàn cầu, đồng thời xóa sổ nghề bốc xếp ở hàng chục cảng. Cả hai mệnh đề đều đúng. Nhiệm vụ của chính sách là làm việc với mệnh đề thứ hai mà không phủ nhận mệnh đề thứ nhất.

Tương tự, sẽ có người cho rằng các quy định sẽ giải quyết vấn đề này, như Liên minh châu Âu đã hành động, các nước khác sẽ theo sau. Đúng là quy định đã chứng minh là có tác dụng thật. Điều đó cần được ghi nhận. Đạo luật Thị trường Số buộc chấm dứt điều khoản ngang giá và mở đường cho thanh toán ngoài nền tảng. Cuối năm 2025, cơ quan cạnh tranh Anh gán quy chế

thị trường chiến lược cho hai kho ứng dụng lớn nhất.[1] Nhưng ba giới hạn cần được chỉ ra. Thứ nhất, quy định vận hành theo nhịp nhiều năm, còn thị phần tác nhân dịch chuyển theo nhịp nhiều tháng. Thứ hai, phần tô giảm ở một mắt xích không tự động chảy về mắt xích khác, như nghiên cứu về giá ứng dụng ở châu Âu đã cho thấy, dù nghiên cứu ấy do một bên có lợi ích liên quan công bố. Thứ ba, cũng là điều quan trọng nhất với người đọc cuốn sách này, phần lớn quy định hiện hành bảo vệ người tiêu dùng và nhà phát triển phần mềm, chứ không bảo vệ địa phương với tư cách một chủ thể kinh tế. Không có điều khoản nào buộc một tác nhân AI phải mô tả đúng một di sản.

- ***Phân bổ chi phí và lợi ích:***

• *Không đặt phí đối với truy cập phổ thông vào phần dữ liệu nền được phép công khai. Có thể thu phí đối với truy cập khối lượng lớn, dữ liệu đã xử lý gia tăng hoặc cam kết chất lượng dịch vụ đặc biệt.*

• *Không đồng đầu tư nếu hợp đồng không quy định rõ ai giữ dữ liệu khi hợp tác chấm dứt.*

• *Không chấp nhận điều khoản buộc địa phương chỉ cung cấp dữ liệu cho một nền tảng duy nhất.*

• *Không coi khoản chi duy trì dữ liệu nền là chi thường xuyên có thể cắt khi ngân sách khó khăn, vì gián đoạn làm mất toàn bộ giá trị đã tích lũy.*

• *Không đo thành công của hợp tác xã dữ liệu bằng số thành viên. Đo bằng tỷ lệ dữ liệu được cập nhật đúng hạn.*

1. Competition and Markets Authority của Anh, quyết định ngày 22 tháng 10 năm 2025 gán quy chế thị trường chiến lược cho nền tảng di động của Apple và của Google theo Đạo luật Thị trường Số, Cạnh tranh và Người tiêu dùng năm 2024, sau quyết định ngày 10 tháng 10 năm 2025 đối với dịch vụ tìm kiếm của Google. Quy chế có hiệu lực năm năm.

### 6.7.4. Ai hưởng lợi và ai có nghĩa vụ đóng góp

Nếu lợi ích lan tỏa rộng thì nghĩa vụ đóng góp cũng nên được phân bổ theo mức hưởng lợi, chứ không theo khả năng bị yêu cầu. Nguyên tắc thực dụng là ai rút được nhiều giá trị nhất từ một lớp dữ liệu thì nên góp nhiều nhất vào việc duy trì lớp ấy. Trên thực tế, các nền tảng lớn hưởng lợi nhiều nhất từ dữ liệu nền do ngân sách công tạo ra, trong khi phần đóng góp ngược lại thường chỉ dừng ở việc cho phép hiển thị.

Có ba công cụ khả thi mà không cần dựng thêm bộ máy. Thứ nhất, điều kiện có đi có lại trong hợp đồng: nền tảng muốn khai thác dữ liệu nền của địa phương thì phải trả lại dữ liệu tổng hợp về dòng khách ở mức không định danh. Thứ hai, gắn quyền truy cập tầng tổng hợp với mức đóng góp dữ liệu, để bên góp nhiều được dùng nhiều. Thứ ba, đưa yêu cầu về chuẩn mở và khả năng xuất dữ liệu vào điều kiện của mọi dự án dùng ngân sách, vì đó là chỗ Nhà nước có đòn bẩy thật.

### 6.7.5. Đo năng lực giữ giá trị

Toàn bộ chương sẽ chỉ là một mô tả nếu không quy được về vài con số mà một địa phương có thể tự đo mỗi quý. Bảng dưới đây đề xuất sáu chỉ số. Hai chỉ số đầu đo sự dịch chuyển dòng tiền, hai chỉ số tiếp theo đo khả năng giữ quan hệ và duy trì dữ liệu, hai chỉ số cuối đo rủi ro về mô tả và mức độ phụ thuộc cấu trúc. Các ngưỡng đi kèm không phải chuẩn mực đã được kiểm định mà là mốc do tác giả đề xuất để kích hoạt một cuộc rà soát. Bảng 6.4 tập hợp bộ chỉ số này.

**Bảng 6.4.** Bộ chỉ số đo năng lực giữ giá trị của một địa phương

| Chỉ số | Cách đo | Ngưỡng kích hoạt rà soát do tác giả đề xuất |
|---|---|---|
| **Tỷ lệ giao dịch qua trung gian** | Ước tính từ mẫu doanh nghiệp đại diện, dữ liệu thuế hoặc báo cáo theo nhóm ngành, không dùng con số tổng doanh thu du lịch | Vượt hai phần ba và tiếp tục tăng |
| **Chi phí phân phối thực** | Hoa hồng cộng phí hiển thị cộng phần giảm giá do điều kiện nền tảng cộng chi phí xử lý hủy phát sinh riêng từ kênh | Vượt một phần tư giá trị đặt chỗ |
| **Mức sở hữu quan hệ khách hàng** | Số khách duy nhất đã đồng ý cho liên hệ lại trên tổng số khách duy nhất | Dưới 5% |
| **Độ tươi dữ liệu** | Chia ba nhóm: dữ liệu tĩnh rà soát mỗi 6 đến 12 tháng, dữ liệu bán động mỗi 30 đến 90 ngày, dữ liệu vận hành trong ngày | Bất kỳ nhóm nào trễ hơn chu kỳ của chính nó |
| **Độ chính xác khi máy mô tả** | Điểm sai lệch có trọng số trên một bộ câu hỏi chuẩn, thử trên nhiều tác nhân AI, trọng số cao cho định danh, ranh giới và di sản | Xuất hiện sai lệch có trọng số cao, hoặc điểm tổng xấu đi hai quý liên tiếp |
| **Mức phụ thuộc một nền tảng** | Tỷ trọng của nền tảng lớn nhất trong tổng lượt đặt | Vượt một nửa |

*Nguồn: Tác giả xây dựng. Hai chỉ số đầu đo sự dịch chuyển dòng tiền, hai chỉ số tiếp theo đo khả năng giữ quan hệ và duy trì dữ liệu, hai chỉ số cuối đo rủi ro về mô tả và mức độ phụ thuộc cấu trúc. Các ngưỡng là mốc do tác giả đề xuất để khởi động một cuộc rà soát, không phải chuẩn mực đã được kiểm định. Nếu không có dữ liệu toàn địa phương cho chỉ số thứ nhất, có thể ước tính từ mẫu doanh nghiệp đại diện theo bốn nhóm lưu trú, vận tải, vé tham quan và hoạt động trải nghiệm. Điều quan trọng là dùng nhất quán một cách đo theo thời gian, không phải đạt độ chính xác tuyệt đối ngay từ kỳ đầu.*

Bốn nội dung dưới đây rút các chỉ số ấy về mức hành động cụ thể cho từng vai trò.

- ***Nếu bạn là chủ tịch tỉnh hoặc thành phố:***

• *Đặt một câu hỏi kiểm tra trong cuộc họp tới: địa phương có danh sách khách quay lại của riêng mình hay không, ai đang giữ nó.*

• *Xếp lớp dữ liệu nền và định danh dùng chung vào nhóm hạ tầng số thiết yếu của địa phương trong kế hoạch đầu tư công trung hạn, thay vì để hoàn toàn trong ngân sách xúc tiến.*

• *Đưa điều khoản xuất dữ liệu ở định dạng mở thành yêu cầu mặc định trong các hợp đồng do địa phương hoặc đơn vị trực thuộc ký kết, không chấp nhận điều khoản độc quyền hiển thị trong những hợp đồng ấy, đồng thời khuyến nghị doanh nghiệp tránh các điều khoản làm hạn chế khả năng phân phối đa kênh.*

- ***Nếu bạn là giám đốc sở quản lý du lịch:***

• *Xây một tầng dữ liệu nền dùng chung cho toàn tỉnh trước khi từng xã tự làm, vì chi phí lặp lại là khoản lãng phí lớn nhất của giai đoạn này.*

• *Chọn một rổ cố định gồm khoảng năm trợ lý AI đang được sử dụng rộng rãi, hỏi cùng một bộ câu hỏi về địa phương mỗi quý, lưu lại câu trả lời để theo dõi sai lệch. Chỉ thay đổi rổ khi có lý do rõ ràng, để giữ được khả năng so sánh theo thời gian.*

• *Đàm phán tập thể với nền tảng thay cho từng doanh nghiệp, vì quy mô là một trong số ít đòn bẩy giúp cải thiện sức thương lượng, bên cạnh chất lượng tồn kho, tỷ lệ chuyển đổi và khả năng bán trực tiếp.*

▪ ***Nếu bạn là chủ khách sạn, homestay hoặc điểm trải nghiệm:***

• *Tính lại chi phí thật của kênh nền tảng, gồm hoa hồng, khuyến mại bắt buộc và tỷ lệ hủy, rồi so với chi phí kênh trực tiếp.*

• *Giữ lại quyền liên hệ với khách sau chuyến đi trên cơ sở có sự đồng ý rõ ràng và một lý do có giá trị thật, chẳng hạn bảo hành dịch vụ hoặc chương trình khách quen, đúng theo quy định về bảo vệ dữ liệu cá nhân.*

• *Tham gia hạ tầng dữ liệu dùng chung của địa phương thay vì tự xây kết nối riêng với từng nền tảng.*

▪ ***Nếu bạn là giám đốc trung tâm xúc tiến đầu tư và du lịch:***

• *Chuyển một phần ngân sách chiến dịch sang tài sản dữ liệu dùng lại được, rồi báo cáo phần chuyển đổi ấy như một khoản đầu tư, không phải một khoản cắt giảm.*

• *Lập hồ sơ bằng chứng số cho những giá trị khó đo của địa phương, để chúng có mặt trong dữ liệu thay vì chỉ có mặt trong lời giới thiệu.*

• *Công bố một bộ dữ liệu mở tối thiểu về sự kiện, di sản và giao thông, vì đó là phần rẻ nhất nhưng lan tỏa nhất của lớp dữ liệu nền.*

Trước khi sang phần kết, một khoảng nghỉ. Ba việc vừa nêu đều đo được bằng số liệu ngân sách và dữ liệu hiện có của địa phương và người đọc nên thử đo ít nhất một việc trước khi đồng ý với phần còn lại.

## 6.8. KẾT LUẬN CHƯƠNG

Chương 4 giải thích cách máy hình thành tri thức về địa phương. Chương 5 phân tích cách máy phân bổ sự chú ý và dòng khách. Chương 6 khép lại phần cơ chế bằng câu hỏi ai giữ lại phần giá trị kinh tế. Sáu cơ chế được phân tích trong chương có thể đọc như một chuỗi, chia làm hai nhóm. Nhóm thứ nhất nói về quyền lực và tích lũy, cơ chế tách rời giải thích vì sao người tạo giá trị và người giữ giá trị không còn là một, cơ chế điểm nghẽn chỉ ra chỗ phần chênh lệch ấy đọng lại, cơ chế lãi kép của dữ liệu giải thích vì sao chỗ đọng ấy ngày càng sâu. Nhóm thứ hai nói về hệ quả đối với địa phương, cơ chế lũy thoái cho biết ai chịu chi phí nặng nhất khi cố gắng bước vào cuộc chơi, cơ chế nén giao diện giải thích vì sao những giá trị khó đo lại rơi ra khỏi cuộc so sánh. cơ chế tái trung gian hóa cảnh báo rằng mỗi lời hứa xóa bỏ trung gian thường kết thúc bằng một trung gian mới đứng gần điểm quyết định hơn. Điểm chung của cả sáu cơ chế là chúng đều mô tả cấu trúc, không mô tả ý đồ. Không cần giả định ai muốn làm hại địa phương thì các kết quả trên vẫn xảy ra.

Kết luận trung tâm là tạo ra giá trị không còn đồng nghĩa với giữ được giá trị. Nói cho chính xác thì đây là một xu hướng đang hình thành, không phải một trạng thái đã hoàn tất. Phí không mặc nhiên là tô, trung gian không đồng nghĩa với trích xuất, dữ liệu hóa không tất yếu làm mất bản sắc. Nguy cơ địa phương bị đẩy về vai trò gia công tăng mạnh khi mất đồng thời giao diện, dữ liệu và quyền giao dịch. Mất bất kỳ hai trong ba quyền ấy cũng đã tạo ra mức phụ thuộc đáng kể.

Hệ quả chiến lược vì thế khá gọn. Địa phương cần bảo vệ ba tài sản liên kết với nhau, quan hệ khách hàng, dữ liệu hành vi và khả năng tự hoàn tất giao dịch. Trong nhiều trường hợp, khả

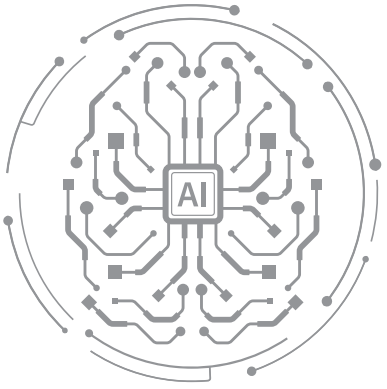
năng hoàn tất giao dịch là điều kiện để quan hệ khách hàng hình thành, còn quan hệ ấy lại tạo ra dữ liệu cho lần phục vụ tiếp theo. Khi cả ba vòng liên kết này đều thuộc về chủ thể khác, trải nghiệm địa phương dù tốt đến đâu cũng có nguy cơ chỉ còn là đầu vào cho một chuỗi do người khác điều phối.

Đó cũng là điểm nối tự nhiên sang Chương 7. Nếu Chương 6 trả lời câu hỏi "Ai giữ giá trị trong nền kinh tế tác nhân?", thì Chương 7 sẽ đi xa hơn để trả lời câu hỏi "Địa phương phải thiết kế chính mình như thế nào để trở thành một thực thể mà AI có thể nhận diện, hiểu, tin cậy, lựa chọn và hành động trên nó trong chuỗi giá trị mới?". Từ đây, trọng tâm không còn chỉ là kinh tế học của giá trị, mà chuyển sang kiến trúc chiến lược của một địa phương có khả năng cạnh tranh trong kỷ nguyên AI. Câu hỏi của thế kỷ trước là tạo ra được bao nhiêu. Câu hỏi của thế kỷ này là giữ lại được bao nhiêu.

## CHƯƠNG 7

# THIẾT KẾ ĐỊA PHƯƠNG CHO MÁY

*Từ khả đọc sang khả hành động: địa phương như một đồ thị, tồn kho giao dịch được, bốn điều kiện để tác nhân dám giao dịch và bảng điều khiển khả hành động thay cho bảng điều khiển kết quả.*


AI

Ba chương trước của cuốn sách lần lượt trả lời ba câu hỏi nền tảng của nền kinh tế tác nhân. Chương 4 giải thích máy hình thành tri thức về địa phương bằng cách nào và vì sao ký ức thuật toán khiến một hình ảnh có thể tồn tại rất lâu sau khi thực tế đã thay đổi. Chương 5 phân tích cách những hệ thống suy luận cạnh tranh để giành quyền đề xuất, từ đó quyết định dòng người, dòng tiền và sự chú ý sẽ được phân bổ về đâu. Chương 6 tiếp tục chỉ ra rằng khi quyền lựa chọn dịch chuyển sang tác nhân, giá trị kinh tế cũng dịch chuyển theo, khiến người tạo ra trải nghiệm không còn mặc nhiên là người giữ được khách hàng hay phần lớn lợi nhuận.

Đến đây, một câu hỏi thực tiễn xuất hiện. Nếu địa phương đã hiểu cách máy đọc, cách máy tin và cách máy lựa chọn, thì phải thiết kế mình như thế nào để trở thành một đối tượng mà máy có thể hành động? Đây là câu hỏi trung tâm của Chương 7.

Trong nền kinh tế trải nghiệm, một địa phương thành công khi con người muốn đến. Trong nền kinh tế tác nhân, điều đó vẫn đúng nhưng chưa đủ. Trước khi một con người quyết định, thường đã có một hệ thống suy luận tìm kiếm thông tin, so sánh các lựa chọn, kiểm tra khả năng cung ứng, đánh giá rủi ro và thậm chí hoàn tất việc đặt dịch vụ. Nếu toàn bộ chuỗi hành động ấy không thể được thực hiện một cách tự động, địa phương có nguy cơ bị loại khỏi chuỗi hành động trước khi giá trị kịp chuyển hóa thành giao dịch, dù bản thân trải nghiệm có hấp dẫn đến đâu.

Đó là lý do bộ lọc thứ tư của mô hình bốn bộ lọc không còn hỏi máy có nhìn thấy, có tin hay có đề xuất địa phương hay không. Nó hỏi máy có thể hành động trên địa phương ấy hay không và vì thế được gọi là bộ lọc khả hành động. Một địa phương có thể được nhận diện, được mô tả chính xác và thường xuyên được đề xuất, nhưng nếu máy không kiểm tra được tình trạng dịch vụ, không xác nhận được giá, không đặt được chỗ hoặc không hoàn tất được thanh toán thì chuỗi giá trị dừng lại ở bước gợi ý. Trong nền kinh tế tác nhân, được biết đến không còn đồng nghĩa với được lựa chọn. Chỉ khi máy có thể tiếp tục hành động thì giá trị mới đi trọn chuỗi phân phối mới, dù phần giao dịch do chính con người hoàn tất vẫn tồn tại và vẫn tạo ra giá trị thật.

Vì vậy, trọng tâm của chương này không còn là truyền thông, thương hiệu hay quảng bá. Đây là chương đầu tiên của cuốn sách chuyển từ phân tích sang thiết kế. Đối tượng được thiết kế là năng lực vận hành của địa phương dưới góc nhìn của máy, không phải một chiến dịch marketing. Câu hỏi không còn là "chúng ta muốn kể câu chuyện gì", mà là "hệ thống suy luận có thể thực hiện những hành động gì dựa trên dữ liệu của chúng ta". Đó là sự chuyển dịch từ tư duy nội dung sang tư duy hạ tầng, từ hình ảnh sang khả năng thực thi.

Khả hành động không thể do một chủ thể đơn lẻ tạo ra. Chính quyền giữ dữ liệu hành chính, quy hoạch và hạ tầng công. Doanh nghiệp giữ dịch vụ, tồn kho, giá và năng lực phục vụ. Các nền tảng số cung cấp giao diện và chuẩn kết nối. Hạ tầng dữ liệu bảo đảm thông tin lưu thông theo thời gian thực. Chỉ khi những thành phần ấy vận hành như một hệ thống thống nhất thì địa phương mới trở thành thực thể mà tác nhân có thể đọc, suy luận và hành động trọn vẹn. Đây không còn là bài toán của riêng ngành du lịch, công nghệ thông tin hay truyền thông mà là bài toán điều phối của cả hệ sinh thái.

Các mục tiếp theo lần lượt xem xét bảy câu hỏi: thế nào là một điểm đến sẵn sàng cho máy, địa phương được biểu diễn ra sao ở dạng dữ liệu, các thực thể nối với nhau bằng gì, thứ máy giao dịch thực chất là gì, cơ chế nào cho phép máy đặt được dịch vụ, quản trị địa phương bằng bảng chỉ số nào và làm sao kiểm toán chính mô hình mà máy đang dựng về địa phương.

Nếu Chương 6 trả lời câu hỏi ai thu được giá trị kinh tế, thì Chương 7 trả lời câu hỏi địa phương phải được thiết kế thế nào để vẫn nằm trong chuỗi tạo ra giá trị ấy. Khả hành động không phải một tiện ích bổ sung sau cùng mà là điều kiện để trải nghiệm, thương hiệu và tài sản của địa phương được chuyển hóa thành những hành động có hiệu lực, trong đó giao dịch là biểu hiện kinh tế quan trọng nhất. Đó cũng là lý do đầu tư cho khả hành động không thể thay bằng đầu tư truyền thông và là lý do tương lai của quản trị địa phương phụ thuộc ngày càng nhiều vào khả năng dựng một liên minh bền vững giữa chính quyền, doanh nghiệp và hạ tầng số.

Trong nền kinh tế trải nghiệm, địa phương được thiết kế để con người muốn đến. Trong nền kinh tế tác nhân, nó còn phải được thiết kế để máy có thể làm việc cùng. Bộ lọc cuối cùng không hỏi địa phương có hấp dẫn hay không. Nó hỏi máy có hoàn thành được việc hay không.

## 7.1. ĐIỂM ĐẾN SẴN SÀNG CHO MÁY NGHĨA LÀ GÌ

Trong phần lớn lịch sử phát triển du lịch, một điểm đến được coi là tốt khi có tài nguyên hấp dẫn, hạ tầng thuận tiện, dịch vụ chất lượng và thương hiệu đủ mạnh để thu hút du khách. Những tiêu chí ấy vẫn đúng, nhưng chúng được xây dựng trên một giả định quan trọng: con người là chủ thể trực tiếp tìm kiếm, so sánh và ra quyết định. Chính con người đọc tài liệu giới thiệu, truy cập trang web, xem video quảng bá, hỏi bạn bè rồi mới lựa chọn nơi sẽ đến.

Nền kinh tế tác nhân làm thay đổi trình tự đó. Trước khi một du khách nhìn thấy một điểm đến, thường đã có một hoặc nhiều hệ thống suy luận thay họ đọc dữ liệu, đánh giá các phương án và chuẩn bị sẵn một tập lựa chọn. Điều này không có nghĩa máy thay thế hoàn toàn con người, mà có nghĩa ngày càng nhiều quyết định ban đầu được hình thành trước khi con người tiếp xúc trực tiếp với thông tin. Vì vậy, một điểm đến tốt đối với con người chưa chắc đã là một điểm đến tốt đối với máy.

Sự khác biệt này không nằm ở chất lượng trải nghiệm mà nằm ở khả năng biến trải nghiệm thành hành động. Một khu nghỉ dưỡng có thể được đánh giá rất cao nhưng nếu hệ thống suy luận không xác định được tình trạng phòng, không biết giá hiện tại, không thể đặt chỗ hoặc không thể xác nhận giao dịch, thì đối với tác nhân, khu nghỉ dưỡng ấy là một lựa chọn chưa hoàn chỉnh. Ngược lại, một dịch vụ có trải nghiệm chưa thật sự vượt trội nhưng có dữ liệu đầy đủ, tình trạng luôn được cập nhật và toàn bộ quy trình giao dịch có thể thực hiện tự động lại có nhiều cơ hội xuất hiện trong chuỗi lựa chọn hơn. Tiêu chuẩn cạnh tranh vì thế đã mở rộng chứ không thay thế. Chất lượng trải nghiệm quyết định giá trị được tạo ra, còn khả năng thực thi quyết định giá trị ấy có đi trọn chuỗi phân phối hay không.

Đó là lý do cần phân biệt giữa một điểm đến nổi tiếng và một điểm đến sẵn sàng cho máy. Nổi tiếng là trạng thái của nhận thức, còn sẵn sàng cho máy là trạng thái của năng lực vận hành. Một địa phương có thể được hàng triệu người biết đến nhưng vẫn chưa sẵn sàng để các tác nhân số sử dụng, bởi giữa việc biết đến một nơi và có thể hoàn tất một chuỗi hành động tại nơi đó là một khoảng cách rất lớn.

Khoảng cách ấy ngày càng trở thành yếu tố quyết định trong nền kinh tế tác nhân. Máy không chỉ cần biết địa phương tồn tại, mà còn phải biết địa phương đang mở hay đóng, dịch vụ nào còn khả dụng, mức giá nào đang áp dụng, ai chịu trách nhiệm cung cấp, điều kiện đặt chỗ ra sao và kết quả giao dịch có được xác nhận hay không. Nói cách khác, khi phải hành động, một câu chuyện hấp dẫn là chưa đủ. Máy còn cần một hệ thống có thể vận hành.

Vì vậy, khái niệm "điểm đến sẵn sàng cho máy" không nên được hiểu như một khái niệm công nghệ. Đây trước hết là một khái niệm quản trị. Nó phản ánh mức độ mà toàn bộ hệ sinh thái của địa phương, từ cơ quan quản lý, doanh nghiệp, đơn vị vận hành hạ tầng đến các nhà cung cấp dịch vụ, có thể phối hợp để biến tài sản và trải nghiệm thành những đối tượng mà hệ thống suy luận không chỉ hiểu được mà còn có thể truy vấn và giao dịch được.

Đây cũng chính là ý nghĩa của bộ lọc thứ tư trong mô hình bốn bộ lọc. Nếu ba bộ lọc trước trả lời các câu hỏi máy có nhìn thấy không, máy có tin không và máy có đề xuất không, thì bộ lọc cuối cùng đặt ra câu hỏi quan trọng nhất: máy có thể hoàn thành công việc hay không. Chỉ khi câu trả lời là có, địa phương mới thật sự bước vào chuỗi giá trị của nền kinh tế tác nhân. Khi đó, mục tiêu của thiết kế địa phương không còn là làm cho mình xuất hiện đẹp hơn trước máy, mà là làm cho máy có thể làm việc cùng mình.

Trong nền kinh tế trải nghiệm, địa phương cạnh tranh để được nhớ. Trong nền kinh tế tác nhân, nó cạnh tranh để được giao việc.

Trong chương này, được giao việc là trạng thái mà một tác nhân lựa chọn một thực thể hoặc một địa phương để thực hiện một nhiệm vụ cụ thể thay cho người dùng, từ lập lịch trình, kiểm tra khả dụng, giữ chỗ cho đến hoàn tất giao dịch. Được giao việc đi xa hơn được đề xuất vì bao gồm cả phần thực thi, nhưng là một trạng thái khó đạt hơn và do đó chỉ áp dụng cho một tập thực thể hẹp hơn. Về tập hợp, được biết đến chứa được đề xuất và được đề xuất chứa được giao việc.

### 7.1.1. Từ "được biết đến" sang "được giao việc"

Trong suốt lịch sử tiếp thị địa phương, mục tiêu lớn nhất là được biết đến. Các địa phương đầu tư vào quảng bá, sự kiện, truyền thông và thương hiệu để gia tăng mức độ nhận diện. Thành công thường được đo bằng lượng người đã nghe tên, lượng tìm kiếm, số lượt xem, số bài báo hay số khách ghé thăm. Logic rất đơn giản: càng nhiều người biết đến thì cơ hội được lựa chọn càng lớn. Logic ấy được hình thành trong một thế giới mà con người là người trực tiếp ra quyết định.

Sự khác biệt giữa hai khái niệm này rất lớn. Một địa phương có thể xuất hiện dày đặc trên truyền thông, được hàng triệu người biết đến và sở hữu một thương hiệu mạnh, nhưng nếu tác nhân không thể truy cập dữ liệu, không thể xác nhận dịch vụ, không thể kiểm tra tình trạng còn chỗ hay không thể hoàn tất giao dịch, thì quá trình lựa chọn vẫn dừng lại. Địa phương được nhìn thấy nhưng không được sử dụng.

Ngược lại, một điểm đến ít nổi tiếng hơn nhưng có dữ liệu đầy đủ, thông tin nhất quán, dịch vụ luôn cập nhật và khả năng

kết nối trực tiếp với các hệ thống suy luận có thể được tác nhân ưu tiên trong nhiều tình huống. Không phải vì trải nghiệm của nó tốt hơn, mà vì nó giúp máy hoàn thành công việc với chi phí thấp hơn, ít rủi ro hơn và xác suất thành công cao hơn.

Sự khác biệt giữa hai trạng thái này là khác biệt về loại tài sản kinh tế. Được biết đến tạo ra xác suất được cân nhắc, một tài sản mang tính thống kê và dễ bốc hơi. Được giao việc tạo ra quyền tiếp nhận nhu cầu và cơ hội tham gia vào dòng tiền phát sinh từ nhu cầu ấy. Nói vậy chưa phải nói rằng người được giao việc sẽ giữ được phần lớn giá trị: Chương 6 đã cho thấy phần chênh lệch có thể vẫn nằm ở nền tảng, đại lý hoặc lớp thanh toán. Điều thay đổi là vị trí trong chuỗi giao dịch, không phải tỷ lệ phân chia. Trước đây truyền thông tạo ra sự chú ý, còn con người chuyển sự chú ý thành giao dịch. Nay một phần ngày càng lớn của khâu chuyển hóa ấy được trung gian bởi máy, nên thứ quyết định không còn chỉ là địa phương được nhắc đến bao nhiêu lần mà còn là địa phương có được ủy nhiệm hay không.

Bởi vậy, cuộc cạnh tranh mới không còn là cuộc đua giành thêm vài triệu lượt xem hay vài nghìn bài viết. Đó là cuộc đua trở thành một đối tác đáng tin cậy của hệ thống suy luận. Trong nền kinh tế tác nhân, thương hiệu mạnh không còn chỉ được định nghĩa bằng mức độ nổi tiếng. Nó còn thể hiện ở mức độ mà các tác nhân sẵn sàng đưa địa phương vào quyết định và chuỗi giao dịch. Được biết đến là bước khởi đầu. Được giao việc là biểu hiện kinh tế rõ nhất của khả hành động.

Con người vẫn là người quyết định cuối cùng, nên lo lắng về máy là quá sớm. Đúng ở phần kết luận và sai ở phần quy trình. Con người vẫn bấm nút xác nhận, nhưng tập lựa chọn đặt trước mặt họ đã được thu hẹp từ nhiều bước trước đó. Ảnh hưởng của một hệ thống không nằm ở chỗ nó ra quyết định thay ai, mà ở chỗ nó quyết định những phương án nào được nhìn thấy. Giới

hạn của lập luận nằm ở chỗ hiện chưa có dữ liệu đáng tin cậy đo tỷ lệ các quyết định du lịch tại các điểm đến đã bị tác nhân định hình từ giai đoạn đầu. Vì vậy có thể xác định hướng tác động nhưng chưa nên khẳng định độ lớn của nó.

### 7.1.2. Bốn cấp độ trưởng thành của một điểm đến

Không phải mọi địa phương đều tham gia nền kinh tế tác nhân theo cùng một mức độ. Có nơi gần như vô hình trước các hệ thống suy luận. Có nơi được nhận diện khá đầy đủ nhưng tác nhân chưa thể hoàn tất hành động. Cũng có nơi đã trở thành một phần của chuỗi tự động từ tìm kiếm, lập kế hoạch đến xác nhận hoặc giao dịch. Vì vậy, thay vì chỉ hỏi một địa phương có "chuyển đổi số" hay chưa, cần một mô hình trưởng thành phản ánh đúng cách máy nhìn và sử dụng địa phương. Có thể hình dung quá trình ấy qua bốn cấp độ liên tiếp.

***Cấp độ thứ nhất*** là có mặt (present). Đây là điều kiện tối thiểu để một địa phương tồn tại trong thế giới số. Các thực thể như điểm tham quan, khách sạn, nhà hàng, sự kiện hay tuyến giao thông đã có dữ liệu, có định danh và có thể được hệ thống tìm thấy. Tuy nhiên, sự hiện diện này mới chỉ trả lời câu hỏi: địa phương có tồn tại hay không. Nhiều địa phương hiện nay mới dừng ở cấp độ này. Họ có trang web, có bản đồ, có thông tin rải rác trên nhiều nền tảng, nhưng dữ liệu thiếu nhất quán, thiếu cấu trúc và thường xuyên không được cập nhật.

***Cấp độ thứ hai*** là được hiểu (understood). Không chỉ biết địa phương tồn tại, máy còn hiểu địa phương là gì. Các thực thể được phân loại đúng, quan hệ giữa chúng được xác lập và thông tin từ nhiều nguồn có thể đối chiếu với nhau. Hệ thống suy luận bắt đầu hình thành một mô hình tương đối đầy đủ về địa

phương: đâu là điểm tham quan, đâu là bảo tàng, đâu là khách sạn, sự kiện nào diễn ra ở đâu, doanh nghiệp nào vận hành dịch vụ nào. Đây là kết quả của ba chương trước của cuốn sách, nơi khả đọc, thẩm quyền và ký ức thuật toán tạo nên năng lực nhận thức của máy.

***Cấp độ thứ ba*** là được đề xuất (recommended). Khi đã hiểu, hệ thống mới có thể so sánh và lựa chọn. Địa phương không chỉ tồn tại trong cơ sở dữ liệu mà còn xuất hiện trong tập ứng viên mỗi khi người dùng yêu cầu một hành trình, một nơi lưu trú hay một trải nghiệm phù hợp. Đây là cấp độ mà các thuật toán xếp hạng, các hệ thống gợi ý và các tác nhân bắt đầu phân phối sự chú ý. Tuy nhiên, được đề xuất vẫn chưa đồng nghĩa với tạo ra giá trị kinh tế. Một gợi ý chỉ mới là khả năng. Giao dịch mới là kết quả.

***Cấp độ thứ tư*** là có thể hành động (actionable). Đây là cấp độ cao nhất và cũng là trọng tâm của chương này. Ở cấp độ này, hệ thống suy luận không chỉ đề xuất mà còn có thể tiếp tục hành động: kiểm tra dịch vụ còn hay hết, biết mức giá hiện tại, biết điều kiện sử dụng, giữ chỗ, xác nhận và hoàn tất. Giao dịch là dạng hành động kinh tế quan trọng nhất ở cấp độ này, nhưng không phải dạng duy nhất. Tác nhân còn có thể lập lịch, điều phối, cảnh báo, nộp hồ sơ dịch vụ công hay xác nhận thông tin mà không phát sinh thanh toán. Điểm chung là máy làm thay đổi được trạng thái của thế giới thật chứ không chỉ mô tả nó. Chính tại đây, khả đọc chuyển hóa thành khả hành động.

Bốn cấp độ này tạo thành một chuỗi phụ thuộc lẫn nhau, không phải bốn trạng thái tách biệt. Không thể được hiểu nếu chưa có mặt. Không thể được đề xuất nếu chưa được hiểu. Và cũng chưa thể hành động được nếu mới chỉ được đề xuất. Mỗi cấp độ đều kế thừa thành quả của cấp độ trước và bổ sung một

năng lực mới. Chi tiết đáng lưu ý nằm ở chỗ khác. Nhiều nguồn lực được đầu tư để số hóa dữ liệu, xây dựng cổng thông tin hay quảng bá trực tuyến, nhưng rất ít địa phương thiết kế toàn bộ hệ thống để tiến tới cấp độ cuối cùng. Họ giúp máy biết mình là ai nhưng chưa giúp máy làm được việc cùng mình.

Đó cũng là lý do chương này tập trung vào khái niệm khả hành động. Trong nền kinh tế tác nhân, thước đo trưởng thành của một điểm đến không còn dừng ở mức độ hiện diện hay mức độ nổi tiếng. Thước đo cuối cùng là liệu hệ thống suy luận có đủ tin tưởng và đủ khả năng để giao cho địa phương thực hiện một nhiệm vụ có hiệu lực trong thế giới thật hay không. Chỉ khi đạt tới cấp độ ấy, địa phương mới thực sự bước vào chuỗi giá trị mới của nền kinh tế tác nhân. Bảng 7.1 mô tả bốn cấp độ trưởng thành theo thứ tự ấy.

### 7.1.3. Bộ lọc cuối cùng

Nếu ba bộ lọc đầu tiên quyết định một địa phương có được nhìn thấy, được hiểu và được đề xuất hay không, thì bộ lọc cuối cùng quyết định một điều quan trọng hơn nhiều: hệ thống suy luận có tiếp tục sử dụng địa phương ấy hay không. Đây là điểm phân biệt giữa một thực thể chỉ tồn tại trong tri thức của máy và một thực thể thực sự tham gia vào nền kinh tế tác nhân.

Điều này giải thích một hiện tượng ngày càng phổ biến: nhiều địa phương, nhiều khách sạn hay nhiều điểm tham quan rất nổi tiếng nhưng lại không xuất hiện trong các chuỗi hành động tự động của tác nhân. Vấn đề không nằm ở danh tiếng. Vấn đề nằm ở chỗ danh tiếng không giúp máy hoàn thành công việc.

Hãy hình dung một tác nhân được giao nhiệm vụ xây dựng lịch trình cho một gia đình trong ba ngày. Hệ thống có thể biết

**Bảng 7.1.** Bốn cấp độ trưởng thành của một điểm đến trước máy

| Cấp độ | Câu hỏi máy đặt ra | Điều kiện phải có | Dấu hiệu chưa đạt |
|---|---|---|---|
| **1. Có mặt** (Present) | Nơi này có tồn tại không | Thực thể được định danh, có tọa độ, tìm thấy được | Thông tin rải rác nhiều nguồn, không có mã định danh chung |
| **2. Được hiểu** (Understood) | Nơi này là cái gì | Phân loại đúng, quan hệ được khai báo, nhiều nguồn đối chiếu được | Máy mô tả sai loại hình hoặc gộp nhầm với thực thể khác |
| **3. Được đề xuất** (Recommended) | Có nên đưa vào tập lựa chọn không | Đủ tín hiệu chất lượng, đủ độ tin cậy, đủ độ phủ chủ đề | Được nhắc tên nhưng không xuất hiện khi có yêu cầu cụ thể |
| **4. Có thể hành động** (Actionable) | Tôi có hoàn thành được việc không | Trạng thái hiệu lực, chủ thể xác nhận, quyền thực hiện và cơ chế phản hồi kết quả. Riêng với giao dịch thương mại cần thêm tồn kho, giá và thanh toán | Tác nhân bắt đầu nhưng không làm thay đổi được trạng thái hoặc không nhận được xác nhận kết quả |

rất rõ một điểm đến nổi tiếng, có hàng chục nghìn bài đánh giá và xuất hiện dày đặc trên truyền thông. Nhưng khi tác nhân cần kiểm tra giờ mở cửa, số lượng vé còn lại, khả năng đặt trước, giá theo từng khung giờ hay xác nhận thanh toán, chuỗi hành động lập tức bị đứt nếu những thông tin ấy không tồn tại hoặc không thể truy cập theo cách máy sử dụng được. Trong tình huống đó, hệ thống thường sẽ chuyển sang một lựa chọn khác có ít danh tiếng hơn nhưng có thể hoàn thành toàn bộ quy trình với độ tin cậy cao hơn.

Đó là logic rất khác với cách con người lựa chọn. Con người có thể chấp nhận tìm kiếm thêm, gọi điện hỏi, nhắn tin xác nhận

hoặc chờ phản hồi. Máy được thiết kế để giảm tối đa những bước không chắc chắn. Mỗi lần phải dừng lại để chờ xác minh là một lần chi phí xử lý tăng lên và xác suất hoàn thành nhiệm vụ giảm xuống. Vì vậy, trong nhiều trường hợp, tác nhân không ưu tiên lựa chọn tốt nhất theo cảm nhận của con người, mà ưu tiên lựa chọn có thể thực hiện được với mức rủi ro thấp nhất.

Bộ lọc cuối cùng vì thế không đánh giá mức độ nổi tiếng, mà đánh giá mức độ có thể hành động. Đối với hệ thống suy luận, một dịch vụ không thể đặt, một giá không thể xác minh, một lịch hoạt động không được cập nhật hay một tồn kho không thể kiểm tra đều là những tín hiệu làm giảm độ tin cậy của toàn bộ chuỗi quyết định. Mỗi điểm đứt nhỏ đều có thể khiến một địa phương bị loại khỏi tập lựa chọn cuối cùng.

Điều này cũng làm thay đổi ý nghĩa của cạnh tranh. Trong nền kinh tế trải nghiệm, địa phương cạnh tranh để được chú ý. Trong nền kinh tế tác nhân, nó còn phải cạnh tranh để được ủy nhiệm thực hiện hành động. Quyền lực kinh tế vì thế không dừng ở việc được nhắc đến mà nằm ở việc được giao thực hiện hành động, đặc biệt là những hành động dẫn tới giao dịch.

Đây cũng là bộ lọc mà đầu tư truyền thông không thể thay thế và là bộ lọc khó vượt qua nhất. Ba bộ lọc đầu chủ yếu liên quan đến dữ liệu và tri thức, còn bộ lọc thứ tư liên quan đến năng lực vận hành của cả hệ sinh thái. Một chiến dịch quảng bá làm tăng lượng tìm kiếm nhưng không tự tạo ra dữ liệu thời gian thực, không kết nối tồn kho, không đồng bộ lịch hoạt động và không giúp tác nhân hoàn tất việc đặt dịch vụ. Muốn vượt qua bộ lọc này, cả hệ sinh thái phải bảo đảm tác nhân đi được từ yêu cầu đến một kết quả được xác nhận và không một doanh nghiệp hay cơ quan nào làm được điều đó một mình.

Bộ lọc cuối. Xác suất một địa phương được tác nhân chọn để thực hiện một nhiệm vụ bằng tích xác suất vượt qua từng bộ lọc: phân giải, chứng thực, xếp hạng và khả hành động. Vì là một phép nhân, kết quả đặc biệt nhạy với bộ lọc yếu nhất. Mọi bộ lọc đều góp phần làm giảm tích số, nhưng một bộ lọc bằng không sẽ triệt tiêu toàn bộ kết quả dù các bộ lọc còn lại đạt điểm rất cao.

Không thay thế giữa truyền thông và khả hành động. Chi tiêu cho truyền thông và chi tiêu cho hạ tầng khả hành động không phải hai cách khác nhau để mua cùng một thứ. Khi bộ lọc thứ tư chưa thông, hiệu suất biên của đồng chi tiêu truyền thông tiếp theo giảm dần về không, bởi mọi sự chú ý tạo thêm đều dừng lại ở bước gợi ý. Quan hệ giữa hai khoản chi là bổ sung, không phải thay thế.

Các phép thử trên những bộ chuẩn hàng đầu cho thấy du lịch vẫn là một trong những môi trường khó nhất đối với tác nhân AI. Trên WebVoyager, trong cấu hình Agent-E được công bố năm 2024, cùng một tác nhân đạt 70,7% trên trang thương mại điện tử và 85,7% trên nền tảng học liệu, nhưng chỉ đạt 27,3% trên trang đặt phòng du lịch.[1] Đáng chú ý, tỷ lệ thất bại mà chính tác nhân không tự nhận biết lên tới gần 70%.[2] Xu hướng tương tự cũng xuất hiện trên WebArena, bộ chuẩn gồm 812 nhiệm vụ do Đại học Carnegie Mellon công bố năm 2023. Thành tích của các tác nhân đơn lẻ đã tăng mạnh, từ 14,4% với cấu hình GPT-4 ban

1. Tamer Abuelsaad và cộng sự, Emergence AI, "Agent-E: From Autonomous Web Navigation to Foundational Design Principles in Agentic Systems", arXiv:2407.13032, tháng 7 năm 2024, https://arxiv.org/abs/2407.13032. Bộ chuẩn WebVoyager gồm 643 nhiệm vụ tại 15 trang, tác nhân chỉ dùng văn bản, mô hình GPT-4-Turbo. Booking 27,3%, Google Flights 35,7%, Amazon 70,7%, Coursera 85,7%, WolframAlpha 95,7%, toàn bộ 73,1%. Bài tiền ấn phẩm, chưa qua bình duyệt.

2. Cùng nguồn đã dẫn ở chú thích 1 trang 455. Nhóm tác giả tách thất bại thành loại tác nhân tự nhận biết và loại không nhận biết. Trên trang đặt phòng, tổng tỷ lệ thất bại là 72,7%, trong đó 4,5% tự nhận biết và 68,2% không tự nhận biết. Trên toàn bộ bộ chuẩn, tác nhân tự nhận biết hơn 52% số lần thất bại, nên mức 68,2% là một ngoại lệ.

đầu lên khoảng 60% trên các bảng xếp hạng công khai cập nhật đến giữa năm 2026. Tuy nhiên, con số này vẫn thấp hơn đáng kể so với 78,24% của con người trên cùng bộ nhiệm vụ.[1] Các kết quả này phụ thuộc vào mô hình, cấu hình và điều kiện thử nghiệm, vì vậy nên được hiểu như thứ tự về năng lực hơn là những giá trị tuyệt đối. Điều quan trọng là một quy luật nhất quán chứ không phải tác nhân nào hay mô hình nào đang dẫn đầu: các tác vụ du lịch khó hơn đáng kể so với nhiều lĩnh vực khác. Chúng đòi hỏi xử lý đồng thời nhiều ràng buộc như lịch trình, giá, tình trạng còn chỗ, chính sách hủy, phương thức thanh toán và khả năng phối hợp giữa nhiều dịch vụ, trong khi trạng thái của hệ thống có thể thay đổi ngay trong quá trình giao dịch. Chính vì vậy, chất lượng dữ liệu và khả năng chuẩn hóa trở thành yếu tố quyết định. Khi dữ liệu sạch, nhất quán và có thể hành động bằng máy, tác nhân AI giảm đáng kể số bước suy luận, ít mắc lỗi hơn và có khả năng hoàn tất giao dịch cao hơn. Trong những bài toán nhiều ràng buộc như du lịch, lợi thế cạnh tranh không chỉ nằm ở mô hình AI mạnh hơn mà còn ở hạ tầng dữ liệu tốt hơn.

Tác nhân hiện còn quá kém, chưa đáng để đầu tư hạ tầng. Phần mô tả hiện trạng là chính xác. Trên bộ tiêu chuẩn WebArena, tỷ lệ hoàn thành nhiệm vụ của tác nhân đơn lẻ mới ở khoảng 60% so với 78% của con người và khoảng cách giữa phòng thí nghiệm với môi trường thật còn lớn hơn. Nhưng kết luận rút ra lại ngược với trực giác. Chính vì tác nhân còn yếu, nó càng tránh những nơi khó và càng dồn về những nơi mà chuỗi hành động chắc chắn hoàn tất. Năng lực máy càng thấp thì phần thưởng dành cho hạ tầng sạch càng cao.

1. Shuyan Zhou và cộng sự, "WebArena: A Realistic Web Environment for Building Autonomous Agents", arXiv:2307.13854, Đại học Carnegie Mellon, tháng 7 năm 2023: 812 nhiệm vụ dựng từ 241 khuôn mẫu, tác nhân GPT-4 đạt 14,41% trong khi con người đạt 78,24%. Các con số cập nhật về sau lấy từ bảng xếp hạng công khai, phụ thuộc mô hình và cấu hình nên không so sánh trực tiếp được với kết quả gốc.

## 7.2. MÔ HÌNH VẬN HÀNH CỦA ĐỊA PHƯƠNG Ở DẠNG DỮ LIỆU

Nếu ở mục trước, chương sách khẳng định rằng một địa phương chỉ thật sự tồn tại trong nền kinh tế tác nhân khi máy có thể hành động trên nó, thì bước tiếp theo là đặt ra câu hỏi quan trọng hơn: máy hình dung địa phương như thế nào để có thể đưa ra những hành động ấy? Câu trả lời không nằm ở giao diện ứng dụng, cũng không nằm ở một giao diện lập trình ứng dụng (API) hay một mô hình AI cụ thể, mà nằm ở mô hình dữ liệu logic mà hệ thống sử dụng để biểu diễn thế giới.

Máy không trải nghiệm địa phương như một tổng thể sống động, nơi cảnh quan, con người, văn hóa và lịch sử hòa quyện thành một cảm nhận thống nhất. Để suy luận và hành động, nó phải biểu diễn tổng thể ấy thành một mạng lưới gồm thực thể, thuộc tính, trạng thái và quan hệ. Thành phố không phải là một khối duy nhất mà là tập hợp của hàng nghìn địa điểm, doanh nghiệp, sự kiện, tuyến đường, dịch vụ, con người và hoạt động, tất cả đều được kết nối bằng những mối quan hệ có thể truy vấn và suy luận.

Ở góc độ này, địa phương trong nền kinh tế tác nhân gần với khái niệm song sinh số hơn là một trang web quảng bá. Tuy nhiên, đây không phải là "bản sao số" theo nghĩa mô phỏng hạ tầng vật lý hay cảm biến thời gian thực. Điều quan trọng hơn là một song sinh số ở cấp độ tri thức: một mô hình biểu diễn đầy đủ các thực thể, quan hệ và quy tắc vận hành của địa phương để máy có thể hiểu, suy luận và hành động. Đây là "bản đồ nhận thức" của máy, không phải "bản đồ địa lý" của con người.

Chính vì vậy, câu hỏi chiến lược cũng thay đổi. Trước đây người ta hỏi đã có trang web chưa, đã có ứng dụng chưa, đã

quảng bá đủ mạnh chưa. Trong nền kinh tế tác nhân, câu hỏi quan trọng hơn là: nếu một tác nhân phải tự vận hành trọn một chuyến đi mà không hỏi lại con người, nó đã có đủ mô hình dữ liệu để làm việc đó chưa?

Mục này vì thế không bàn về công nghệ triển khai hay các chuẩn kết nối cụ thể, vốn còn thay đổi theo từng thế hệ công nghệ. Điều bền hơn là logic biểu diễn địa phương. Một mô hình dữ liệu được thiết kế theo chuẩn mở, có cơ chế truy cập và xác thực phù hợp, sẽ giảm đáng kể chi phí để các thế hệ tác nhân khác nhau dùng lại cùng một nền tảng tri thức.

> ▪ ***Bốn lớp khái niệm cần phân biệt:***
> • *Mô hình dữ liệu là khái niệm chính, chỉ toàn bộ cách địa phương được biểu diễn cho máy.*
> • *Đồ thị tri thức là cấu trúc dùng để biểu diễn, gồm nút và cạnh có ngữ nghĩa.*
> • *Song sinh số là trạng thái của mô hình khi nó phản ánh được vận hành đang diễn ra chứ không chỉ danh mục tài sản.*
> • *Hệ điều hành chỉ là một ẩn dụ quản trị dùng để nói về khả năng điều phối, không phải một định nghĩa kỹ thuật và cũng không hàm ý địa phương có thể được lập trình toàn phần.*

### 7.2.1. Một địa phương không phải trang web mà là mô hình dữ liệu

Trong nhiều năm, chuyển đổi số của địa phương thường được hiểu là xây dựng trang web, cổng thông tin, ứng dụng di động hay các kênh truyền thông trực tuyến. Những công cụ này rất quan trọng đối với con người vì chúng giúp truyền tải thông tin và quảng bá hình ảnh. Tuy nhiên, trong nền kinh tế tác nhân,

chúng không còn là trung tâm của vấn đề. Một tác nhân không tiếp nhận trang web chủ yếu như một bề mặt thẩm mỹ. Nó quan tâm nhiều hơn đến cấu trúc thông tin, khả năng truy cập và khả năng hoàn tất hành động. Giao diện vẫn ảnh hưởng, đôi khi ảnh hưởng rất mạnh, như các kết quả đánh giá tác nhân duyệt web ở phần sau cho thấy. Nhưng thứ quyết định vẫn là bên dưới: một cấu trúc dữ liệu đủ đầy để máy đọc, suy luận và hành động.

Điều này đánh dấu sự thay đổi từ tư duy "xuất bản thông tin" sang tư duy "xây dựng mô hình". Trang web chỉ là một cách trình bày dữ liệu cho con người, trong khi mô hình dữ liệu là cách biểu diễn địa phương để máy hiểu. Hai trang web có thể giống hệt nhau về nội dung nhưng hoàn toàn khác nhau về khả năng được AI sử dụng nếu phía sau chúng là hai cấu trúc dữ liệu khác nhau.

Một địa phương vì thế không còn được nhìn như một tập hợp các trang giới thiệu mà như một hệ thống các thực thể được kết nối với nhau. Một điểm tham quan phải liên kết với lịch mở cửa, giá vé, tuyến giao thông, bãi đỗ xe, doanh nghiệp lữ hành, nhà hàng lân cận, sự kiện đang diễn ra và nhiều thực thể khác. Một lễ hội không chỉ là một bài viết giới thiệu mà còn là một đối tượng có thời gian, địa điểm, đơn vị tổ chức, khả năng đặt vé và mối liên hệ với các hoạt động phụ trợ. Chính những mối liên kết này mới tạo nên khả năng suy luận của máy.

Có thể hình dung sự khác biệt bằng một phép so sánh đơn giản. Trang web giống như một cuốn sách hướng dẫn du lịch, nơi người đọc phải tự tìm kiếm, tổng hợp và ra quyết định. Mô hình dữ liệu giống như bản đồ vận hành của cả địa phương, nơi mọi thành phần đều được định danh và kết nối để một tác nhân có thể tự dựng hành trình, đặt dịch vụ và tối ưu trải nghiệm với mức can thiệp thủ công tối thiểu.

Điều quan trọng là mô hình dữ liệu không thay thế trang web mà tồn tại ở tầng sâu hơn. Trang web, ứng dụng, chatbot hay trợ lý AI chỉ là các giao diện khác nhau cùng truy cập vào một mô hình tri thức chung. Nếu lớp dữ liệu bên dưới đầy đủ và nhất quán, địa phương có thể xuất hiện đồng thời trên nhiều nền tảng mà không phải xây dựng lại từ đầu. Ngược lại, nếu lớp dữ liệu rời rạc hoặc thiếu liên kết, việc đầu tư thêm giao diện sẽ chỉ làm tăng chi phí truyền thông chứ không cải thiện đáng kể khả năng được AI lựa chọn.

Vì vậy, tài sản số nền tảng của một địa phương trong nền kinh tế tác nhân không nằm ở riêng lượt truy cập trang web, mà ở chỗ mô hình dữ liệu có phản ánh đúng thực tế vận hành hay không. Trang web là nơi con người nhìn thấy địa phương. Mô hình dữ liệu mới là nơi máy thực sự hiểu địa phương.

### 7.2.2. Từ sơ đồ tổ chức sang sơ đồ tri thức

Các địa phương từ lâu đã quen quản lý bằng sơ đồ tổ chức. Mỗi sở, ban, ngành chịu trách nhiệm về một lĩnh vực riêng, mỗi doanh nghiệp, đơn vị sự nghiệp hay tổ chức xã hội hoạt động trong phạm vi chức năng của mình. Cách tổ chức này rất phù hợp với quản trị hành chính, nhưng lại không phản ánh cách một tác nhân AI nhìn thế giới.

Máy không quan tâm một nhà hàng thuộc sự quản lý của cơ quan nào, một lễ hội do sở nào tổ chức hay một tuyến xe buýt nằm trong phòng ban nào. Điều máy cần biết là nhà hàng đó phục vụ món gì, còn chỗ hay không, cách điểm tham quan bao xa, có phù hợp với sở thích của du khách hay không và có thể đặt bàn ngay lúc này không. Nói cách khác, AI không nhìn theo cấu trúc quản lý, mà nhìn theo cấu trúc quan hệ.

Trong mô hình tri thức của máy, mọi thực thể đều được kết nối bằng các quan hệ có ý nghĩa. Một khách sạn liên kết với các loại phòng, mức giá, đánh giá, bãi biển gần nhất, phương tiện di chuyển và các hoạt động giải trí xung quanh. Một bảo tàng kết nối với hiện vật, hướng dẫn viên, lịch sự kiện, trường học, tuyến xe công cộng và các quán cà phê lân cận. Những quan hệ này không chỉ mô tả hiện trạng mà còn cho phép AI suy luận ra các hành động tiếp theo.

Đó là lý do vì sao cùng một địa phương nhưng AI có thể tạo ra vô số hành trình khác nhau cho từng người dùng. Máy không đi theo sơ đồ hành chính mà đi theo mạng lưới tri thức. Khi một du khách yêu cầu "một ngày khám phá văn hóa địa phương phù hợp với gia đình có trẻ nhỏ", AI không truy vấn xem sở văn hóa đang quản lý những gì. Nó lần theo các quan hệ giữa di sản, khoảng cách, thời gian mở cửa, độ tuổi phù hợp, giao thông, nhà hàng và các hoạt động bổ trợ để xây dựng một kế hoạch hoàn chỉnh.

Tư duy này cũng làm thay đổi cách các địa phương xây dựng dữ liệu. Thay vì mỗi cơ quan sở hữu một cơ sở dữ liệu tách biệt, điều quan trọng hơn là mọi dữ liệu đều có khả năng kết nối với nhau thông qua các định danh chung và các quan hệ rõ ràng. Giá trị không còn nằm ở từng tập dữ liệu riêng lẻ mà nằm ở mạng lưới liên kết giữa chúng.

Có thể nói, sơ đồ tổ chức giúp con người biết ai chịu trách nhiệm, còn sơ đồ tri thức giúp máy biết điều gì liên quan đến điều gì. Trong nền kinh tế tác nhân, chính mạng lưới quan hệ ấy mới là nền tảng để AI hiểu địa phương, suy luận các phương án tối ưu và tự động thực hiện những hành động tạo ra giá trị kinh tế. Chuyển từ sơ đồ tổ chức sang sơ đồ tri thức vì thế không đơn thuần là thay đổi cách lưu trữ dữ liệu, mà là thay đổi cách địa phương được biểu diễn trong nhận thức của máy.

### 7.2.3. Địa phương như một hệ điều hành

Trong nền kinh tế tác nhân, cách hình dung phù hợp nhất về một địa phương không còn là một tập hợp các công trình hay các cơ quan hành chính, mà là một hệ điều hành. Cũng như hệ điều hành của một máy tính quản lý tài nguyên, điều phối tiến trình và cung cấp dịch vụ cho các ứng dụng, địa phương tương lai cũng cần được biểu diễn như một hệ thống vận hành thống nhất để các tác nhân AI có thể tương tác.

Trong mô hình này, con người là các chủ thể của hệ thống. Đó có thể là cư dân, du khách, doanh nghiệp, nghệ nhân, hướng dẫn viên, nhà đầu tư hay cán bộ quản lý. Mỗi nhóm có vai trò, quyền hạn, nhu cầu và hành vi khác nhau. Máy không chỉ cần biết họ là ai, mà còn cần hiểu họ có thể tương tác với những thực thể nào và trong những bối cảnh nào.

Tài sản là lớp tài nguyên của hệ điều hành địa phương. Đó không chỉ là di sản, cảnh quan, bảo tàng hay công viên, mà còn bao gồm tri thức, thương hiệu, làng nghề, sản phẩm địa phương, cộng đồng và các giá trị văn hóa sống. Đây là những nguồn lực mà hệ thống có thể huy động để tạo ra trải nghiệm và giá trị kinh tế.

Dịch vụ là lớp chức năng của hệ thống. Khách sạn, nhà hàng, vận chuyển, đặt vé, thanh toán, hướng dẫn viên, chăm sóc y tế, sự kiện, hỗ trợ khẩn cấp hay cấp phép đều là các dịch vụ mà AI có thể gọi đến để hoàn thành một nhiệm vụ. Một địa phương càng có nhiều dịch vụ được số hóa và kết nối, khả năng tự động hóa của các tác nhân càng cao.

Cuối cùng, luồng dữ liệu chính là "mạch máu" của toàn bộ hệ thống. Tình trạng giao thông, lượng khách, lịch sự kiện, thời

tiết, công suất khách sạn, giá vé, mức độ đông đúc hay phản hồi của người dùng đều liên tục thay đổi. Nếu không có các dòng dữ liệu này, AI chỉ nhìn thấy một bức ảnh tĩnh của địa phương thay vì một hệ thống đang vận động.

Điểm đáng chú ý là bốn thành phần trên không tồn tại độc lập mà liên tục tương tác với nhau. Con người sử dụng dịch vụ. Dịch vụ khai thác tài sản. Việc sử dụng tạo ra dữ liệu mới. Dữ liệu lại làm thay đổi cách AI phân bổ nguồn lực và đề xuất trải nghiệm. Địa phương vì thế là một hệ thống động chứ không phải một danh sách các điểm đến, nơi mọi thành phần liên tục tác động lẫn nhau.

Cách nhìn này cũng làm thay đổi tư duy quản trị. Thay vì tối ưu từng ngành riêng lẻ như du lịch, giao thông hay văn hóa, địa phương cần tối ưu toàn bộ "hệ điều hành". Khi mọi thành phần được kết nối trong cùng một mô hình dữ liệu, AI mới có thể phối hợp chúng thành những trải nghiệm liền mạch và tạo ra giá trị vượt xa tổng giá trị của từng bộ phận riêng lẻ.

Cần đặt một giới hạn cho ẩn dụ này. Hệ điều hành chỉ có giá trị ở cấp độ điều phối. Địa phương không phải một hệ thống khép kín có thể tối ưu hoàn toàn bằng dữ liệu. Con người, cộng đồng và văn hóa cần được thừa nhận quyền từ chối và quyền quyết định phần nào được biểu diễn, phần nào giữ mờ. Đây là một nguyên tắc chuẩn tắc mà cuốn sách đề xuất, chưa phải một quyền đã được pháp luật hiện hành định danh. Đọc ẩn dụ này theo nghĩa đen sẽ dẫn thẳng tới tư duy kỹ trị, tức là xem cư dân và di sản như tài nguyên có thể lập trình.

### 7.2.4. Vì sao dữ liệu phải phản ánh vận hành thật

Nhiều địa phương đã đầu tư xây dựng cơ sở dữ liệu số nhưng vẫn chưa trở thành những địa phương "khả hành động" đối với AI. Nguyên nhân không nằm ở số lượng dữ liệu mà ở bản chất của dữ liệu. Phần lớn các hệ thống hiện nay giống một danh mục tĩnh điện tử, tức là tập hợp các thông tin giới thiệu, trong khi điều AI cần lại là một mô hình phản ánh cách địa phương thực sự vận hành.

Một danh mục tĩnh trả lời câu hỏi "địa phương có những gì". Nó liệt kê danh lam thắng cảnh, khách sạn, nhà hàng, lễ hội hay sản phẩm đặc trưng. Đây là dữ liệu hữu ích cho việc quảng bá nhưng rất hạn chế đối với việc ra quyết định.

Ngược lại, một mô hình vận hành, gần với logic của song sinh số, trả lời câu hỏi "địa phương đang vận hành như thế nào". Nó không chỉ biết một khách sạn tồn tại mà còn biết khách sạn đó còn bao nhiêu phòng trống, thời gian nhận phòng, khả năng kết nối với phương tiện giao thông, mức độ đông đúc của khu vực xung quanh và tác động của thời tiết đến hành trình. Máy không cần một bức ảnh tĩnh mà cần một mô hình phản ánh trạng thái đang diễn ra.

Điều này đặc biệt quan trọng đối với các tác nhân AI. Một trợ lý du lịch không thể lập lịch trình tối ưu nếu dữ liệu về thời gian mở cửa đã lỗi thời. Một hệ thống điều phối giao thông không thể phân luồng nếu không biết tình trạng ùn tắc hiện tại. Một AI đặt dịch vụ cũng không thể hành động nếu không biết năng lực phục vụ thực tế của từng doanh nghiệp. Chất lượng quyết định luôn bị giới hạn bởi mức độ phản ánh đúng của mô hình dữ liệu.

Mô hình ấy vì thế không phải là một bản sao đẹp hơn của trang web mà là một biểu diễn động của hoạt động thực tế. Cần

nói thêm rằng có giá, có lịch và có công suất thì mới đạt tới một dạng song sinh số ở cấp độ vận hành, chưa phải bản sao đầy đủ của địa phương theo nghĩa mà kỹ thuật thường dùng cho hạ tầng vật lý. Giá trị của nó không nằm ở việc lưu trữ nhiều thông tin hơn, mà ở khả năng mô tả chính xác các thực thể, các quan hệ và trạng thái vận hành của địa phương theo thời gian.

Đây cũng là khác biệt căn bản giữa chuyển đổi số truyền thống và chuyển đổi sang nền kinh tế tác nhân. Chuyển đổi số trước đây chủ yếu số hóa tài liệu và quy trình. Trong nền kinh tế tác nhân, mục tiêu cao hơn là xây dựng một mô hình vận hành số của địa phương. Khi mô hình này đủ đầy và đủ chính xác, AI không chỉ đọc được địa phương mà còn có thể suy luận, dự báo và thực hiện hành động trên chính mô hình đó.

Nói cách khác, mô hình vận hành không phải là một cuốn danh mục tĩnh được số hóa. Nó là một mô hình sống về trạng thái vận hành của địa phương trong thế giới dữ liệu. Chỉ khi dữ liệu phản ánh đúng cách địa phương vận hành ngoài đời thực, các tác nhân AI mới có thể đưa ra những quyết định đúng trong thế giới thực.

Các nền tảng lớn đã xây dựng đồ thị tri thức rồi, địa phương không cần làm lại. Có phần đúng và đây là một trong hai chỗ chương này không bác bỏ được trọn vẹn. Các nền tảng thực sự đã bao phủ phần lớn thực thể phổ thông. Nhưng chủ thể nào định nghĩa lược đồ thì chủ thể đó quyết định điều gì được coi là có thật. Phần tri thức không khớp với lược đồ có sẵn, chẳng hạn tri thức nghề, tri thức cộng đồng hay các thực thể văn hóa không có mã phân loại quốc tế, sẽ bị lược bỏ chứ không bị ghi nhận sai. Phản bác chỉ được trả lời một nửa: địa phương không cần dựng lại thứ đã có, nhưng cần giữ quyền định nghĩa phần mà lược đồ chung bỏ sót. Với máy, điều không được biểu diễn thì gần như không tồn tại.

## 7.3. ĐỊA PHƯƠNG NHƯ MỘT ĐỒ THỊ

Nếu mục trước trả lời câu hỏi địa phương được biểu diễn bằng những thành phần nào, thì mục này đi sâu hơn vào câu hỏi quan trọng hơn: những thành phần ấy liên kết với nhau ra sao. Đối với một tác nhân AI, bản thân các thực thể chỉ là những điểm dữ liệu rời rạc. Giá trị thực sự nằm ở mạng lưới quan hệ kết nối chúng. Chính các quan hệ này mới tạo nên khả năng suy luận, lập kế hoạch và hành động.

Đây cũng là điểm khác biệt căn bản giữa một cơ sở dữ liệu truyền thống và một đồ thị tri thức (knowledge graph). Trong cơ sở dữ liệu, người ta chủ yếu lưu trữ các bản ghi: một khách sạn, một nhà hàng, một lễ hội hay một tuyến xe buýt. Trong đồ thị tri thức, điều quan trọng hơn là khách sạn ở gần địa điểm nào, nhà hàng phục vụ món gì, lễ hội diễn ra tại đâu, tuyến xe kết nối những khu vực nào và hướng dẫn viên thông thạo ngôn ngữ gì. Mỗi quan hệ đều mang ý nghĩa và có thể trở thành một bước trong chuỗi suy luận của AI.

Điều này có nghĩa là máy không nhìn địa phương như một danh mục các đối tượng độc lập mà như một mạng lưới liên kết. Khi người dùng yêu cầu "hãy lên kế hoạch cho một ngày khám phá làng nghề phù hợp với gia đình có trẻ nhỏ", AI không đi tìm một bản ghi mang tên "làng nghề". Nó lần theo các quan hệ giữa làng nghề, khoảng cách, thời gian di chuyển, độ tuổi phù hợp, nhà hàng, khu vui chơi, phương tiện giao thông và lịch hoạt động để tạo thành một hành trình hoàn chỉnh. Mỗi bước suy luận đều là một bước đi trên mạng lưới quan hệ đó.

Chính vì vậy, trong nền kinh tế tác nhân, quan hệ quan trọng không kém thực thể. Một địa phương có thể sở hữu hàng nghìn

tài sản giá trị nhưng nếu các tài sản ấy không được kết nối trong cùng một mô hình tri thức thì đối với AI, chúng vẫn chỉ là những "hòn đảo dữ liệu". Ngược lại, một mạng lưới có cấu trúc tốt cho phép máy phát hiện những kết hợp mà con người không dễ dàng nhận ra, từ đó tạo ra các trải nghiệm mới, tối ưu hóa hành trình và phân bổ nguồn lực hiệu quả hơn.

Mục này vì thế không đi theo cách phân loại quen thuộc như con người, địa điểm, doanh nghiệp hay sự kiện, mà đi theo logic của đồ thị. Các tiểu mục dưới đây lần lượt xem xét bốn câu hỏi: vì sao một thực thể chỉ có nghĩa khi nằm trong mạng lưới, các loại quan hệ nào tạo ra ngữ nghĩa, vì sao quan hệ thường trội hơn thuộc tính khi máy phải hành động và đồ thị đứt ở đâu. Khi ghép các quan hệ này lại, địa phương không còn là một tập hợp thông tin, mà trở thành một mạng lưới tri thức sống, nơi AI có thể di chuyển từ một thực thể đến bất kỳ thực thể nào thông qua những liên kết có ý nghĩa.

Có thể nói, nếu thực thể là các nút (nodes) thì quan hệ là các cạnh (edges). Các nút tạo nên "từ vựng" của địa phương, nhưng chính các cạnh mới tạo nên "ngữ pháp" để AI hiểu và vận hành trên địa phương. Trong nền kinh tế tác nhân, năng lực cạnh tranh không chỉ phụ thuộc vào việc địa phương có bao nhiêu dữ liệu, mà còn phụ thuộc vào việc mạng lưới quan hệ giữa các dữ liệu ấy phong phú, chính xác và phản ánh đúng thế giới thực đến mức nào.

### 7.3.1. Thực thể sống trong mạng lưới

Một thực thể không có nhiều ý nghĩa khi đứng riêng lẻ. Biết rằng một bảo tàng tồn tại, một nghệ nhân đang hoạt động hay một lễ hội được tổ chức hằng năm mới chỉ cung cấp những mảnh thông

tin rời rạc. Để máy hiểu được vai trò của chúng trong đời sống địa phương, mỗi thực thể phải được đặt vào một mạng lưới quan hệ rộng hơn.

Có thể gọi mạng lưới đó là đồ thị địa phương (place graph). Trong đồ thị này, địa điểm, con người, tổ chức, sự kiện, sản phẩm, dịch vụ và trải nghiệm trở thành các nút. Những mối liên hệ giữa chúng trở thành các cạnh. Một địa phương không còn được mô tả bằng danh sách những gì nó sở hữu, mà bằng cấu trúc cho thấy các thành phần ấy phụ thuộc, hỗ trợ và tạo giá trị cho nhau như thế nào.

Một ngôi đình chẳng hạn không chỉ là một điểm trên bản đồ. Nó có thể thuộc về một làng, thờ một nhân vật lịch sử, là nơi diễn ra một lễ hội, được một cộng đồng quản lý, liên quan đến một truyền thuyết và nằm trong một tuyến tham quan. Khi những quan hệ này được biểu diễn đầy đủ, ngôi đình không còn là một địa điểm đơn lẻ mà trở thành một nút trong mạng lưới văn hóa, xã hội, không gian và kinh tế của địa phương.

Tương tự, một nghệ nhân không chỉ là một hồ sơ cá nhân. Người đó có thể thuộc một làng nghề, thực hành một kỹ thuật truyền thống, đào tạo học viên, sản xuất một nhóm sản phẩm, tham gia lễ hội và tổ chức trải nghiệm cho du khách. Nếu dữ liệu chỉ ghi tên, tuổi và nghề nghiệp, máy biết nghệ nhân tồn tại nhưng không biết phải kết nối người đó với hành trình, sản phẩm hay trải nghiệm nào. Chính mạng lưới quan hệ mới biến một con người thành một thành phần có thể được tìm thấy và huy động trong quá trình tạo giá trị.

Đồ thị địa phương vì vậy không phải là bản đồ địa lý theo nghĩa thông thường. Bản đồ địa lý chủ yếu trả lời câu hỏi một thực thể

nằm ở đâu. Đồ thị địa phương trả lời một tập hợp câu hỏi rộng hơn: thực thể đó thuộc về cộng đồng nào, phục vụ mục đích gì, liên quan đến tài sản nào, được vận hành bởi ai, xuất hiện vào thời gian nào và có thể kết hợp với những thực thể nào khác.

Đối với AI, đây là sự khác biệt giữa việc nhận biết một địa điểm và hiểu một địa phương. Máy có thể nhận biết hàng nghìn điểm quan tâm từ tọa độ và tên gọi. Nhưng chỉ khi các điểm ấy được đặt trong một đồ thị địa phương, máy mới có thể suy luận rằng một chợ sáng phù hợp với trải nghiệm ẩm thực, một nghệ nhân có thể kết nối với lớp học thủ công, hay một lễ hội có thể tạo thêm nhu cầu cho lưu trú, vận tải và thương mại địa phương.

Một đồ thị địa phương mạnh không nhất thiết là mạng lưới có nhiều nút nhất. Điều quan trọng hơn là các nút có được kết nối đúng hay không. Một địa phương có thể sở hữu hàng nghìn thực thể nhưng vẫn khó được máy hiểu nếu phần lớn chúng đứng cô lập. Ngược lại, một địa phương nhỏ với ít tài sản hơn vẫn có thể trở nên rất rõ ràng trong nhận thức của AI nếu các quan hệ giữa tài sản, con người, dịch vụ và trải nghiệm được mô tả đầy đủ.

Bởi vậy, đơn vị căn bản của địa phương trong thế giới dữ liệu không phải là thực thể riêng lẻ mà là thực thể trong quan hệ. Không có mạng lưới, thực thể chỉ là một bản ghi. Trong mạng lưới, nó trở thành một phần của câu chuyện, một mắt xích trong vận hành và một điểm xuất phát cho hành động của máy.

### 7.3.2. Quan hệ tạo ra ngữ nghĩa

Tên gọi và thuộc tính giúp máy nhận diện một thực thể, nhưng chính quan hệ mới giúp máy hiểu thực thể đó có ý nghĩa gì. Một

tòa nhà có thể là nhà ở, khách sạn, bảo tàng, trường học hoặc nơi tổ chức sự kiện. Hình dạng vật lý không đủ để xác định vai trò của nó. Ý nghĩa xuất hiện khi thực thể được đặt trong quan hệ với các thực thể khác.

Trong đồ thị tri thức, quan hệ không phải là đường nối trang trí giữa hai nút. Mỗi loại quan hệ là một phát biểu có ý nghĩa về thế giới. Khi hệ thống biết một nhà hàng ở gần một bảo tàng, một phòng trưng bày nằm bên trong một trung tâm văn hóa, một nghệ nhân thuộc về một làng nghề, một lễ hội được đăng cai bởi một địa phương hay một di tích được vận hành bởi một tổ chức, máy bắt đầu hình thành được một mô hình có thể suy luận.

Quan hệ ở gần (near) tạo ra ngữ nghĩa không gian và cho biết những thực thể nào có thể ghép chung một hành trình. Nhưng gần không đồng nghĩa với khoảng cách nhỏ. Hai địa điểm cách nhau một kilômét có thể rất gần nếu có đường đi bộ thuận tiện và rất xa nếu bị chia cắt bởi sông, đường cao tốc hay địa hình khó tiếp cận. Quan hệ này phải phản ánh khả năng di chuyển thực tế chứ không chỉ tọa độ.

Quan hệ nằm bên trong (inside) tạo ra cấu trúc bao chứa: một cửa hàng trong chợ, một phòng triển lãm trong bảo tàng, một di tích trong quần thể, một làng trong vùng di sản. Nó giúp máy hiểu rằng thực thể nhỏ hơn chịu ràng buộc bởi quy tắc, giờ hoạt động và điều kiện tiếp cận của không gian lớn hơn. Nếu khu phức hợp đóng cửa, mọi thực thể bên trong cũng không dùng được, dù dữ liệu riêng của chúng vẫn ghi là đang hoạt động.

Quan hệ thuộc về (belong to) thể hiện tư cách thành viên, nguồn gốc hoặc bản sắc: nghệ nhân thuộc một cộng đồng nghề, món ăn thuộc một truyền thống ẩm thực, doanh nghiệp thuộc

một hiệp hội. Đây là một trong những loại quan hệ quan trọng nhất đối với bản sắc, bên cạnh các quan hệ như bắt nguồn từ, được thực hành bởi, được truyền lại bởi hay được cộng đồng công nhận. Nó giúp máy phân biệt một sản phẩm chỉ được bán tại địa phương với một sản phẩm thực sự bắt nguồn từ địa phương.

Quan hệ được đăng cai bởi (hosted by) mô tả chủ thể tiếp nhận và tổ chức một hoạt động, nối sự kiện với địa điểm, năng lực tổ chức và các dịch vụ phụ trợ. Nó cũng cho phép máy suy luận những biến động tạm thời như lượng khách tăng, giao thông thay đổi hay nhu cầu lưu trú phát sinh.

Quan hệ được vận hành bởi (operated by) xác định chủ thể chịu trách nhiệm cho hoạt động thực tế, tách khỏi quyền sở hữu. Một bảo tàng có thể thuộc sở hữu nhà nước nhưng do một đơn vị sự nghiệp vận hành. Một tuyến xe có thể do địa phương quy hoạch nhưng do doanh nghiệp khai thác. Phân biệt hai quyền này là cần thiết, vì tác nhân phải biết ai có thẩm quyền xác nhận thông tin, cung cấp dịch vụ và chịu trách nhiệm khi có thay đổi.

Mỗi quan hệ trong dữ liệu chỉ mang một lớp thông tin riêng. Tuy nhiên, khi nhiều quan hệ được liên kết với nhau, máy không còn chỉ trả lời từng câu hỏi đơn lẻ mà có thể thực hiện những suy luận nhiều bước. Ví dụ, AI biết một nghệ nhân thuộc về một làng nghề, làng nghề đó nằm trong một khu vực di sản, khu vực này ở gần một bến xe và bến xe được vận hành bởi một doanh nghiệp vận tải. Từ chuỗi quan hệ ấy, hệ thống không chỉ giới thiệu một nghệ nhân, mà còn có thể đề xuất cách di chuyển, sắp xếp hành trình và giải thích bối cảnh văn hóa của điểm đến. Đó chính là sức mạnh của đồ thị tri thức, máy không suy nghĩ bằng các mẩu thông tin rời rạc mà bằng mạng lưới các thực thể và quan hệ, đi từ nút này sang nút khác để ghép thành một câu trả lời hoặc một hành động hoàn chỉnh.

Ba thuộc tính sau đây quyết định chất lượng của một quan hệ. Thứ nhất là loại, mỗi quan hệ phải thuộc một loại xác định, vì máy suy luận theo loại chứ không theo đường nối chung chung. Thứ hai là hướng, lễ hội được đăng cai bởi thành phố và thành phố đăng cai lễ hội mô tả cùng một liên kết nhưng phục vụ hai hướng truy vấn khác nhau, nên máy phải nhận biết được hướng của quan hệ và, khi cần, suy ra quan hệ nghịch đảo tương ứng. Thứ ba là hiệu lực theo thời gian, một địa điểm gần tuyến giao thông trong điều kiện bình thường có thể khó tiếp cận khi đường bị đóng, một sự kiện chỉ được đăng cai trong một khoảng thời gian, một đơn vị vận hành có thể thay đổi sau khi hợp đồng kết thúc.

Điều này cho thấy xây dựng đồ thị địa phương không đơn thuần là nối thật nhiều thực thể với nhau. Các quan hệ phải chính xác, có loại hình rõ ràng, có hướng, có thời hạn và phản ánh đúng thực tế. Nếu quan hệ sai, máy không chỉ hiểu sai một thực thể mà còn có thể suy luận sai cả chuỗi. Trong ngôn ngữ, từ ngữ chỉ tạo ra ý nghĩa đầy đủ khi được đặt trong câu. Trong đồ thị địa phương, thực thể cũng chỉ có ý nghĩa đầy đủ khi được đặt trong quan hệ. Thực thể là từ vựng, quan hệ là ngữ pháp. Nhờ ngữ pháp ấy, máy không chỉ biết địa phương có gì mà còn hiểu địa phương vận hành như thế nào.

### 7.3.3. Khi hành động, quan hệ thường quan trọng hơn thuộc tính

Trong các hệ thống dữ liệu truyền thống, thực thể thường được mô tả bằng thuộc tính. Một khách sạn có tên, địa chỉ, số phòng, hạng sao, mức giá và danh sách tiện nghi. Một bảo tàng có năm thành lập, diện tích, thời gian mở cửa và số lượng hiện vật. Những thuộc tính này giúp máy nhận diện và phân loại đối

tượng, nhưng chưa đủ để xác định giá trị của đối tượng trong một tình huống cụ thể.

Trong nhiều nhiệm vụ hành động, giá trị không nằm bên trong thực thể mà nằm trong quan hệ giữa thực thể với nhu cầu, thời gian và những thực thể khác. Một khách sạn năm sao có thể rất hấp dẫn khi được xem riêng lẻ. Nhưng đối với một hành khách phải lên chuyến tàu lúc sáu giờ sáng, khách sạn hai sao nằm cạnh ga có thể là lựa chọn tốt hơn. Hạng sao là thuộc tính của khách sạn. Gần ga là quan hệ giữa khách sạn và một hạ tầng di chuyển. Khi AI phải hành động, quan hệ ấy có thể quan trọng hơn toàn bộ danh sách tiện nghi.

Điều này không có nghĩa thuộc tính mất giá trị. Thuộc tính cho biết thực thể là gì và có gì, còn quan hệ cho biết thực thể làm được gì trong một mạng lưới cụ thể. Vì thế một địa phương không thể nâng chất lượng dữ liệu chỉ bằng cách thêm trường mô tả. Thêm số điện thoại, ảnh, mức giá hay danh mục tiện nghi làm hồ sơ đầy đủ hơn nhưng chưa chắc làm thực thể hữu dụng hơn. Điều máy cần biết là thực thể ấy nối với ai, nằm trong hệ thống nào, tiếp cận bằng cách nào và kết hợp được với những lựa chọn nào khác.

Một bãi biển đẹp là thuộc tính hấp dẫn. Nhưng với một gia đình có trẻ nhỏ, ý nghĩa thực tế của nó phụ thuộc vào các quan hệ: có gần nơi lưu trú không, có nằm trong vùng nước an toàn không, có dịch vụ cứu hộ không, có nhà vệ sinh và chỗ ăn uống lân cận không. Tương tự, một làng nghề ba trăm năm tuổi giàu giá trị văn hóa, nhưng để biến giá trị ấy thành trải nghiệm, máy phải biết làng nghề gắn với nghệ nhân nào, có hoạt động nào cho khách, diễn ra khi nào, đặt trước qua đâu và đi tới bằng phương tiện gì. Tuổi đời là thuộc tính. Khả năng tham gia là kết quả của mạng lưới.

Sức mạnh lớn nhất của quan hệ không nằm ở việc lưu trữ thông tin, mà ở khả năng tạo ra những suy luận vượt ra ngoài dữ liệu được ghi trực tiếp. Khi AI biết một khách sạn ở gần nhà ga, nhà ga kết nối với sân bay và sân bay phục vụ nhiều chuyến bay quốc tế vào sáng sớm, hệ thống có thể suy luận rằng khách sạn đó là lựa chọn phù hợp cho hành khách quá cảnh hoặc những người cần khởi hành sớm. Không cần tồn tại một trường dữ liệu ghi rõ *"phù hợp với khách đi chuyến bay sáng"*. Kết luận này được hình thành từ việc máy lần theo nhiều quan hệ liên tiếp trong đồ thị tri thức. Đó là điểm khác biệt căn bản giữa đồ thị tri thức và danh mục dữ liệu tĩnh. Danh mục tĩnh chỉ lưu những gì đã được mô tả sẵn, còn đồ thị cho phép máy kết hợp các quan hệ để tạo ra tri thức mới. Nói cách khác, danh mục trả lời câu hỏi *"điểm đến này là gì"*, còn đồ thị giúp AI trả lời câu hỏi *"điểm đến này phù hợp với ai, trong hoàn cảnh nào và vì sao"*. Chính khả năng suy luận theo ngữ cảnh này là nền tảng để các tác nhân AI xây dựng hành trình, đưa ra đề xuất và thực hiện những quyết định mà không cần mọi kết luận đều được con người viết sẵn.

Quan hệ cũng làm giá trị của một thực thể thay đổi theo bối cảnh. Cùng một khách sạn có thể rất hợp với khách công vụ vì gần trung tâm hội nghị, không hợp với khách nghỉ dưỡng vì xa biển, nhưng lại hợp với người cao tuổi vì cạnh bệnh viện. Thực thể không đổi, thuộc tính cũng không đổi. Chỉ mục tiêu và quan hệ được kích hoạt là khác. Nguyên tắc thiết kế dữ liệu rút ra từ đó: không chỉ hỏi thực thể này có những thuộc tính gì, mà hỏi thêm nó có quan hệ gì với những thực thể mà con người và máy cần dùng.

Cần giữ mệnh đề này trong đúng phạm vi của nó. Với một số nhu cầu, thuộc tính có tính loại trừ tuyệt đối và không quan hệ nào bù được: khả năng tiếp cận bằng xe lăn, thành phần gây dị

ứng, giới hạn độ tuổi, yêu cầu thị thực, tiêu chuẩn an toàn cháy nổ. Trong những trường hợp ấy, thuộc tính là điều kiện cần và máy phải loại trừ trước khi xét đến bất kỳ quan hệ nào. Điều chương này khẳng định hẹp hơn: khi các điều kiện loại trừ đã được thỏa mãn, việc bổ sung thêm một quan hệ đúng thường làm tăng khả năng được chọn nhiều hơn việc bổ sung thêm một trường mô tả.

Bởi vậy, một thực thể giàu thuộc tính nhưng nghèo quan hệ vẫn có thể bị cô lập trong đồ thị. Ngược lại, một thực thể bình thường nhưng được kết nối đúng có thể giữ vai trò trung tâm trong hàng nghìn hành trình. Thuộc tính mô tả giá trị nội tại, còn quan hệ kích hoạt giá trị sử dụng. Khi máy chuyển từ đọc sang hành động, giá trị sử dụng thường là thứ có tiếng nói cuối cùng.

Quan hệ trội hơn thuộc tính. Khi máy phải hành động chứ không chỉ mô tả, giá trị sử dụng của một thực thể do vị trí của nó trong mạng lưới quyết định, không do độ dài danh sách thuộc tính. Thêm một trường mô tả làm hồ sơ đầy đủ hơn. Thêm một quan hệ đúng làm hồ sơ khả dụng hơn. Thuộc tính mô tả giá trị nội tại. Quan hệ kích hoạt giá trị sử dụng.

### 7.3.4. Điểm đứt của đồ thị

Một đồ thị tri thức không thất bại chỉ khi dữ liệu sai. Nó còn thất bại khi những liên kết cần thiết không tồn tại, khi một thực thể bị nhận diện sai hoặc khi cùng một đối tượng bị chia thành nhiều bản ghi không thể nối lại với nhau. Đây là những điểm đứt của đồ thị: nơi chuỗi suy luận của máy bị gián đoạn dù các dữ liệu riêng lẻ vẫn có vẻ chính xác.

Thực thể vỡ là khi một đối tượng ngoài đời thật bị tách thành nhiều định danh không nối lại được với nhau. Thực thể sai là khi định danh đúng nhưng thuộc tính gắn với nó sai. Thực thể gộp sai là khi nhiều đối tượng khác nhau bị nhập thành một định danh duy nhất. Ba dạng này đòi hỏi ba cách xử lý khác nhau: hợp nhất hồ sơ, sửa dữ liệu và tách hồ sơ.

Điểm đứt thứ nhất là liên kết bị thiếu (missing link). Một bảo tàng có thể được định danh đúng, thời gian mở cửa chính xác và địa chỉ đầy đủ. Một bãi đỗ xe gần đó cũng có dữ liệu riêng. Nhưng nếu không tồn tại quan hệ cho biết bãi đỗ xe phục vụ bảo tàng, AI không thể chắc chắn rằng người dùng có thể kết hợp hai thực thể này. Đối với con người địa phương, mối liên hệ có thể hiển nhiên. Đối với máy, điều không được biểu diễn gần như không tồn tại.

Một liên kết bị thiếu thường nhỏ đến mức khó nhận thấy. Trang của lễ hội ghi rõ địa điểm tổ chức, còn trang của địa điểm không nhắc đến lễ hội. Dữ liệu giao thông xác nhận một tuyến xe đi qua khu di sản nhưng không có quan hệ nối điểm dừng với cổng vào. Danh sách nghệ nhân đầy đủ nhưng không ai được liên kết với sản phẩm hay làng nghề mà họ đại diện. Mỗi kho dữ liệu đều đúng, song toàn bộ hành trình vẫn không hình thành. Đây là nghịch lý của dữ liệu phân mảnh: địa phương sở hữu mọi mảnh ghép mà máy vẫn không ghép được thành bức tranh.

Một đồ thị tri thức chỉ mạnh khi các quan hệ giữa các thực thể được khai báo đầy đủ. Chỉ cần thiếu một mắt xích, toàn bộ chuỗi suy luận của AI có thể bị đứt đoạn. Ví dụ, AI biết một du khách muốn đến làng nghề, đồng thời cũng biết vị trí của các điểm đón xe buýt và mạng lưới giao thông công cộng. Nhưng nếu không tồn tại quan hệ cho biết điểm đón nào thực sự phục vụ làng nghề, hệ thống sẽ không thể nối các thông tin này thành

một hành trình hoàn chỉnh. Khi đó, AI chỉ còn ba lựa chọn: đưa ra một gợi ý với độ tin cậy thấp, yêu cầu người dùng tự xác minh, hoặc loại bỏ làng nghề khỏi danh sách đề xuất. Dù theo cách nào, trải nghiệm của người dùng đều suy giảm và cơ hội của địa phương cũng bị mất đi. Điều đáng chú ý là nguyên nhân không nằm ở việc thiếu xe buýt hay thiếu điểm đến. Dịch vụ vẫn tồn tại, tài sản vẫn hiện hữu, nhưng mối quan hệ giữa chúng không được biểu diễn trong dữ liệu. Đối với máy, một quan hệ chưa được khai báo gần như tương đương với một quan hệ không tồn tại. Vì vậy, trong nền kinh tế tác nhân, kết nối dữ liệu nhiều khi quan trọng không kém bản thân dữ liệu. Chỉ khi các thực thể được liên kết đúng, AI mới có thể suy luận, lập kế hoạch và hoàn thành một hành trình trọn vẹn.

Điểm đứt thứ hai là thực thể bị vỡ (broken entity). Hiện tượng này xảy ra khi một đối tượng ngoài đời thực không được nhận diện như một thực thể thống nhất trong thế giới dữ liệu. Một khách sạn có thể xuất hiện dưới tên tiếng Việt trên trang web địa phương, tên tiếng Anh trên nền tảng đặt phòng, tên viết tắt trên bản đồ và tên pháp nhân trong hệ thống đăng ký kinh doanh. Nếu máy không xác định được đó là cùng một đối tượng, dữ liệu sẽ bị chia nhỏ thành nhiều hồ sơ.

Khi thực thể bị vỡ, mỗi hồ sơ chỉ giữ một phần sự thật. Một bản ghi có địa chỉ, một bản ghi có đánh giá, một bản ghi có khả năng đặt phòng và một bản ghi có thông tin pháp lý. Không hồ sơ nào đủ hoàn chỉnh để tạo ra độ tin cậy cao. Tài sản ngoài đời thực chỉ có một, nhưng trong nhận thức của máy nó bị phân tán thành nhiều thực thể yếu.

Chiều ngược lại cũng nguy hiểm: nhiều đối tượng khác nhau có thể bị gộp sai thành một thực thể. Hai nhà hàng cùng tên tại

hai địa điểm khác nhau có thể bị nhập làm một. Một lễ hội truyền thống có thể bị nhầm với sự kiện thương mại sử dụng tên tương tự. Một di tích và khu du lịch xây dựng xung quanh nó có thể bị coi là cùng một địa điểm. Khi đó, đánh giá, tọa độ, thời gian hoạt động và thông tin giao dịch bị trộn lẫn.

Sai lầm trong phân giải thực thể không dừng lại ở một bản ghi riêng lẻ. Vì mỗi thực thể là một nút trong đồ thị, một nút bị tách vỡ hoặc gộp nhầm sẽ kéo theo hàng loạt quan hệ sai, nhà hàng có thể bị nối với sai khách sạn, nghệ nhân bị gắn với sai làng nghề, còn sự kiện bị đặt tại sai địa điểm. Từ đó hình thành một chuỗi lan truyền, thực thể vỡ dẫn đến quan hệ vỡ, quan hệ vỡ dẫn đến suy luận vỡ và suy luận vỡ cuối cùng dẫn đến hành động vỡ. Mỗi lỗi ở tầng nhận diện làm sai các cạnh gắn với thực thể, còn mỗi cạnh sai lại khiến một nhánh suy luận chệch hướng. Trong nền kinh tế nội dung, hậu quả có thể chỉ là một câu trả lời sai. Nhưng trong nền kinh tế tác nhân, hậu quả có thể là một hành trình sai, một yêu cầu gửi nhầm hoặc một giao dịch được thực hiện với sai chủ thể. Một tác nhân AI có thể đề xuất sai nơi, dẫn người dùng đến sai địa chỉ, đặt nhầm dịch vụ hoặc không thể hoàn tất giao dịch. Trong nền kinh tế nội dung, lỗi như vậy chủ yếu gây bất tiện. Trong nền kinh tế tác nhân, nó có thể tạo ra chi phí thực, tranh chấp và mất niềm tin.

Điểm đứt của đồ thị còn xuất hiện khi quan hệ đã từng đúng nhưng không còn đúng. Một doanh nghiệp ngừng vận hành địa điểm nhưng vẫn được ghi là đơn vị quản lý. Một tuyến xe đã thay đổi lộ trình nhưng quan hệ cũ chưa bị xóa. Một lễ hội chuyển địa điểm nhưng các nguồn khác vẫn giữ thông tin trước đây. Đồ thị không chỉ cần được xây dựng. Nó phải được duy trì như một hệ thống sống.

Vì vậy, chất lượng của đồ thị địa phương không nên đo bằng số lượng thực thể hay số lượng liên kết. Chỉ số quan trọng hơn là tính liên tục của những đường đi mà máy cần dùng: từ nhu cầu của người dùng đến tài sản, từ tài sản đến dịch vụ, từ dịch vụ đến giao dịch, có tồn tại một chuỗi liên kết đầy đủ và đáng tin hay không. Có thể hình dung như mạng lưới đường sá: số con đường không quyết định khả năng lưu thông, chỉ một cây cầu đứt cũng đủ cô lập cả một khu vực.

Bởi vậy, nhiệm vụ quan trọng của địa phương không chỉ là bổ sung dữ liệu mới mà còn là liên tục tìm kiếm các điểm đứt: thực thể nào đang bị cô lập, quan hệ nào còn thiếu, hồ sơ nào bị trùng lặp và đường suy luận nào không thể đi đến hành động. Một đồ thị mạnh không phải là đồ thị không có lỗi, mà là đồ thị có khả năng nhận ra và sửa các điểm đứt trước khi chúng trở thành quyết định sai của máy.

Có bốn dạng lỗi phổ biến làm suy giảm chất lượng của đồ thị tri thức và mỗi dạng đều ảnh hưởng trực tiếp đến khả năng suy luận của tác nhân AI:

- ***Liên kết thiếu*** xảy ra khi hai thực thể đều đúng nhưng không có quan hệ kết nối. Có thể phát hiện bằng cách tìm những nút có rất ít hoặc không có liên kết so với các thực thể cùng loại. Khi gặp dạng lỗi này, AI thường bỏ điểm đến khỏi hành trình vì không đủ căn cứ để kết nối với các bước tiếp theo.

- ***Thực thể vỡ*** là trường hợp một đối tượng ngoài đời thực bị tách thành nhiều hồ sơ khác nhau. Việc đối sánh tên, tọa độ, mã số thuế hoặc số điện thoại giữa các nguồn dữ liệu giúp phát hiện lỗi này. Hệ quả là mỗi hồ sơ chỉ chứa một phần thông tin, khiến không hồ sơ nào đủ đầy để máy tin tưởng.

- ***Thực thể gộp sai*** xảy ra khi nhiều đối tượng khác nhau bị nhập thành một thực thể. Dấu hiệu thường thấy là tọa độ phân tán bất thường hoặc các đánh giá, giờ mở cửa và thông tin giao dịch mâu thuẫn với nhau. Khi đó, AI sẽ trộn lẫn dữ liệu của nhiều nơi khác nhau và đưa ra các đề xuất sai.

- ***Quan hệ hết hạn*** là những liên kết từng đúng nhưng không còn phản ánh thực tế, chẳng hạn đơn vị vận hành, tuyến kết nối hay cơ quan quản lý đã thay đổi. Cách khắc phục là gắn thời hạn hiệu lực cho mọi quan hệ và yêu cầu xác nhận định kỳ. Nếu không được cập nhật, AI có thể tiếp tục dẫn khách đến những đơn vị đã ngừng hoạt động hoặc không còn chịu trách nhiệm cung cấp dịch vụ.

### 7.3.5. Khi cả địa phương trở thành đồ thị tri thức

Một trong những ngộ nhận phổ biến nhất khi xây dựng dữ liệu địa phương là cho rằng chỉ cần thu thập thật nhiều thông tin thì AI sẽ hiểu tốt hơn. Trong lập luận của cuốn sách này, điều quyết định năng lực suy luận của tác nhân không phải số lượng dữ liệu mà là số lượng và chất lượng của các mối quan hệ giữa dữ liệu.

Con người có thể tự suy luận từ những mảnh thông tin rời rạc nhờ kinh nghiệm và trực giác. Máy thì không. Một tác nhân chỉ có thể suy luận khi tồn tại những đường kết nối rõ ràng giữa các thực thể. Nếu các nút trong đồ thị đứng riêng lẻ, AI chỉ nhìn thấy một tập hợp đối tượng. Chỉ khi các nút được liên kết, AI mới nhìn thấy một hệ thống.

Giả sử hệ thống biết rằng một khách sạn nằm gần ga tàu, ga tàu kết nối với tuyến metro, tuyến metro đi tới trung tâm hội

nghị, còn trung tâm hội nghị đang diễn ra một sự kiện quốc tế. Không cần bất kỳ dòng mô tả nào nói rằng khách sạn này "phù hợp cho khách công tác", tác nhân vẫn có thể tự suy luận điều đó từ chuỗi quan hệ đã tồn tại trong đồ thị.

Đây chính là sức mạnh của đồ thị tri thức. Giá trị không nằm ở từng thực thể riêng lẻ mà nằm ở mạng lưới kết nối giữa chúng. Mỗi liên kết mới đều mở ra thêm những con đường suy luận mà trước đó không tồn tại.

Trong khoa học mạng lưới, số lượng quan hệ thường tăng nhanh hơn rất nhiều so với số lượng nút. Khi địa phương bổ sung thêm một điểm tham quan, giá trị không chỉ đến từ bản thân điểm đó mà còn đến từ việc nó tạo thêm hàng chục quan hệ mới với khách sạn, nhà hàng, tuyến xe buýt, lễ hội, tuyến tham quan, khu dân cư, doanh nghiệp và các dịch vụ công. Đồ thị vì thế trở nên giàu ngữ nghĩa theo cấp số nhân chứ không chỉ tăng tuyến tính.

Điều này cũng giải thích vì sao hai địa phương có cùng số lượng dữ liệu nhưng năng lực được AI khai thác lại rất khác nhau. Một nơi sở hữu hàng trăm nghìn bản ghi nhưng gần như không có liên kết sẽ chỉ là một kho dữ liệu lớn. Trong khi đó, một nơi có ít dữ liệu hơn nhưng được tổ chức thành mạng lưới quan hệ đầy đủ sẽ cho phép tác nhân lập kế hoạch, giải thích, đề xuất và phối hợp hành động hiệu quả hơn nhiều.

Trong mô hình đó, mọi thành phần đều trở thành một nút của đồ thị: con người, doanh nghiệp, địa danh, tài sản văn hóa, hạ tầng, sự kiện, phương tiện giao thông, dịch vụ công, sản phẩm du lịch, nguồn dữ liệu và cả các tác nhân AI. Chúng không tồn tại độc lập mà được kết nối bằng hàng triệu quan hệ thể hiện vị trí,

quyền sở hữu, chức năng, lịch hoạt động, khả năng phục vụ và ảnh hưởng lẫn nhau.

Khi địa phương được biểu diễn dưới dạng đồ thị tri thức, tác nhân không còn phải đọc từng trang web, từng tập tài liệu hay từng cơ sở dữ liệu riêng biệt. Chúng có thể truy cập trực tiếp vào mạng lưới tri thức, lần theo các quan hệ, kiểm chứng thông tin và tự xây dựng chuỗi suy luận để hoàn thành nhiệm vụ.

Sự chuyển đổi này tạo ra một bước nhảy căn bản. Địa phương không còn là một tập hợp thông tin để con người tìm kiếm, mà trở thành một hệ thống tri thức mà máy có thể hiểu, phân tích và hành động trên đó. Đây chính là nền tảng để tác nhân lập kế hoạch hành trình, điều phối giao thông, đặt dịch vụ, xử lý sự cố, hỗ trợ quản trị đô thị hay tối ưu hóa trải nghiệm của từng cá nhân trong thời gian thực.

Ở góc nhìn rộng hơn, đồ thị tri thức cũng là lớp kết nối giữa thế giới vật lý và thế giới số. Mỗi thay đổi ngoài đời thực, từ một tuyến đường mới, một khách sạn đóng cửa, một lễ hội được tổ chức đến một dịch vụ mới xuất hiện, đều được phản ánh vào đồ thị và ngay lập tức trở thành tri thức có thể khai thác. Song sinh số, hệ thống điều hành thông minh và nền kinh tế tác nhân đều dựa trên chính lớp tri thức này để hoạt động.

Vì vậy, mục tiêu cuối cùng của thiết kế dữ liệu địa phương không phải là xây dựng thêm một cổng thông tin hay một cơ sở dữ liệu lớn hơn. Mục tiêu là kiến tạo một đồ thị tri thức sống, nơi mọi thực thể được định danh, mọi quan hệ được biểu diễn và mọi thay đổi của địa phương được cập nhật liên tục. Chỉ khi đó, địa phương mới thực sự trở thành một thực thể có thể được máy hiểu, suy luận và hợp tác. Đây cũng là nền móng cho mục tiếp

theo, nơi câu hỏi không còn là địa phương được mô tả như thế nào, mà là địa phương có gì để tác nhân thực sự giao dịch.

Chuẩn hóa dữ liệu sẽ bóp chết bản sắc. Lo ngại này có cơ sở lịch sử vững chắc và không nên bị gạt đi. Mọi hệ thống phân loại đều làm nghèo đối tượng được phân loại. Tuy nhiên, phản bác nhầm đối tượng: cái chết của bản sắc không đến từ việc mô tả một lễ hội bằng lược đồ, mà đến từ việc mô tả sai hoặc để người khác mô tả thay. Chuẩn hóa là điều kiện để cộng đồng có thể quyết định phần nào công bố, chứ không phải nghĩa vụ công bố tất cả. Không đưa vào lược đồ phần tri thức mà cộng đồng chủ thể đã từ chối công bố. Quyền mờ phải được thiết kế vào hệ thống từ đầu, không xử lý bằng ngoại lệ về sau.

## 7.4. TỒN KHO CÓ THỂ GIAO DỊCH

Nếu mục 7.1 đặt ra yêu cầu địa phương phải được giao việc, mục 7.2 mô tả địa phương như một mô hình dữ liệu và mục 7.3 giải thích cách các thực thể nối thành đồ thị, thì mục này trả lời câu hỏi quan trọng nhất ở cấp độ kinh tế: máy giao dịch cái gì?

Trong ngành du lịch truyền thống, "tồn kho" thường được hiểu là số lượng phòng khách sạn, ghế máy bay hay vé tham quan còn trống. Đó là cách nhìn của con người và của các hệ thống đặt chỗ thế hệ trước. Nhưng trong nền kinh tế tác nhân, khái niệm tồn kho mở rộng đáng kể. Đối với tác nhân AI, bất kỳ tài nguyên nào có thể được mô tả, định danh, kiểm tra trạng thái và thực hiện giao dịch đều có thể trở thành một dạng tồn kho.

Một địa phương vì thế không còn chỉ sở hữu tài nguyên, mà sở hữu một danh mục tài sản có khả năng giao dịch (tradable inventory). Mỗi phòng nghỉ, mỗi bàn ăn, mỗi suất biểu diễn, mỗi khung giờ khám bệnh, mỗi chỗ đỗ xe, mỗi vé tham quan, mỗi hướng dẫn viên, mỗi xe điện, mỗi lớp học trải nghiệm, thậm chí mỗi khoảng thời gian còn trống của một nghệ nhân hay chuyên gia đều có thể trở thành một đơn vị giao dịch nếu chúng được số hóa đúng cách.

Đây là khác biệt căn bản giữa số hóa thông tin và số hóa khả năng giao dịch. Một trang web có thể giới thiệu rằng địa phương có một làng nghề nổi tiếng, nhưng tác nhân không thể đặt một buổi trải nghiệm nếu không tồn tại một thực thể đại diện cho buổi trải nghiệm đó cùng các thuộc tính như thời gian, số chỗ còn lại, giá, điều kiện tham gia và API giao dịch. Nói cách khác, thông tin chỉ giúp máy biết một tài sản tồn tại, còn tồn kho giao dịch mới giúp máy biến tài sản đó thành doanh thu.

Điều này cũng làm thay đổi cách nhìn về phát triển địa phương. Giá trị của một điểm đến không còn chỉ nằm ở số lượng tài sản mà nó sở hữu, mà ở tỷ lệ tài sản có thể được đưa vào chuỗi giao dịch của nền kinh tế tác nhân. Hai địa phương có cùng số lượng khách sạn, nhà hàng và điểm tham quan có thể có giá trị hoàn toàn khác nhau nếu một nơi đã chuẩn hóa toàn bộ tồn kho để tác nhân có thể tìm kiếm, so sánh và đặt chỗ tự động, trong khi nơi còn lại vẫn phụ thuộc vào điện thoại, tin nhắn hoặc xác nhận thủ công.

Theo nghĩa đó, tồn kho không còn là khái niệm của quản trị vận hành nội bộ mà trở thành hạ tầng của khả hành động. Chỉ khi tài sản được biểu diễn thành tồn kho có thể giao dịch, tác nhân mới có thể lập kế hoạch, tối ưu hành trình, phối hợp nhiều dịch vụ và hoàn tất giao dịch mà không cần con người can thiệp. Đây chính là bước chuyển từ đồ thị tri thức sang đồ thị giao dịch: từ mạng lưới tri thức mô tả thế giới sang mạng lưới các tài nguyên mà máy có thể trực tiếp sử dụng để tạo ra giao dịch.

Ninh Bình là ví dụ gần nhất cho khoảng cách giữa tài sản và tồn kho. Quần thể Tràng An là di sản kép được UNESCO ghi danh, mật độ điểm tham quan trên một bán kính nhỏ thuộc loại cao nhất nước và gần như toàn bộ trải nghiệm cốt lõi là thuyền, xe điện, vé vào cửa theo khung giờ. Đó đều là những thứ về bản chất có thể biểu diễn thành tồn kho: một chuyến thuyền có số ghế, một khung giờ có số suất. Nhưng phần lớn vẫn bán tại quầy, nên với một tác nhân thì Tràng An là một nơi đáng đến mà không đặt được. Khoảng cách ở đây không phải khoảng cách hạ tầng, nó là khoảng cách giữa việc có tồn kho và việc công bố tồn kho.

Vì vậy, trong nền kinh tế tác nhân, câu hỏi chiến lược không còn là "địa phương có gì?", mà là "bao nhiêu phần trăm tài sản

của địa phương đã trở thành tồn kho mà máy có thể giao dịch?". Đó sẽ là một trong những chỉ số quan trọng nhất phản ánh mức độ sẵn sàng của một điểm đến trong kỷ nguyên AI.

### 7.4.1. Máy không mua trải nghiệm, máy mua tồn kho

Con người đi du lịch để tìm kiếm trải nghiệm. Họ mong muốn được thư giãn, khám phá, học hỏi hay tạo nên những ký ức đáng nhớ. Nhưng tác nhân AI không thể trực tiếp mua những khái niệm trừu tượng như "một kỳ nghỉ tuyệt vời" hay "một trải nghiệm đáng nhớ". Điều mà tác nhân có thể tìm kiếm, so sánh và giao dịch chỉ là những đơn vị tài nguyên cụ thể tạo nên trải nghiệm đó. Nói cách khác, con người mua giá trị cảm nhận, còn máy mua khả năng thực hiện.

Một chuyến du lịch đến Phú Quốc không phải là một sản phẩm duy nhất. Đó là sự kết hợp của hàng chục đơn vị tồn kho khác nhau: phòng khách sạn trong hai đêm, hai vé máy bay, một bàn ăn tối, hai vé cáp treo Hòn Thơm, một suất tour lặn ngắm san hô ở quần đảo An Thới, một xe đưa đón sân bay và một hướng dẫn viên trong ba giờ. Đối với con người, tất cả hợp thành một trải nghiệm thống nhất. Nhưng đối với tác nhân, đó là một bài toán ghép nối các đơn vị tồn kho độc lập.

Điều này giải thích vì sao AI không thể giao dịch chỉ dựa trên nội dung quảng bá. Một video đẹp, một câu chuyện hấp dẫn hay một chiến dịch truyền thông thành công có thể khiến con người muốn đến một nơi. Nhưng nếu tác nhân không tìm thấy các đơn vị tồn kho có thể đặt ngay lập tức, quá trình giao dịch sẽ dừng lại. Máy không thể mua một lời hứa. Máy chỉ có thể mua một tài nguyên có trạng thái xác định.

Vì vậy, trong nền kinh tế tác nhân, khoảng cách giữa "được yêu thích" và "được giao dịch" chính là khoảng cách giữa trải nghiệm và tồn kho. Địa phương chỉ thực sự tham gia chuỗi giá trị mới khi mọi thành phần của trải nghiệm đều được chuyển hóa thành các đơn vị giao dịch mà tác nhân có thể nhận diện, kiểm tra tình trạng và đặt chỗ tự động.

Nói theo cách khác, trải nghiệm là sản phẩm cuối cùng dành cho con người, còn tồn kho là ngôn ngữ giao dịch dành cho máy. Mọi trải nghiệm đều phải được phân rã thành những đơn vị tồn kho trước khi tác nhân có thể mua, kết hợp và tạo thành một hành trình hoàn chỉnh.

### 7.4.2. Điều gì thực sự được bán: phòng, ghế, vé, suất, khung giờ

Khi quan sát nền kinh tế tác nhân, có vẻ như tác nhân đang mua một chuyến đi, một kỳ nghỉ hay một gói trải nghiệm. Nhưng ở tầng dữ liệu, thứ thực sự được giao dịch là những đơn vị năng lực phục vụ trong một khoảng thời gian xác định. Khách sạn không bán một kỳ nghỉ mà bán một phòng còn trống vào một đêm cụ thể. Nhà hát không bán nghệ thuật mà bán một ghế trong một suất diễn. Phòng khám không bán sức khỏe mà bán một khung giờ của một bác sĩ. Bảng 7.2 liệt kê đầy đủ hơn phép quy đổi này cho từng nhóm chủ thể.

Tất cả những đơn vị này đều có một đặc điểm chung: chúng tồn tại với số lượng hữu hạn, có trạng thái thời gian, có thể được giữ chỗ, được xác nhận, được thanh toán và không thể bán hai lần cho cùng một thời điểm. Đó chính là bản chất của tồn kho trong nền kinh tế số.

Nhìn từ góc độ này, mọi doanh nghiệp dịch vụ thực chất đều đang quản lý một tập hợp các đơn vị tồn kho có khả năng giao dịch. Giá, công suất, tỷ lệ lấp đầy và doanh thu đều được quyết định bởi cách các đơn vị tồn kho đó được phân bổ theo thời gian. Tác nhân không cần hiểu toàn bộ doanh nghiệp vận hành ra sao. Nó chỉ cần biết đơn vị tồn kho nào đang khả dụng, với điều kiện gì và có thể đặt ngay hay không.

Vì vậy, bài toán chuyển đổi số của địa phương không phải là số hóa doanh nghiệp theo nghĩa chung chung, mà là biến mọi năng lực phục vụ thành những đơn vị tồn kho có định danh, có trạng thái thời gian và có thể giao dịch qua API. Chỉ khi đó, các tác nhân mới có thể tự động kết hợp hàng trăm đơn vị tồn kho từ nhiều nhà cung cấp khác nhau để tạo thành một trải nghiệm hoàn chỉnh cho con người. Đây chính là nền tảng của một điểm đến lắp ghép được (composable destination), nơi mọi dịch vụ đều có thể được máy lắp ghép, tối ưu và giao dịch theo thời gian thực. Bảng 7.2 *(trang sau)* minh họa bước chuyển từ tài sản sang đơn vị tồn kho giao dịch được.

### 7.4.3. Tồn kho phải sống: khả dụng theo thời gian thực

Trong nền kinh tế tác nhân, tồn kho không phải là một danh sách tĩnh mà là một thực thể sống. Giá trị của nó không nằm ở việc tồn tại trên trang web, mà ở việc phản ánh chính xác trạng thái hiện tại của thế giới thực.

Một phòng khách sạn có thể còn trống lúc 10 giờ sáng nhưng đã được đặt lúc 10 giờ 05. Một bàn ăn vừa được khách hủy. Một chuyến xe điện vừa đầy chỗ. Một bác sĩ vừa kéo dài thời gian khám. Những thay đổi này diễn ra liên tục và làm thay đổi ngay lập tức khả năng giao dịch của toàn bộ hệ thống.

Đối với con người, thông tin chậm vài phút đôi khi vẫn chấp nhận được. Nhưng đối với tác nhân, dữ liệu cũ đồng nghĩa với quyết định sai. Nếu tác nhân đặt một phòng đã hết, đề xuất của nó thất bại. Nếu tác nhân xây dựng một hành trình dựa trên những vé không còn tồn tại, toàn bộ kế hoạch phải tính toán lại từ đầu. Vì vậy, tác nhân không chỉ cần biết tồn kho là gì, mà còn phải biết tồn kho đang ở trạng thái nào. Mỗi đơn vị tồn kho phải liên tục cập nhật các thuộc tính như còn hay hết, số lượng còn lại, thời điểm khả dụng tiếp theo, thời gian giữ chỗ, điều kiện hủy và các thay đổi phát sinh trong quá trình vận hành.

Đó là lý do khả dụng trở thành một thành phần của dữ liệu chứ không còn là thông tin phụ. Một đồ thị tri thức chỉ mô tả thế giới là chưa đủ. Nền kinh tế tác nhân cần một đồ thị sống, một đồ thị mà các thực thể luôn thay đổi theo thời gian thực. Nói cách khác, tồn kho chỉ thực sự tồn tại đối với tác nhân khi nó luôn "sống". Dữ liệu càng gần với trạng thái vận hành thực tế, tác nhân càng có thể hành động chính xác và đáng tin cậy.

### 7.4.4. Giá là một trạng thái của tồn kho

Trong tư duy truyền thống, giá được xem là trung tâm của giao dịch. Doanh nghiệp công bố một bảng giá, khách hàng lựa chọn rồi quyết định mua hay không. Nhưng trong nền kinh tế tác nhân, giá chỉ còn là một thuộc tính của tồn kho, giống như vị trí, sức chứa hay thời gian phục vụ.

Tác nhân quan tâm tới giá của đúng đơn vị tồn kho tại đúng thời điểm, không tới một bảng giá cố định. Một phòng tiêu chuẩn vào tối thứ Bảy có thể có giá khác hoàn toàn cùng căn phòng vào chiều thứ Ba. Một ghế ở hàng đầu của buổi hòa nhạc có giá khác

**Bảng 7.2.** Từ tài sản sang đơn vị tồn kho giao dịch được

| Chủ thể | Tài sản như con người nhìn | Đơn vị tồn kho như máy giao dịch | Thuộc tính bắt buộc |
|---|---|---|---|
| **Khách sạn** | Nơi lưu trú | Một phòng thuộc một hạng cho một đêm cụ thể | Còn hay hết, giá hiệu lực, điều kiện hủy |
| **Nhà hàng** | Ẩm thực | Một bàn hoặc một chỗ trong một khung giờ | Sức chứa còn lại, thời gian giữ chỗ |
| **Điểm tham quan** | Di sản | Một vé cho một khung giờ vào cửa | Hạn ngạch theo khung giờ, chính sách vé |
| **Nhà hát** | Nghệ thuật | Một ghế trong một suất diễn xác định | Sơ đồ chỗ, hạng vé |
| **Cơ sở y tế** | Chăm sóc sức khỏe | Một khung giờ khám của một bác sĩ | Chuyên khoa, thời lượng, điều kiện tiếp nhận |
| **Nghệ nhân, hướng dẫn viên** | Tri thức và kỹ năng | Một khoảng thời gian phục vụ | Lịch trống, ngôn ngữ, sức chứa nhóm |
| **Bãi đỗ, phương tiện** | Hạ tầng | Một chỗ hoặc một chuyến trong một khoảng thời gian | Vị trí, sức chứa, thời gian khả dụng |

ghế cuối. Một khung giờ khám vào buổi tối khác với buổi sáng. Giá không còn gắn với sản phẩm nói chung mà gắn với từng đơn vị tồn kho cụ thể.

Vì vậy, định giá cho máy là khả năng để tác nhân truy cập mức giá đang có hiệu lực ngay khi ra quyết định, không phải việc công bố một danh mục giá. Giá trở thành dữ liệu động, thay đổi theo công suất, nhu cầu, thời gian, chính sách khách hàng thân thiết, chương trình khuyến mại hay nhiều điều kiện kinh doanh khác.

Quan trọng hơn, tác nhân không tối ưu theo giá thấp nhất. Nó tối ưu theo giá trị tổng thể. Một khách sạn đắt hơn nhưng gần ga tàu, có thể nhận phòng ngay và giúp giảm chi phí di chuyển vẫn có thể là lựa chọn tối ưu. Vì thế, giá chỉ là một biến trong bài toán tối ưu đa mục tiêu mà tác nhân giải quyết. Trong nền kinh tế tác nhân, doanh nghiệp không còn cạnh tranh bằng bảng giá đẹp, mà bằng khả năng cung cấp giá đúng, cho đúng tồn kho, tại đúng thời điểm và dưới đúng điều kiện giao dịch.

Tồn kho sống. Giá trị của một đơn vị tồn kho đối với máy suy giảm theo độ trễ cập nhật và suy giảm nhanh hơn nhiều so với đối với con người. Với con người, dữ liệu chậm là dữ liệu kém tiện. Với tác nhân, dữ liệu chậm là dữ liệu sai, vì nó dẫn tới một hành động thất bại chứ không dẫn tới một câu hỏi bổ sung.

Đồ thị tri thức cho tác nhân khả năng suy luận. Tồn kho số cho tác nhân khả năng thực thi. Thiếu đồ thị, máy không hiểu địa phương. Thiếu tồn kho số, máy hiểu nhưng không hoàn tất được bất kỳ hành động nào, nên dù địa phương xuất hiện trong mọi câu trả lời của AI, phần giá trị kinh tế vẫn dừng ở bước đề xuất. Đó là lý do chuyển đổi số trong nền kinh tế tác nhân không kết thúc ở trang web, dữ liệu mở hay giao diện tra cứu.

Ví dụ dưới đây là một kịch bản do tác giả dựng để làm rõ cấu trúc chi phí, không phải số liệu khảo sát của một địa phương có thật. Giả sử một thị xã có khoảng 1.200 thực thể du lịch đã được định danh, ngân sách xúc tiến và chuyển đổi số hằng năm 12,5 tỷ đồng. Ở kịch bản A, gần như toàn bộ ngân sách dành cho truyền thông, chỉ 0,5 tỷ đồng dành cho duy trì dữ liệu. Ở kịch bản B, ngân sách được chia lại thành 7 tỷ đồng cho xúc tiến và 5,5 tỷ đồng cho dữ liệu, tín hiệu và năng lực giao dịch. Tổng chi hai kịch bản bằng nhau. Điều thay đổi không phải quy mô chi

mà là vị trí của đồng chi trong chuỗi bốn bộ lọc. Điểm cần nhấn mạnh, kịch bản B không hứa hẹn nhiều khách hơn ngay trong năm đầu. Nó chỉ chuyển địa phương từ chỗ không thể được giao việc sang chỗ có thể.[1]

Doanh nghiệp nhỏ không đủ năng lực xây dựng API. Hoàn toàn đúng nếu hiểu yêu cầu theo nghĩa mỗi hộ kinh doanh tự viết một giao diện riêng. Nhưng đó không phải cách thị trường vận hành. Phần lớn năng lực giao dịch của doanh nghiệp nhỏ sẽ đến qua các hệ thống có sẵn mà họ đã dùng: phần mềm bán hàng, phần mềm quản lý lưu trú, ví điện tử, nền tảng đặt chỗ dùng chung. Nhiệm vụ của chính quyền là bảo đảm những hệ thống trung gian ấy nói cùng một chuẩn định danh, không phải dạy hộ kinh doanh viết mã.

Đối với các doanh nghiệp nhỏ, chỉ cần thực hiện tốt bốn việc sau đã có thể cải thiện đáng kể khả năng được tác nhân AI lựa chọn:

- ***Luôn duy trì ba dữ liệu quan trọng nhất:*** giờ hoạt động, tình trạng còn chỗ và giá đang áp dụng. Đây là ba tín hiệu ảnh hưởng trực tiếp đến khả năng được đề xuất và hoàn tất giao dịch.

- ***Không tự phát triển hệ thống kết nối.*** Hãy sử dụng một phần mềm quản lý đã hỗ trợ các chuẩn kết nối phổ biến, sau đó tập trung bảo đảm dữ liệu trong hệ thống luôn phản ánh đúng tình trạng vận hành thực tế.

1. Tính toán của tác giả trên một kịch bản dựng để minh họa cơ chế, không phải số liệu quan sát của một doanh nghiệp hay địa phương có thật.

- ***Chỉ sử dụng một định danh duy nhất trên mọi nền tảng:*** cùng một tên pháp lý, một địa chỉ và một tọa độ. Hồ sơ trùng lặp sẽ phân tán đánh giá, làm suy giảm độ tin cậy và khả năng được nhận diện.

- ***Rút ngắn thời gian xác nhận đặt chỗ*** xuống dưới ngưỡng mà tác nhân AI có thể chờ. Trong nền kinh tế tác nhân, một yêu cầu được xác nhận sau nhiều giờ thường có giá trị gần như một yêu cầu không được xác nhận.

Danh mục chỉ chứa điều đã được viết ra. Đồ thị cho phép suy ra điều chưa ai viết.

## 7.5. TỪ TỒN KHO ĐẾN GIAO DỊCH

Nếu mục 7.4 xác định thứ tác nhân giao dịch là những đơn vị tồn kho có trạng thái, thì mục này xem xét cơ chế để tác nhân truy cập, giữ chỗ, xác nhận và thanh toán cho những đơn vị ấy. Một bên là đối tượng, một bên là đường dẫn tới đối tượng. Thiếu bên nào thì giao dịch cũng không hình thành.

Đây là ranh giới giữa khả đọc bằng máy (machine-readable) và khả đặt bằng máy (machine-bookable). Một địa phương có thể được AI mô tả rất chính xác, được đưa vào câu trả lời, thậm chí được đề xuất cho hàng triệu người dùng. Nhưng nếu tác nhân không hoàn tất được bước đặt phòng, mua vé, giữ chỗ hay thanh toán thì toàn bộ quá trình dừng ở mức gợi ý.

Trong nền kinh tế tác nhân, tài sản quan trọng nhất không còn là một bài giới thiệu hay một chiến dịch quảng bá. Thứ quyết định dòng tiền là khả năng để máy hoàn thành giao dịch mà không cần con người can thiệp. Một phòng khách sạn có thể đặt ngay, một vé tham quan còn chỗ theo thời gian thực, một khung giờ khám còn trống hay một bàn ăn có thể xác nhận tức thì đều trở thành những đơn vị kinh tế mà AI có thể trực tiếp phân phối.

Điều này làm thay đổi cách nhìn về dịch vụ. Trước đây, doanh nghiệp thiết kế dịch vụ để khách hàng dễ tìm và dễ mua. Trong nền kinh tế tác nhân, dịch vụ còn phải được thiết kế để máy có thể hiểu, so sánh, lựa chọn và hoàn tất giao dịch. Nói cách khác, dịch vụ không chỉ cần thân thiện với con người mà còn phải thân thiện với các tác nhân số. Vì vậy, cạnh tranh giữa các địa phương không còn chỉ diễn ra ở chất lượng trải nghiệm hay sức mạnh thương hiệu. Nó diễn ra ngay tại chỗ chuyển ý định của máy thành doanh thu thực tế. Hai địa phương có thể hấp dẫn

như nhau trong mắt hệ thống, nhưng nơi có dịch vụ đọc được, tồn kho cập nhật và giao dịch tự động sẽ nhận dòng khách trước.

### 7.5.1. Từ khả đọc bằng máy đến khả đặt bằng máy

Một địa phương được máy đọc được chưa có nghĩa là được máy mua được. Sự khác biệt giữa hai cấp độ không nằm ở lượng thông tin mà ở loại tác động. Khả đọc bằng máy là trạng thái cho phép máy quan sát: nơi này ở đâu, thuộc loại gì, giá khoảng bao nhiêu, phù hợp với ai. Khả đặt bằng máy là trạng thái cho phép máy yêu cầu hệ thống chuyển từ trạng thái này sang trạng thái khác và nhận lại một cam kết có hiệu lực. Đọc là quan sát trạng thái. Đặt là can thiệp vào trạng thái. Một lời gọi kỹ thuật chưa phải giao dịch nếu thiếu giữ chỗ, xác nhận, chủ thể chịu trách nhiệm và cơ chế xử lý thất bại.

Ở trạng thái thứ hai, tác nhân không chỉ biết rằng còn phòng trống mà biết phòng nào còn, giá hiện tại bao nhiêu, điều kiện hủy ra sao, thanh toán bằng phương thức nào và nhận được xác nhận ngay trong cùng phiên làm việc. Khoảng cách giữa hai cấp độ này chính là khoảng cách giữa được biết đến và thu được tiền.

Điều này cũng làm thay đổi tiêu chuẩn đánh giá chuyển đổi số của địa phương. Một trang web đẹp, nhiều hình ảnh hay nội dung phong phú có thể giúp tăng khả năng được đọc. Nhưng nếu việc đặt dịch vụ vẫn phải gọi điện, nhắn tin, chờ xác nhận hoặc nhập lại thông tin nhiều lần, địa phương mới chỉ đạt được khả đọc bằng máy chứ chưa đạt khả đặt bằng máy. Trong nền kinh tế tác nhân, khoảng cách giữa hai cấp độ này chính là khoảng cách giữa được biết đến và thu được tiền.

### 7.5.2. API là cánh cửa, không phải mục tiêu

Khi nói đến khả đặt bằng máy, nhiều người lập tức nghĩ đến API. Điều đó đúng nhưng chưa đủ. API không phải giá trị kinh tế. API chỉ là cánh cửa để máy bước vào hệ thống.

Một cánh cửa rộng nhưng phía sau không có dữ liệu chính xác, không có tồn kho thời gian thực, không có cơ chế xác nhận, không có thanh toán hay không có quy trình thực hiện dịch vụ thì AI vẫn không thể hoàn tất hành động. Ngược lại, một hệ thống có dữ liệu vận hành đầy đủ nhưng không có cách để tác nhân truy cập cũng giống như một cửa hàng khóa kín. Hàng hóa vẫn ở đó nhưng không ai mua được. Vì vậy, điều đáng hỏi không phải địa phương có API hay không mà là API dẫn tới một hệ thống có thể giao dịch.

Nhìn dưới góc độ kinh tế, API làm giảm mạnh chi phí giao dịch giữa các tác nhân. Thay vì phải gọi điện, gửi email, xác nhận thủ công hay nhập lại dữ liệu, phần lớn quá trình có thể diễn ra trong vài giây và chi phí biên của việc trao đổi, xử lý giảm rất mạnh. Cần nói rõ rằng đây là chi phí trao đổi thông tin, không phải toàn bộ chi phí giao dịch: xác thực danh tính, vận hành, hoàn tiền, xử lý tranh chấp và bảo đảm an ninh vẫn tốn kém và trong một số trường hợp còn tăng lên. Khi chi phí giao dịch giảm, nhiều giao dịch nhỏ trước đây không đáng để thực hiện nay cũng trở nên có lợi về mặt kinh tế. Chính điều đó mở rộng quy mô của thị trường. Vì thế, API nên được xem như đường cao tốc của nền kinh tế tác nhân chứ không phải đích đến của chuyển đổi số. Đích đến luôn là khả năng để mọi dịch vụ có thể được khám phá, lựa chọn, đặt chỗ, thanh toán và thực hiện một cách tự động.

Một địa phương sẽ không chiến thắng vì sở hữu nhiều API hơn đối thủ. Họ chiến thắng khi các API ấy kết nối trực tiếp với

hệ thống vận hành thật, phản ánh đúng tình trạng dịch vụ và cho phép hàng triệu tác nhân số giao dịch liên tục mà gần như không cần con người can thiệp. Đó mới là nền tảng tạo nên tăng trưởng kinh tế trong kỷ nguyên tác nhân.

Từ cuối năm 2025 đến giữa năm 2026, có thể nhận diện ít nhất năm sáng kiến và giao thức cạnh tranh nhằm trở thành hạ tầng chuẩn cho việc tác nhân AI mua hàng và giao dịch thay con người. Google giới thiệu Agent Payments Protocol với sự tham gia của hơn 60 đối tác như Mastercard, PayPal, American Express và Coinbase, dựa trên cơ chế ủy quyền bằng mandate ký số. OpenAI và Stripe phát triển Agentic Commerce Protocol, triển khai đầu tiên qua tính năng thanh toán ngay trong ChatGPT. Google và Shopify tiếp tục công bố Universal Commerce Protocol vào tháng 1/2026. Cùng thời gian đó, Coinbase hồi sinh mã trạng thái HTTP 402 thông qua giao thức x402 và đến tháng 4/2026, Linux Foundation thành lập quỹ bảo trợ để phát triển chuẩn này. Cũng trong tháng 4/2026, Google chuyển Agent Payments Protocol sang cho FIDO Alliance quản lý.[1]

Điều quan trọng đối với các địa phương không phải là dự đoán giao thức nào sẽ chiến thắng. Các chuẩn ở tầng giao dịch có thể thay đổi rất nhanh. Ngược lại, tầng nền gồm định danh thực thể, chất lượng dữ liệu và khả năng kết nối gần như không phụ thuộc vào kết quả của cuộc cạnh tranh đó. Đầu tư vào dữ liệu nền vì thế có tuổi thọ dài hơn nhiều so với đầu tư theo một giao thức cụ thể.

1. Google Cloud, thông cáo Agent Payments Protocol ngày 16 tháng 9 năm 2025 với hơn 60 đối tác, https://cloud.google.com/blog/products/ai-machine-learning/announcing-agents-to-payments-ap2-protocol. Universal Commerce Protocol do Google công bố tháng 1 năm 2026 tại hội nghị National Retail Federation, xây dựng cùng Shopify. Linux Foundation công bố lập x402 Foundation tháng 4 năm 2026, tổ chức này vận hành đầy đủ từ tháng 7 năm 2026. Google chuyển Agent Payments Protocol cho FIDO Alliance ngày 28 tháng 4 năm 2026 kèm bản 0.2.

Nhìn từ góc độ địa phương, các giao thức này có thể chia thành bốn nhóm:

***Nhóm 1:*** Kết nối công cụ và dữ liệu giúp tác nhân AI truy cập các hệ thống bên ngoài thông qua một giao diện thống nhất. Đại diện tiêu biểu là Model Context Protocol (MCP). Điều địa phương cần chuẩn bị không phải là tự xây giao thức mới, mà là bảo đảm dữ liệu như tồn kho, lịch hoạt động hay tình trạng còn chỗ có thể được máy truy cập qua các giao diện chuẩn.

***Nhóm 2:*** Thương mại và thanh toán chuẩn hóa toàn bộ quy trình từ giỏ hàng, ủy quyền thanh toán đến vòng đời đơn hàng. Các sáng kiến như Agentic Commerce Protocol, Agent Payments Protocol hay Universal Commerce Protocol đều thuộc nhóm này. Đối với địa phương, giải pháp thực tế là lựa chọn các hệ thống trung gian đã hỗ trợ ít nhất một trong những chuẩn phổ biến, thay vì tự triển khai từ đầu.

***Nhóm 3:*** Thanh toán máy với máy hướng tới việc tác nhân AI tự chi trả để truy cập dữ liệu hoặc sử dụng dịch vụ số. Những giao thức như x402 hay Machine Payments Protocol mở ra khả năng thương mại hóa dữ liệu và các dịch vụ số của địa phương theo từng lượt sử dụng, thay vì chỉ bán các sản phẩm du lịch truyền thống.

***Nhóm 4:*** Ủy quyền và trách nhiệm là lớp bảo đảm tác nhân AI chỉ hành động khi đã được người dùng cho phép và có thể xác định rõ trách nhiệm nếu xảy ra sai sót. Cơ chế phổ biến là mandate ký số dựa trên các chuẩn chứng thư có thể xác minh. Với địa phương, đây không chỉ là vấn đề kỹ thuật mà còn là câu hỏi quản trị: khi tác nhân đặt nhầm, thanh toán sai hoặc cung cấp thông tin sai, ai là người có trách nhiệm xác nhận, sửa lỗi và bồi hoàn. Đây sẽ là một trong những nền tảng quan trọng của nền kinh tế tác nhân trong những năm tới.

### 7.5.3. Bốn điều kiện để tác nhân dám giao dịch

Con người có thể chấp nhận sự không chắc chắn. Nếu chưa rõ giá, họ có thể gọi điện hỏi. Nếu chưa biết còn chỗ hay không, họ có thể nhắn tin. Nếu chưa có xác nhận ngay, họ vẫn sẵn sàng chờ. Tác nhân thì không.

Một tác nhân số chỉ hành động khi đủ điều kiện để tin rằng giao dịch có thể hoàn thành với xác suất thành công cao. Nếu một mắt xích còn thiếu, tác nhân thường sẽ chuyển sang lựa chọn khác thay vì tiếp tục dò hỏi như con người. Có bốn điều kiện quyết định niềm tin đó:

- ***Thứ nhất là giá.*** Tác nhân cần biết không chỉ mức giá mà còn biết giá đang có hiệu lực, bao gồm các điều kiện đi kèm như thuế, phụ phí, chính sách hoàn hủy hay ưu đãi. Một mức giá không rõ ràng đồng nghĩa với một giao dịch không thể tính toán.
- ***Thứ hai là lịch và tình trạng sẵn sàng phục vụ.*** Một dịch vụ chỉ thực sự tồn tại khi tác nhân biết nó còn khả dụng vào đúng thời điểm người dùng yêu cầu. Lịch mở cửa, số lượng còn lại, khung giờ trống hay công suất phục vụ đều là dữ liệu quyết định.
- ***Thứ ba là khả năng xác nhận.*** Tác nhân phải biết rằng sau khi gửi yêu cầu, hệ thống có thể trả lời "đã xác nhận", "đã giữ chỗ" hay "không còn chỗ" trong thời gian rất ngắn. Một giao dịch treo nhiều giờ hoặc nhiều ngày gần như không còn phù hợp với logic của nền kinh tế tác nhân.
- ***Cuối cùng là thanh toán.*** Giao dịch chỉ kết thúc khi giá trị được chuyển giao. Nếu không có cơ chế thanh toán rõ ràng, tác nhân chỉ mới hoàn thành một yêu cầu thông tin chứ chưa hoàn thành một hành động kinh tế.

Bốn điều kiện này tạo nên niềm tin giao dịch của máy. Chúng hoạt động như những tín hiệu giúp tác nhân đánh giá rằng một giao dịch có thể hoàn tất với chi phí thấp và rủi ro thấp. Địa phương nào cung cấp được những tín hiệu này một cách nhất quán sẽ trở thành nơi mà các tác nhân ưu tiên lựa chọn, bởi với máy, độ tin cậy của quy trình quan trọng không kém sức hấp dẫn của trải nghiệm.

Bốn điều kiện trên là phần lõi. Ngoài chúng còn một lớp bảo đảm gồm hai thành phần mà thực tiễn cho thấy không thể bỏ qua. Thứ nhất là định danh chủ thể chịu trách nhiệm: khi giao dịch sai, tác nhân phải biết gửi khiếu nại tới ai và ai có nghĩa vụ bồi hoàn. Thứ hai là cơ chế hoàn hủy và xử lý ngoại lệ: một hệ thống chỉ cho phép đặt mà không cho phép hủy hoặc sửa sẽ bị tác nhân xếp vào nhóm rủi ro cao, bởi hành động không đảo ngược được là loại hành động mà mọi hệ thống tự động đều tránh. Bốn điều kiện lõi quyết định máy có bắt đầu hay không. Lớp bảo đảm quyết định máy có quay lại lần sau hay không.

### 7.5.4. Khi mọi dịch vụ đều trở thành một hàm có thể gọi

Trong nền kinh tế tác nhân, dịch vụ không còn chỉ là một sản phẩm được mô tả trên trang web. Đối với tác nhân, mỗi dịch vụ dần trở thành một năng lực có thể được gọi khi cần.

Một khách sạn không chỉ là nơi lưu trú mà trở thành một hàm có thể thực hiện lệnh đặt phòng. Một nhà hàng trở thành một hàm đặt bàn. Một bảo tàng trở thành một hàm mua vé. Một bệnh viện trở thành một hàm đặt lịch khám. Một công ty vận chuyển trở thành một hàm đặt xe.

Điểm chung của các hàm này là chúng đều có đầu vào, quy tắc xử lý và đầu ra rõ ràng. Tác nhân gửi yêu cầu với các tham số như thời gian, số lượng người, ngân sách hay sở thích, hệ thống xử lý, sau đó trả về một kết quả xác định: đặt thành công, hết chỗ, cần thay đổi thời gian hoặc đề xuất phương án khác.

Điều này tạo ra một thay đổi rất lớn trong cách tổ chức nền kinh tế. Trước đây, doanh nghiệp cạnh tranh bằng khả năng thu hút khách truy cập vào trang web hoặc ứng dụng của mình. Trong nền kinh tế tác nhân, doanh nghiệp cạnh tranh bằng việc dịch vụ của mình có thể được tác nhân gọi và sử dụng dễ dàng đến đâu.

Nói cách khác, đơn vị cạnh tranh không còn là trang web hay ứng dụng mà là khả năng (capability). Một tác nhân không quan tâm giao diện đẹp đến mức nào. Nó quan tâm liệu có thể gọi đúng chức năng, nhận đúng kết quả và hoàn thành giao dịch với độ tin cậy cao hay không.

Khi ngày càng nhiều dịch vụ được chuẩn hóa theo cách này, địa phương cũng thay đổi vai trò. Thay vì chỉ là tập hợp các doanh nghiệp, địa phương trở thành một thư viện các năng lực kinh tế mà hàng triệu tác nhân có thể kết hợp để xây dựng hành trình cho người dùng. Mỗi khách sạn, nhà hàng, điểm tham quan hay phương tiện vận chuyển trở thành một "khối chức năng" có thể lắp ghép với các khối khác để tạo thành vô số trải nghiệm khác nhau.

Đó là ý nghĩa sâu xa của khả đặt bằng máy. Mục tiêu cuối cùng không phải là số hóa từng dịch vụ riêng lẻ, mà là biến toàn bộ năng lực của một địa phương thành những khả năng có thể được tác nhân khám phá, gọi và phối hợp, qua đó mở rộng quy mô giao dịch mà không cần mở rộng tương ứng quy mô của bộ máy trung gian.

Ẩn dụ hàm cần một giới hạn rõ ràng. Không phải toàn bộ trải nghiệm có thể biến thành hàm. Chỉ phần giao dịch và điều phối là chuẩn hóa được: giữ chỗ, xác nhận, thanh toán, đổi lịch, hoàn tiền. Chất lượng tiếp xúc con người, khả năng ứng biến trước tình huống bất thường và ý nghĩa văn hóa của một trải nghiệm vẫn nằm ngoài lời gọi kỹ thuật và chính chúng mới là thứ khiến người ta quay lại. Chuẩn hóa phần gọi được để bảo vệ phần không gọi được, chứ không phải để thay thế nó.

Tháng 5 năm 2026, Expedia Group xác nhận sẽ triển khai máy chủ theo giao thức Model Context Protocol (MCP) cho khối khách hàng doanh nghiệp, cho phép các tác nhân AI của đối tác kết nối trực tiếp tới hệ thống lưu trú, vé máy bay, thuê xe và hoạt động trải nghiệm mà không cần xây dựng các tích hợp riêng cho từng dịch vụ. Cùng thời điểm, Expedia cho biết hạ tầng doanh nghiệp của họ đang xử lý khoảng 21 tỷ lượt gọi API mỗi ngày. Đây là số liệu do chính doanh nghiệp công bố, chưa được kiểm toán độc lập, nên phù hợp để hiểu như một chỉ dấu về quy mô hơn là một phép đo chính xác. Điều đáng chú ý không nằm ở bản thân con số 21 tỷ, mà ở xu hướng mà nó phản ánh. Trong ngành du lịch, một phần rất lớn các quyết định tìm kiếm, so sánh, kiểm tra tồn kho và đặt dịch vụ đã diễn ra dưới dạng máy gọi máy (machine-to-machine), hoàn toàn phía sau giao diện mà người dùng nhìn thấy. Khi tác nhân AI trở thành người thực hiện các bước này, khả năng xuất hiện trong các luồng API và giao thức kết nối quan trọng không kém, thậm chí quan trọng hơn, việc sở hữu một trang web hay chiến dịch quảng bá hấp dẫn.

Nói cách khác, một địa phương có thể rất nổi bật trên màn hình của con người nhưng vẫn hoàn toàn vô hình đối với tác nhân AI nếu dữ liệu và dịch vụ của mình không thể được truy cập qua các giao diện chuẩn. Trong nền kinh tế tác nhân, không

được máy kết nối cũng đồng nghĩa với không được máy lựa chọn. Máy không mua một lời hứa. Máy mua một tài nguyên có trạng thái xác định.

### 7.5.5. Địa phương không còn bán thông tin, địa phương bán khả năng thực hiện

Trong hơn ba thập niên của internet, phần lớn chiến lược số của các địa phương xoay quanh một mục tiêu là cung cấp thông tin. Họ xây dựng cổng thông tin du lịch, thiết kế trang web đẹp hơn, viết bài giới thiệu hấp dẫn hơn, đăng nhiều hình ảnh hơn, tối ưu công cụ tìm kiếm và đầu tư cho truyền thông số. Toàn bộ logic ấy dựa trên một giả định: người quyết định cuối cùng là con người. Khi có đủ thông tin, khách du lịch sẽ tự gọi điện, tự đặt phòng, tự mua vé và tự tổ chức chuyến đi. Nền kinh tế tác nhân làm giả định đó không còn đúng.

Khi tác nhân trở thành người thực hiện phần lớn quá trình lựa chọn và giao dịch, giá trị của thông tin giảm xuống. Điều mà tác nhân cần không phải là một bài viết dài hơn hay một video hấp dẫn hơn. Điều nó cần là khả năng hành động. Một mô tả về khách sạn không giúp tác nhân hoàn thành nhiệm vụ. Một chức năng đặt phòng thì có. Một bài giới thiệu về nhà hàng không tạo ra doanh thu. Một khả năng giữ bàn ngay lập tức thì có. Một bài viết về lễ hội không tạo ra dòng tiền. Một chức năng mua vé và xác nhận tham dự mới tạo ra giá trị kinh tế.

Sự thay đổi này gợi lại quá trình chuyển đổi của thương mại điện tử hai thập niên trước. Khi thương mại điện tử xuất hiện, doanh nghiệp không còn cạnh tranh bằng danh mục tĩnh sản phẩm mà bằng khả năng giao hàng. Trong nền kinh tế tác nhân,

địa phương cũng không còn cạnh tranh bằng lượng thông tin mà bằng khả năng thực hiện dịch vụ. Nói cách khác, đơn vị giá trị đã thay đổi. Trước đây, địa phương cung cấp tri thức. Ngày nay, địa phương phải cung cấp năng lực.

Mỗi dịch vụ của địa phương cần trả lời được một câu hỏi rất cụ thể: Tác nhân có thể hoàn thành việc này ngay bây giờ hay không? Nếu câu trả lời là không, dù thông tin có đầy đủ đến đâu, địa phương vẫn chỉ đang cung cấp tài liệu tham khảo. Đó là lý do vì sao khả năng thực hiện (execution capability) trở thành tài sản kinh tế mới của địa phương. Giá trị không còn nằm ở việc AI biết bao nhiêu về nơi này, mà nằm ở việc AI có thể làm được bao nhiêu việc với nơi này.

Điều này cũng làm thay đổi cách đánh giá mức độ phát triển số của một địa phương. Trước đây, người ta hỏi:

- *Có trang web không?*
- *Có ứng dụng không?*
- *Có bao nhiêu bài viết?*
- *Có bao nhiêu lượt truy cập?*

Trong nền kinh tế tác nhân, những câu hỏi quan trọng hơn là:

- *Có bao nhiêu dịch vụ tác nhân có thể gọi trực tiếp?*
- *Bao nhiêu giao dịch được hoàn tất hoàn toàn tự động?*
- *Bao nhiêu tồn kho được cập nhật theo thời gian thực?*
- *Bao nhiêu hoạt động kinh tế có thể diễn ra mà không cần con người can thiệp?*

Đó là các chỉ số phản ánh năng lực kinh tế, chứ không chỉ phản ánh năng lực truyền thông. Chuyển đổi số vì thế không còn

dừng ở việc số hóa thông tin để con người đọc, mà phải số hóa phần năng lực có thể được máy khám phá, gọi và thực hiện. Khi điều đó xảy ra, địa phương không chỉ hiện diện trên internet mà trở thành một nền kinh tế có thể phối hợp bằng phần mềm. Cụm nền kinh tế lập trình được, nếu dùng, chỉ nên hiểu là ẩn dụ cho lớp điều phối và giao dịch, không phải cho toàn bộ đời sống của một địa phương. Giá trị không đến từ việc được mô tả nhiều hơn mà từ khả năng biến ý định thành hành động hoàn chỉnh.

Mở rộng biên thị trường. Khi chi phí tìm kiếm, phối hợp và xử lý giao dịch giảm đủ sâu, thị trường có thể mở rộng về phía những giao dịch giá trị nhỏ trước đây không đáng thực hiện: một giờ của nghệ nhân, một suất học nghề, một chỗ trên chuyến thuyền chiều, một khung giờ trống của phòng khám. Nếu đúng, phần tăng trưởng lớn nhất của nền kinh tế tác nhân tại Việt Nam sẽ nằm ở khu vực hộ kinh doanh chứ không ở khu vực doanh nghiệp lớn. Chiều ngược lại cũng có lý, chi phí tuân thủ và chi phí chuẩn hóa cố định có thể loại chính nhóm này ra khỏi cuộc chơi.

Nên chờ chuẩn quốc tế ổn định rồi hãy triển khai. Đây là chỗ thứ hai chương này để ngỏ. Lập luận có lý về mặt tài chính: giai đoạn 2025 và 2026 chứng kiến ít nhất năm giao thức cạnh tranh nhau ở tầng thanh toán và thương mại tác nhân, đầu tư sớm vào chuẩn thua cuộc là mất vốn thật. Nhưng phần lớn chi phí thực của địa phương không nằm ở tầng giao thức. Nó nằm ở định danh thực thể, chất lượng dữ liệu và quy trình cập nhật, tức những thứ giữ nguyên giá trị dù giao thức nào thắng. Chờ đợi hợp lý ở tầng trên và tốn kém ở tầng dưới.

## 7.6. BẢNG ĐIỀU KHIỂN KHẢ HÀNH ĐỘNG

Nếu chương này đến đây đã trả lời câu hỏi một địa phương phải được thiết kế như thế nào để máy có thể hiểu, suy luận và giao dịch, thì mục cuối cùng phải trả lời câu hỏi tiếp theo: "Quản trị địa phương như thế nào trong nền kinh tế tác nhân?".

Trong nhiều thập niên, các địa phương được điều hành bằng những bảng chỉ số quen thuộc, số lượt khách, doanh thu du lịch, công suất phòng, thời gian lưu trú, lượng tìm kiếm hay mức độ hài lòng của du khách. Những chỉ số này vẫn quan trọng, nhưng chúng đều phản ánh kết quả sau khi quyết định đã được đưa ra. Khi khách đã đến, tiền đã được chi và báo cáo đã được tổng hợp, mọi cơ hội can thiệp phần lớn đã trôi qua. Trong nền kinh tế tác nhân, quyết định quan trọng nhất xảy ra sớm hơn rất nhiều. Trước khi có lượt khách, đã có một lượt truy hồi. Trước khi có giao dịch, đã có một lần được đưa vào tập lựa chọn của tác nhân. Trước khi có doanh thu, đã có một mô hình máy đánh giá rằng địa phương đủ đáng tin để thực hiện hành động. Điều đó làm thay đổi bản chất của quản trị. Người quản lý không chỉ quan sát dòng khách, mà phải quan sát dòng quyết định của máy. Không chỉ theo dõi kết quả kinh doanh, mà còn phải theo dõi những điều kiện tạo ra kết quả ấy.

Bảng điều khiển của một địa phương vì vậy không còn là bảng tổng hợp các KPI truyền thống. Nó trở thành bảng điều khiển khả hành động, nơi phản ánh mức độ sẵn sàng của toàn bộ hệ sinh thái dữ liệu, thực thể, quan hệ, tồn kho và dịch vụ để các tác nhân có thể hiểu, tin tưởng và giao dịch.

Một địa phương có thể tăng lượng khách trong ngắn hạn nhờ chiến dịch quảng bá, nhưng nếu thực thể vẫn rời rạc, dữ

liệu không đồng bộ, tồn kho không thể truy cập theo thời gian thực và dịch vụ không thể đặt tự động, thì lợi thế đó sẽ nhanh chóng biến mất khi các tác nhân trở thành cổng phân phối chủ yếu. Ngược lại, một địa phương đầu tư đúng vào hạ tầng khả hành động có thể chưa nhìn thấy ngay lượng khách tăng mạnh, nhưng đang xây dựng nền móng để được máy lựa chọn lặp đi lặp lại trong tương lai.

Bởi vậy, bảng điều khiển mới không nhằm trả lời câu hỏi "Hôm nay có bao nhiêu khách?" mà nhằm trả lời những câu hỏi quan trọng hơn:

- *Máy có đọc được địa phương hay không?*
- *Máy có hiểu đúng địa phương hay không?*
- *Máy có đủ niềm tin để đề xuất hay không?*
- *Máy có thể hoàn tất giao dịch hay không?*
- *Sau mỗi giao dịch, hệ thống có học được điều gì để lần sau làm tốt hơn hay không?*

Đó là những chỉ số đi trước doanh thu. Cũng giống như trong nền kinh tế công nghiệp, người quản lý theo dõi điện áp, nhiệt độ và áp suất để ngăn dây chuyền ngừng hoạt động trước khi sản phẩm lỗi xuất hiện, thì trong nền kinh tế tác nhân, nhà quản trị địa phương phải theo dõi các tín hiệu về khả đọc, khả suy luận và khả giao dịch trước khi lượng khách hay doanh thu thay đổi.

Vì vậy, bảng điều khiển mới không phải là sự thay thế hoàn toàn cho KPI truyền thống. Nó là một lớp quản trị cao hơn, lớp quản trị những điều kiện quyết định liệu địa phương có còn hiện diện trong chuỗi giá trị của nền kinh tế tác nhân hay không. Đây cũng là bước chuyển cuối cùng của chương: từ việc thiết kế địa phương cho máy sang điều hành địa phương cùng với máy.

### 7.6.1. Bảng điều khiển kết quả khác bảng điều khiển khả hành động

Trong suốt lịch sử quản trị hiện đại, bảng điều khiển được thiết kế để phục vụ con người. Nhà quản lý nhìn vào màn hình để biết doanh thu hôm nay là bao nhiêu, công suất phòng đạt bao nhiêu phần trăm, lượng khách tăng hay giảm, mức độ hài lòng của du khách thay đổi ra sao. Bảng điều khiển là công cụ giúp con người quan sát, phân tích và ra quyết định.

Nhưng trong nền kinh tế tác nhân, một loại bảng điều khiển thứ hai xuất hiện. Nó không đo kết quả kinh doanh mà đo mức độ mà máy có thể sử dụng địa phương. Gọi tắt là bảng điều khiển khả hành động. Bảng thứ nhất phục vụ việc quan sát kết quả, bảng thứ hai phục vụ việc bảo đảm điều kiện tạo ra kết quả. Đây là một khác biệt mang tính bản chất.

Con người có thể nhìn vào một báo cáo PDF dày hàng trăm trang, đọc một bài viết quảng bá hay xem một video giới thiệu điểm đến rồi tự hình thành nhận định. Máy thì không. Một tác nhân không thể đặt phòng từ một brochure đẹp, cũng không thể hoàn tất hành trình chỉ vì một video hấp dẫn. Điều nó cần là những tín hiệu có cấu trúc cho phép thực hiện từng bước của quá trình ra quyết định. Vì vậy, bảng điều khiển dành cho máy không phải là tập hợp các biểu đồ trực quan mà là tập hợp các trạng thái vận hành có thể đọc được bằng chương trình.

Thay vì hỏi:

- *Hôm nay có bao nhiêu khách?*
- *Doanh thu tháng này tăng bao nhiêu?*
- *Bao nhiêu lượt xem trên mạng xã hội?*

Máy sẽ liên tục kiểm tra những câu hỏi khác:

- *Thực thể này còn tồn tại không?*
- *Dịch vụ này còn khả dụng không?*
- *Giá hiện tại là bao nhiêu?*
- *Khung giờ nào còn trống?*
- *API có phản hồi không?*
- *Mức độ tin cậy của dữ liệu là bao nhiêu?*
- *Giao dịch có thể hoàn tất ngay lúc này không?*

Đó là những thông tin mà bảng điều khiển truyền thống gần như không bao giờ theo dõi.

Có thể hình dung hai phòng điều hành khác nhau. Một bên hiển thị doanh thu, biểu đồ tăng trưởng và bản đồ nguồn khách. Bên kia hiển thị số thực thể hợp lệ, tỷ lệ API còn hoạt động, độ đầy đủ của siêu dữ liệu, số dịch vụ còn tồn kho theo thời gian thực, tỷ lệ giao dịch thành công và thời gian phản hồi của từng hệ thống. Một bên quan sát kết quả. Bên kia quan sát điều kiện tạo ra kết quả.

Đó cũng là lý do các chỉ số kết quả vốn đã là chỉ số đi sau và trong nền kinh tế tác nhân khoảng trễ ấy có thể lớn hơn vì nguyên nhân phát sinh ở tầng dữ liệu và khả hành động trước khi biểu hiện thành sụt giảm doanh thu (lagging indicator). Khi doanh thu giảm, lượng khách sụt hay công suất phòng thấp đi, vấn đề thực sự thường đã xảy ra từ trước: Tác nhân không còn đọc đúng dữ liệu, không còn truy cập được tồn kho, không còn đủ niềm tin để đề xuất hoặc không thể hoàn tất giao dịch.

Ngược lại, bảng điều khiển của máy tập trung vào các chỉ số dẫn dắt (leading indicator). Nếu tỷ lệ thực thể được xác minh

tăng lên, độ phủ đồ thị tri thức đầy đủ hơn, API phản hồi ổn định hơn và tồn kho luôn được cập nhật theo thời gian thực, xác suất địa phương được tác nhân lựa chọn sẽ tăng lên trước khi bất kỳ con số doanh thu nào thay đổi. Điều này buộc các nhà quản trị địa phương phải thay đổi cách nhìn về bảng điều khiển. Bảng điều khiển không còn chỉ là công cụ báo cáo thành tích sau mỗi tháng hay mỗi quý. Nó trở thành hệ thống giám sát liên tục mức độ mức sẵn sàng cho máy của toàn bộ địa phương.

Cuối cùng, có thể xem hai loại bảng điều khiển như hai lớp quản trị bổ sung cho nhau. Bảng điều khiển của con người trả lời câu hỏi: địa phương đang hoạt động như thế nào? Bảng điều khiển của máy trả lời câu hỏi quan trọng hơn trong nền kinh tế tác nhân: địa phương có còn đủ điều kiện để được máy lựa chọn và giao dịch hay không?

Trong nhiều năm tới, những địa phương cạnh tranh hiệu quả nhất sẽ không phải là nơi có nhiều biểu đồ đẹp nhất trên màn hình điều hành, mà là nơi có khả năng duy trì một bảng điều khiển mà cả con người lẫn tác nhân đều có thể sử dụng. Khi đó, bảng điều khiển không chỉ phản ánh hoạt động của địa phương mà trở thành một phần của chính hạ tầng vận hành, nơi mỗi thay đổi trong dữ liệu đều có thể ngay lập tức chuyển hóa thành khả năng được máy đọc, suy luận và hành động.

### 7.6.2. Năm tín hiệu phải theo dõi hằng ngày

Trong quản trị truyền thống, báo cáo thường được lập theo ngày, tuần hoặc tháng. Điều đó đủ khi môi trường kinh doanh thay đổi tương đối chậm. Nhưng trong nền kinh tế tác nhân, tác nhân có thể đưa ra hàng triệu quyết định mỗi giờ. Điều quyết định việc

một địa phương được lựa chọn hay bị bỏ qua không phải là báo cáo cuối tháng, mà là trạng thái dữ liệu tại đúng thời điểm tác nhân truy cập.

Vì vậy, nhà quản trị không thể chỉ theo dõi doanh thu hay lượng khách. Họ phải theo dõi liên tục những tín hiệu phản ánh "sức khỏe số" của toàn bộ hệ sinh thái. Đây là những chỉ số không trực tiếp tạo ra doanh thu, nhưng lại quyết định liệu doanh thu có cơ hội xuất hiện hay không. Độ tươi - Dữ liệu còn mới đến đâu. Đối với tác nhân, dữ liệu cũ đồng nghĩa với dữ liệu kém tin cậy.

***Tín hiệu thứ nhất là độ tươi.*** Một nhà hàng đã đổi giờ mở cửa nhưng chưa cập nhật, một khách sạn đã hết phòng nhưng hệ thống vẫn báo còn, một lễ hội đã kết thúc nhưng vẫn nằm trong lịch sự kiện. Chỉ vài sai lệch như vậy cũng đủ hạ mức tin cậy của cả địa phương trong mắt máy. Độ tươi không đo dữ liệu được tạo ra khi nào mà đo khoảng cách giữa thực tế và bản ghi. Khoảng cách càng nhỏ, tác nhân càng dám hành động.

***Tín hiệu thứ hai là khả dụng***, tức khả năng biến thông tin thành giao dịch. Máy không chỉ muốn biết dịch vụ tồn tại mà muốn biết còn bao nhiêu phòng, bao nhiêu chỗ, bao nhiêu bàn, khung giờ nào trống. Không có trạng thái khả dụng theo thời gian thực thì mọi kế hoạch đều thiếu cơ sở.

***Tín hiệu thứ ba là độ trễ*** và cần tách làm hai phép đo thường bị gộp làm một. Độ trễ cập nhật là thời gian để một thay đổi ngoài đời thật xuất hiện trên các nguồn máy đọc. Độ trễ phản hồi là thời gian hệ thống trả lời một truy vấn. Hai con số này có nguyên nhân khác nhau và được sửa bằng hai loại can thiệp khác nhau, nên không nên báo cáo chung.

Giữa hai dịch vụ tương đương, tác nhân có xu hướng ưu tiên dịch vụ phản hồi nhanh hơn, ổn định hơn và ít lỗi hơn. Khi hàng triệu quyết định được thực hiện mỗi ngày, chỉ vài trăm mili giây khác biệt cũng có thể tạo nên khoảng cách rất lớn về số lượng giao dịch. Độ phủ - Hệ sinh thái đã được số hóa đầy đủ chưa. Một tác nhân không đánh giá từng thực thể riêng lẻ. Nó đánh giá toàn bộ mạng lưới.

***Tín hiệu thứ tư là độ phủ.*** Một địa phương có hàng nghìn doanh nghiệp nhưng chỉ vài trăm đơn vị hiện diện trong đồ thị tri thức sẽ cho máy một bức tranh đứt đoạn: tuyến tham quan gãy giữa chừng, dịch vụ không nối được, hành trình không dựng nổi. Độ phủ vì vậy không đo dung lượng dữ liệu mà đo tỷ lệ tài sản thật đã bước vào thế giới số ở dạng đọc và giao dịch được.

***Tín hiệu thứ năm là tính nhất quán.*** Tác nhân không tin một nguồn mà so sánh nhiều nguồn. Khi trang chính thức ghi giờ đóng cửa là 22 giờ, bản đồ số ghi 21 giờ, nền tảng đặt chỗ ghi 23 giờ và mạng xã hội ghi tạm nghỉ, máy phải xử lý mâu thuẫn trước khi hành động và mỗi lần như vậy xác suất được chọn lại giảm.[1] Trong nền kinh tế tác nhân, sự thật không chỉ cần đúng. Nó còn phải đúng ở mọi nơi.

Năm tín hiệu này tạo thành bảng điều khiển cơ bản của một địa phương trong nền kinh tế tác nhân. Chúng không thay thế các KPI truyền thống như lượng khách hay doanh thu, mà đóng vai trò như những chỉ số dẫn dắt. Khi độ tươi được duy trì, khả

1. Expedia Group, thông cáo "Expedia Group B2B Introduces AI Toolkit and Platform for the Future of Travel Distribution", ngày 20 tháng 5 năm 2026, tại hội nghị đối tác Explore 26, https://ir.expediagroup.com/news-and-events/news/news-details/2026/. Con số 21 tỷ lượt gọi API mỗi ngày được chính thông cáo ghi rõ là dữ liệu nội bộ năm 2025.

dụng luôn chính xác, độ trễ thấp, độ phủ ngày càng đầy đủ và tính nhất quán được bảo đảm, xác suất địa phương được tác nhân đọc, tin tưởng và lựa chọn sẽ tăng lên một cách bền vững. Đến lúc đó, doanh thu không còn là kết quả của các chiến dịch quảng bá ngắn hạn, mà là hệ quả tự nhiên của một hạ tầng số được vận hành đúng cách.

### 7.6.3. Chỉ số suy giảm khả hành động: khi địa phương bắt đầu biến mất

Cần mở rộng khái niệm trước khi đi vào chi tiết. Mất khả đọc là dạng dễ nhận thấy nhất nhưng không phải dạng duy nhất. Một địa phương có thể vẫn được đọc mà bị hiểu sai, vẫn được hiểu mà mất khả năng giao dịch, hoặc vẫn còn dữ liệu mà dữ liệu đã hết hiệu lực. Bốn dạng suy giảm ấy đều dẫn tới cùng một kết quả kinh tế là bị loại khỏi tập lựa chọn, nên nhóm chỉ số dưới đây đo suy giảm khả hành động nói chung, trong đó mất khả đọc là trường hợp đầu tiên và thường xuất hiện sớm nhất.

Một địa phương không biến mất khỏi nền kinh tế tác nhân trong một ngày. Nó biến mất từng chút một. Không phải vì khách ngừng đến ngay lập tức, mà vì tác nhân bắt đầu ngừng nhìn thấy nó. Khi tần suất được truy hồi giảm, xác suất được đưa vào tập lựa chọn giảm theo. Khi số lần được đề xuất ít đi, lượng giao dịch cũng giảm. Cuối cùng, lượng dữ liệu mới, đánh giá mới và tín hiệu mới suy giảm, khiến hệ thống càng có ít lý do để lựa chọn địa phương đó trong tương lai. Đó là một vòng xoáy của mất khả đọc (loss of machine readability).

Điều nguy hiểm là quá trình này thường diễn ra âm thầm. Nhà quản lý vẫn nhìn thấy khách, doanh nghiệp vẫn hoạt động,

doanh thu chưa giảm đáng kể. Nhưng ở lớp hạ tầng thuật toán, địa phương đã bắt đầu mất vị trí trong trí nhớ của máy. Vì vậy, trong nền kinh tế tác nhân, cần xuất hiện một nhóm chỉ số hoàn toàn mới: chỉ số mất khả đọc. Đây không phải là chỉ số đo hiệu quả kinh doanh, mà đo mức độ suy giảm khả năng được máy đọc, hiểu và lựa chọn. Có nhiều tín hiệu báo trước hiện tượng này.

Đầu tiên là tỷ lệ thực thể lỗi tăng lên. Các địa điểm không còn tồn tại nhưng vẫn xuất hiện, doanh nghiệp đã đổi tên nhưng chưa cập nhật, liên kết hỏng, tọa độ sai hoặc dữ liệu bị thiếu. Khi số lượng thực thể lỗi vượt quá một ngưỡng nhất định, tác nhân sẽ giảm mức độ tin cậy đối với toàn bộ hệ sinh thái. Tiếp theo là độ tươi của dữ liệu giảm dần. Nếu hàng nghìn thực thể không được cập nhật trong nhiều tháng hoặc nhiều năm, máy sẽ coi hệ thống ít đáng tin hơn so với một địa phương có dữ liệu luôn được làm mới.

Một tín hiệu khác là sự suy giảm độ phủ của đồ thị tri thức. Các doanh nghiệp mới mở không được thêm vào đồ thị tri thức. Sự kiện mới không được kết nối. Tuyến du lịch mới không có quan hệ với các thực thể hiện hữu. Đồ thị bắt đầu xuất hiện những khoảng trống. Với con người, đó chỉ là vài thông tin thiếu. Với tác nhân, đó là những đoạn đường không thể đi tiếp trong quá trình suy luận.

Mức độ không nhất quán giữa các nguồn cũng là một chỉ báo quan trọng. Khi trang web chính thức, Google Maps, OTA, mạng xã hội và API liên tục đưa ra các thông tin khác nhau, tác nhân buộc phải giảm độ tin cậy. Ban đầu, địa phương chỉ bị giảm thứ hạng. Sau đó, nó có thể bị loại khỏi những tình huống mà tác nhân cần mức độ chắc chắn cao.

Một dấu hiệu tinh vi hơn là tần suất được truy hồi giảm. Nếu trước đây tác nhân thường xuyên tham chiếu đến một địa phương nhưng ngày càng ít làm như vậy, đó là tín hiệu cho thấy địa phương đang dần biến mất khỏi không gian nhận thức của máy. Đây có thể là chỉ số cảnh báo sớm quan trọng nhất, bởi nó xuất hiện trước khi lượng khách hay doanh thu giảm.

Có thể hình dung "mất khả đọc" giống như quá trình một con đường dần bị cỏ mọc phủ kín. Ban đầu xe vẫn đi được. Sau đó ít người sử dụng hơn. Cuối cùng con đường vẫn tồn tại trên bản đồ, nhưng không còn ai lựa chọn. Địa phương trong nền kinh tế tác nhân cũng vậy. Không phải biến mất về mặt vật lý, mà biến mất khỏi tuyến đường suy luận của tác nhân.

Điểm khác biệt lớn nhất của quản trị hiện đại là phải phát hiện quá trình này thật sớm. Khi doanh thu giảm, việc phục hồi có thể mất nhiều năm. Nhưng nếu phát hiện ngay từ khi các chỉ số khả đọc bắt đầu suy giảm, nhà quản lý có thể can thiệp bằng cách cập nhật dữ liệu, sửa các liên kết bị đứt, đồng bộ hóa nguồn thông tin và khôi phục niềm tin của máy trước khi hệ quả kinh tế xuất hiện. Nói cách khác, chỉ số mất khả đọc chính là hệ thống cảnh báo sớm của nền kinh tế tác nhân. Nó giúp địa phương biết mình đang dần biến mất khỏi tầm nhìn của tác nhân trước khi biến mất khỏi dòng khách và dòng tiền.

Suy giảm âm thầm. Khi nguyên nhân suy giảm nằm trong hạ tầng số, các chỉ số khả đọc và khả hành động thường xấu đi trước khi tác động xuất hiện đầy đủ trong doanh thu. Độ trễ giữa hai thời điểm ấy là dư địa can thiệp của nhà quản trị. Mệnh đề này không loại trừ những nguyên nhân nằm ngoài hạ tầng, như thiên tai, suy thoái hay biến động giá, khi doanh thu có thể giảm trước dù dữ liệu vẫn hoàn hảo. Khi doanh thu đã phản ánh vấn đề thì chi phí khôi phục đã cao hơn nhiều lần chi phí phòng ngừa.

Đo lường khả đọc sẽ đẻ ra bệnh thành tích mới. Phản bác này chưa bị bác bỏ. Bất kỳ chỉ số nào trở thành mục tiêu đều thôi làm một thước đo tốt. Có thể hình dung ngay cách chúng bị làm đẹp, sinh thực thể ảo để nâng độ phủ, chạm nhẹ bản ghi để làm mới độ tươi, dựng API trả về dữ liệu tĩnh để hạ độ trễ. Cách phòng vệ duy nhất tương đối bền là neo mọi chỉ số hạ tầng vào một chỉ số kết quả khó làm giả, đó là tỷ lệ chuỗi hành động hoàn tất được xác nhận bởi giao dịch thật.

### 7.6.4. Điều hành bằng tín hiệu, không bằng chiến dịch

Trong nhiều thập niên, quản trị điểm đến được tổ chức theo tư duy chiến dịch. Mỗi năm có một chương trình quảng bá. Mỗi mùa du lịch có một lễ hội. Mỗi khi lượng khách giảm lại triển khai một đợt truyền thông mới. Hoạt động quản lý vì thế diễn ra theo từng đợt cao điểm, xen kẽ những khoảng thời gian dài tương đối yên tĩnh. Nền kinh tế tác nhân làm mô hình đó trở nên lỗi thời.

Tác nhân không chờ đến mùa quảng bá mới đưa ra quyết định. Chúng đọc dữ liệu, đánh giá dịch vụ và thực hiện giao dịch liên tục, từng giây trong ngày. Điều này đồng nghĩa với việc địa phương không còn cạnh tranh bằng những chiến dịch ngắn hạn, mà bằng chất lượng của các tín hiệu được phát ra liên tục. Một chiến dịch có thể tạo ra sự chú ý trong vài tuần. Một tín hiệu đúng có thể ảnh hưởng đến hàng triệu quyết định mỗi ngày.

Khác biệt này rất quan trọng. Chiến dịch là những nỗ lực mang tính gián đoạn. Tín hiệu là dòng dữ liệu liên tục mà tác nhân sử dụng để suy luận. Nếu chiến dịch giống như việc treo một tấm biển quảng cáo lớn, thì tín hiệu giống như việc duy trì nhịp tim

của toàn bộ hệ thống. Tấm biển có thể thu hút sự chú ý trong chốc lát. Nhịp tim quyết định hệ thống còn sống hay không.

Điều hành bằng tín hiệu có nghĩa là mọi thay đổi trong thế giới thực phải nhanh chóng được phản ánh trong thế giới số. Một nhà hàng thay đổi thực đơn. Một khách sạn mở thêm hạng phòng. Một bảo tàng kéo dài giờ mở cửa. Một sự kiện mới được bổ sung. Một tuyến xe buýt được điều chỉnh. Tất cả những thay đổi này phải lập tức trở thành tín hiệu mà tác nhân có thể đọc.

Nhà quản trị vì vậy không còn chỉ lập kế hoạch truyền thông. Họ phải xây dựng một hệ thống vận hành tín hiệu. Hệ thống đó liên tục giám sát độ tươi của dữ liệu, đồng bộ các nguồn, cập nhật tồn kho, kiểm tra API, phát hiện liên kết bị đứt và sửa lỗi trước khi tác nhân gặp phải. Đây cũng là sự chuyển đổi từ quản trị theo sự kiện sang quản trị theo trạng thái.

Trong mô hình cũ, nhà quản lý phản ứng khi có vấn đề: khách giảm, doanh thu giảm hoặc truyền thông tiêu cực. Trong mô hình mới, nhà quản lý theo dõi trạng thái của hệ thống theo thời gian thực. Khi độ trễ tăng, dữ liệu cũ đi, tỷ lệ lỗi API vượt ngưỡng hay mức độ nhất quán suy giảm, hệ thống đã phát tín hiệu cần can thiệp, dù doanh thu lúc đó vẫn chưa thay đổi.

Điều này khiến công việc của nhà quản trị địa phương ngày càng giống người điều hành một mạng lưới hạ tầng hơn là người tổ chức các hoạt động quảng bá. Mục tiêu không phải tạo ra những đợt bùng nổ ngắn hạn, mà duy trì một dòng tín hiệu ổn định, chính xác và đáng tin cậy để tác nhân luôn có đủ cơ sở lựa chọn địa phương.

Đó cũng là sự thay đổi sâu sắc nhất của quản trị trong nền kinh tế tác nhân. Chiến dịch tạo ra sự chú ý. Tín hiệu tạo ra quyết

định. Khi quyền lựa chọn chuyển sang máy, lợi thế cạnh tranh bền vững sẽ thuộc về những địa phương có khả năng phát đi các tín hiệu chất lượng cao mỗi ngày, chứ không phải những địa phương chỉ xuất hiện rực rỡ trong vài chiến dịch truyền thông mỗi năm.

Điều này không có nghĩa chiến dịch hết vai trò. Chiến dịch vẫn cần để tạo ra nhu cầu và để định hình mong muốn, những việc mà một dòng dữ liệu sạch không làm được. Nhưng tín hiệu quyết định nhu cầu ấy có được máy chuyển thành hành động hay không. Một địa phương chi mạnh cho chiến dịch mà bỏ tín hiệu sẽ tạo ra nhu cầu rồi trao nó cho nơi khác thực hiện.

Không phê duyệt một cổng thông tin mới khi lớp định danh thực thể chưa tồn tại. Giao diện thứ năm cho cùng một tập dữ liệu rời rạc không làm tăng khả năng được máy chọn, chỉ làm tăng số nguồn mâu thuẫn. Không công bố dữ liệu mà thiếu ba trường bắt buộc: thời điểm cập nhật gần nhất, chủ thể chịu trách nhiệm và mức độ tin cậy. Dữ liệu không có ba trường này là dữ liệu mà máy buộc phải chiết khấu. Không để hai cơ quan cùng phát hành hai bản sự thật về cùng một thực thể. Với mỗi loại thông tin, phải chỉ định một nguồn duy nhất được coi là gốc, các nguồn còn lại dẫn chiếu về đó. Không tính số bản ghi đã số hóa là kết quả. Chỉ tính số chuỗi hành động mà tác nhân hoàn tất được từ đầu đến cuối. Chiến dịch tạo ra sự chú ý. Tín hiệu tạo ra quyết định.

## 7.7. KIỂM TOÁN ĐỊA PHƯƠNG BẰNG CHÍNH HỆ THỐNG SUY LUẬN

Nếu ở các mục trước, câu hỏi của quản trị địa phương là làm thế nào để được nhìn thấy, thì đến cuối Chương 7 câu hỏi đã thay đổi: các hệ thống suy luận đang mô tả, liên hệ và hành động trên địa phương như thế nào?

Trong nhiều thập niên, thương hiệu địa phương được đo bằng khảo sát nhận biết, lượng khách, số bài báo, mức độ xuất hiện trên mạng xã hội hay giá trị truyền thông quy đổi (AVE). Những chỉ số này vẫn còn giá trị, nhưng chúng chỉ phản ánh điều con người nhìn thấy. Trong nền kinh tế tác nhân, còn tồn tại một lớp thương hiệu khác: thương hiệu tồn tại trong quá trình suy luận của máy.

Một địa phương có thể liên tục xuất hiện trên báo chí nhưng vẫn không được AI đề xuất. Ngược lại, một địa phương ít truyền thông hơn nhưng có dữ liệu đầy đủ, quan hệ thực thể rõ ràng, tồn kho số luôn cập nhật và khả năng giao dịch cao lại thường xuyên được các tác nhân lựa chọn. Điều này cho thấy thương hiệu không còn chỉ nằm trong nhận thức của con người mà còn tồn tại dưới dạng mô hình xác suất trong các hệ thống suy luận.

Vì vậy, đến cuối hành trình thiết kế địa phương cho máy, công việc quan trọng nhất không phải tiếp tục quảng bá mà là kiểm toán xem các hệ thống đang biểu diễn gì, dựa vào nguồn nào và làm được gì với địa phương đó.

Đây là một thay đổi lớn về đối tượng đánh giá. Trước đây, thương hiệu được đo bằng những gì địa phương tự nói về mình. Nay nó được đo bằng những gì hàng loạt hệ thống tự kết luận sau

khi đọc toàn bộ dấu vết số của địa phương. Một cuộc kiểm toán như vậy thường bắt đầu bằng những câu hỏi đơn giản. Nếu được yêu cầu đề xuất nơi nghỉ cuối tuần, địa phương có nằm trong tập lựa chọn không? Nếu người dùng hỏi về ẩm thực, di sản hay du lịch sức khỏe, hệ thống có liên hệ đúng chủ đề không? Nếu phải lập hành trình, dữ liệu có đủ để dựng một kế hoạch trọn vẹn không? Nếu cần đặt phòng hay thanh toán, chuỗi hành động dừng ở bước nào?

**MÔ HÌNH ĐỊA PHƯƠNG TRƯỚC MÁY VÀ THƯƠNG HIỆU TRƯỚC MÁY**

*Cần tách hai khái niệm thường hay bị gộp lẫn. Mô hình địa phương trước máy là toàn bộ biểu diễn mà tác nhân dùng để nhận diện, đánh giá và hành động, gồm thực thể, quan hệ, tồn kho, tín hiệu chất lượng và khả năng giao dịch. Thương hiệu trước máy hẹp hơn: đó là phần liên tưởng, thuộc tính nổi trội và vị trí so sánh mà tác nhân suy ra từ mô hình ấy. Mục 7.7 kiểm toán mô hình tổng thể, còn tiểu mục 7.7.3 kiểm toán riêng phần thương hiệu và bản sắc. Cách phân biệt này giữ cho khái niệm thương hiệu ở Chương 7 vẫn cùng nghĩa với các chương trước.*

Nếu một thông tin sai xuất hiện, hệ thống có đủ bằng chứng để tự sửa hay tiếp tục củng cố sai lệch đó?

Những câu hỏi này không đánh giá hình ảnh bề mặt mà đánh giá khả năng tồn tại của địa phương trong quá trình ra quyết định của máy. Quan trọng hơn, kiểm toán không dừng ở kết quả cuối cùng mà phải truy ngược cả chuỗi suy luận.

- *Địa phương được đề xuất vì tín hiệu nào?*
- *Nguồn dữ liệu nào được AI tin cậy nhất?*
- *Quan hệ thực thể nào đang thiếu?*
- *Những mắt xích nào khiến lập luận bị đứt?*
- *Khả năng giao dịch dừng lại ở bước nào?*

Nếu biết được câu trả lời cho những câu hỏi này, nhà quản trị không còn phải đoán vì sao mình không được lựa chọn. Họ có thể nhìn thấy chính xác điểm nghẽn nằm ở đâu trong chuỗi suy luận của máy. Đây cũng là lý do các bảng điều khiển ở mục trước tập trung vào độ tươi, khả dụng, độ phủ, tính nhất quán và độ trễ. Những tín hiệu đó không chỉ giúp hệ thống vận hành tốt hơn mà còn là những chỉ báo trực tiếp về chất lượng của mô hình suy luận mà AI đang hình thành. Từ đó xuất hiện một chức năng quản trị mới. Nhiệm vụ của nó không phải sáng tạo khẩu hiệu hay mua quảng cáo mà là định kỳ truy vấn nhiều hệ thống, ghi lại câu trả lời, phân tích nguồn được dẫn, xác định điểm nghẽn và chuyển kết quả thành yêu cầu sửa hạ tầng dữ liệu.

Trong tương lai gần, nhiều khả năng các địa phương sẽ không còn hỏi "năm nay thương hiệu mạnh hơn bao nhiêu?" mà sẽ hỏi những câu rất khác:

- *Bao nhiêu tác nhân có thể hoàn thành một giao dịch tại địa phương?*
- *Bao nhiêu mô hình AI hiểu đúng bản sắc của địa phương?*
- *Bao nhiêu truy vấn mà địa phương được đưa vào tập lựa chọn?*
- *Bao nhiêu lần AI từ bỏ giao dịch vì thiếu dữ liệu hoặc thiếu khả năng thực hiện?*
- *Mất bao lâu để một thay đổi ngoài đời thực được phản ánh trong toàn bộ hệ sinh thái AI?*

Những câu hỏi ấy phản ánh đúng bản chất của thương hiệu trước máy, không phải điều địa phương nói về mình, mà là mô hình mà hàng loạt hệ thống tự dựng nên về địa phương.

### 7.7.1. Máy đang mô tả địa phương như thế nào

Trong nhiều thập niên, hoạt động quản trị thương hiệu địa phương bắt đầu bằng một câu hỏi quen thuộc: "Chúng ta muốn được biết đến như thế nào?" Từ câu hỏi đó, các địa phương xây dựng định vị, sáng tạo khẩu hiệu, thiết kế logo, sản xuất video quảng bá và triển khai các chiến dịch truyền thông với kỳ vọng hình ảnh mong muốn sẽ được truyền tải đến công chúng. Trong nền kinh tế tác nhân, câu hỏi này không còn đủ nữa.

Một địa phương có thể dành nhiều nguồn lực để nói rằng mình là "điểm đến xanh", "thành phố sáng tạo", "thủ phủ du lịch sức khỏe" hay "thiên đường nghỉ dưỡng", nhưng điều quyết định không phải là địa phương tự mô tả ra sao. Điều quyết định là các hệ thống AI đã thực sự hiểu địa phương theo cách đó hay chưa.

Giữa điều địa phương nói và điều AI kết luận luôn tồn tại một khoảng cách. Khoảng cách ấy được tạo ra bởi hàng triệu tín hiệu mà AI thu thập từ nhiều nguồn khác nhau: dữ liệu mở, trang web chính thức, báo chí, bản đồ số, đánh giá của người dùng, lịch sự kiện, hình ảnh, video, dữ liệu giao dịch và vô số dấu vết số khác. AI không tin vào một tuyên bố đơn lẻ. Nó xây dựng nhận thức thông qua việc tổng hợp và đối chiếu tất cả những bằng chứng có sẵn.

Vì vậy, bài toán thương hiệu không còn là Brand Messaging mà chuyển thành Brand Interpretation. Điều quan trọng không phải thông điệp được phát đi, mà là thông điệp được diễn giải như thế nào trong quá trình suy luận của máy.

Một địa phương có thể tự nhận là trung tâm du lịch sức khỏe, nhưng nếu AI vẫn liên tục mô tả nơi đó như một điểm nghỉ dưỡng biển thông thường thì thương hiệu mong muốn gần như chưa tồn tại trong hệ sinh thái số. Ngược lại, có những địa phương chưa bao giờ chính thức tuyên bố mình là trung tâm ẩm thực, nhưng AI lại luôn đề xuất nơi đó khi người dùng hỏi về trải nghiệm ăn uống. Trong trường hợp này, thương hiệu số mạnh hơn rất nhiều so với khẩu hiệu chính thức.

Đó là lý do câu hỏi đầu tiên của nhà quản trị không nên là:

- *"Chúng ta đang truyền thông điều gì?"*

Mà phải là:

- *"Nếu hôm nay hàng triệu người hỏi các trợ lý số về lĩnh vực này, chúng sẽ mô tả địa phương của chúng ta ra sao, dẫn nguồn nào và làm được đến bước nào?"*

Cần một lưu ý về cách nói. Cụm máy nghĩ gì là một phép nói tắt tiện dụng nhưng không chính xác. Các hệ thống không có một suy nghĩ thống nhất về địa phương. Kết quả chúng đưa ra thay đổi theo mô hình, thời điểm, cách đặt câu hỏi, công cụ truy hồi được bật, vị trí địa lý và hồ sơ người dùng. Điều có thể quan sát và đo được là hành vi đầu ra: chúng mô tả bằng khái niệm nào, dẫn nguồn nào và dừng ở bước nào. Toàn bộ phương pháp kiểm toán dưới đây được xây dựng trên hành vi đầu ra ấy, không phải trên một trạng thái tinh thần giả định của máy. Đây là sự đảo chiều hoàn toàn trong tư duy quản trị thương hiệu. Trước đây, địa phương chủ động kể câu chuyện của mình rồi hy vọng người khác tiếp nhận. Bây giờ, địa phương phải thường xuyên lắng nghe câu chuyện mà các hệ thống AI đang tự kể về mình.

Quản trị thương hiệu vì thế trở thành quá trình đối chiếu liên tục giữa bản sắc mong muốn và mô hình suy luận thực tế. Mỗi khoảng cách giữa hai mô hình đều là một vấn đề quản trị cần được xử lý bằng dữ liệu, bằng quan hệ thực thể, bằng tín hiệu vận hành và bằng khả năng giao dịch, chứ không chỉ bằng thêm một chiến dịch truyền thông. Trong nền kinh tế tác nhân, thương hiệu không còn nằm trong bộ nhận diện hay slogan. Thương hiệu tồn tại trong câu trả lời mà AI tạo ra khi không có mặt chúng ta.

### 7.7.2. Kiểm toán đa tác nhân

Một sai lầm phổ biến khi đánh giá khả năng hiện diện của địa phương trong AI là chỉ thử hỏi một hệ thống duy nhất rồi coi đó là kết luận cuối cùng. Thực tế, không tồn tại một AI đại diện cho toàn bộ nền kinh tế tác nhân.

Mỗi hệ thống được huấn luyện trên những tập dữ liệu khác nhau, sử dụng những phương pháp truy hồi khác nhau, có mục tiêu tối ưu khác nhau và phục vụ những nhóm người dùng khác nhau. Một trợ lý du lịch, một OTA, một bản đồ số, một mô hình ngôn ngữ hay một siêu ứng dụng đều có thể đưa ra những kết luận không hoàn toàn giống nhau về cùng một địa phương. Bởi vậy, kiểm toán thương hiệu trong thời đại AI không thể thực hiện với một tác nhân duy nhất. Nó phải được tiến hành đồng thời trên nhiều hệ thống để quan sát điểm giao nhau và điểm khác biệt trong các mô hình suy luận. Nguyên tắc đầu tiên là so sánh.

Đặt cùng một câu hỏi cho nhiều tác nhân khác nhau:

- *Đâu là điểm đến phù hợp cho nghỉ dưỡng cuối tuần?*
- *Nơi nào nổi bật về du lịch sức khỏe?*
- *Thành phố nào thích hợp cho gia đình có trẻ nhỏ?*
- *Đâu là điểm đến văn hóa hấp dẫn ở Việt Nam?*

Nếu một địa phương liên tục xuất hiện trong câu trả lời của nhiều tác nhân, đó là dấu hiệu cho thấy thương hiệu đã được hình thành khá ổn định trong hệ sinh thái AI. Nếu chỉ một vài tác nhân đề xuất còn phần lớn bỏ qua, thương hiệu vẫn còn phân mảnh. Nguyên tắc thứ hai là đối chiếu. Không chỉ quan sát địa phương có xuất hiện hay không, mà còn phải phân tích AI mô tả địa phương bằng những khái niệm nào.

- *Các tác nhân có đang dùng cùng một tập từ khóa?*
- *Có cùng nhấn mạnh bản sắc?*
- *Có cùng nhận diện lợi thế cạnh tranh?*
- *Có liên hệ đúng với những thực thể quan trọng khác?*
- *Hay mỗi hệ thống đang xây dựng một hình ảnh hoàn toàn khác nhau?*

Sự khác biệt giữa các mô hình suy luận thường hé lộ những khoảng trống trong dữ liệu hoặc những tín hiệu chưa đủ mạnh để tạo ra nhận thức thống nhất. Nguyên tắc thứ ba là lặp lại. AI không phải là một hệ thống tĩnh. Dữ liệu mới liên tục xuất hiện, thuật toán liên tục thay đổi và các mô hình cũng liên tục được cập nhật. Một cuộc kiểm toán thực hiện hôm nay chỉ phản ánh trạng thái của thời điểm hiện tại. Do đó, việc đánh giá cần được thực hiện định kỳ với cùng một bộ câu hỏi, cùng một nhóm tác nhân và cùng một phương pháp ghi nhận kết quả. Chỉ khi lặp lại theo thời gian, nhà quản trị mới quan sát được xu hướng: địa

phương đang được hiểu đúng hơn hay sai lệch hơn, khả năng được đề xuất đang tăng hay giảm, những thay đổi về dữ liệu có thực sự cải thiện mô hình suy luận của AI hay không.

Vì kết quả nhạy cảm với điều kiện, một vòng kiểm toán chỉ có giá trị so sánh khi các điều kiện được giữ cố định. Tối thiểu cần chuẩn hóa sáu yếu tố: cùng một câu hỏi viết nguyên văn, cùng khung thời gian, cùng thiết lập vị trí, cùng loại tài khoản và mức độ cá nhân hóa, cùng trạng thái bật hay tắt công cụ tìm kiếm và lặp lại mỗi câu hỏi nhiều lần để tách nhiễu ngẫu nhiên khỏi khác biệt thật. Kết quả phải ghi lại cả nguồn được trích dẫn. Quan trọng nhất là tách hai thứ thường bị trộn lẫn: câu trả lời mà hệ thống đưa ra và khả năng nó thực sự hoàn thành nhiệm vụ. Một địa phương có thể được mô tả rất đẹp trong câu trả lời mà vẫn thất bại ở bước đặt dịch vụ.

Điều đáng chú ý là mục tiêu của kiểm toán không phải tìm ra AI nào đúng nhất, mà là tìm ra những tín hiệu được nhiều hệ thống cùng xác nhận. Khi nhiều tác nhân đi đến một kết luận tương tự, có thể xem đó là mô hình nhận thức đang dần ổn định. Nhưng đồng thuận chỉ có sức nặng khi các tác nhân thực sự khác nhau về nhà cung cấp mô hình, nguồn truy hồi và cơ chế công cụ. Nếu nhiều hệ thống cùng phụ thuộc một nguồn dữ liệu nền, sự giống nhau có thể chỉ phản ánh một điểm nghẽn chung chứ không phải nhiều bằng chứng độc lập. Vì vậy, kiểm toán thương hiệu trong nền kinh tế tác nhân không còn giống một cuộc khảo sát thị trường được thực hiện mỗi năm một lần. Nó giống một quy trình quan sát liên tục các hệ thống suy luận, nơi mỗi tác nhân đóng vai trò như một "kiểm toán viên" độc lập. Càng nhiều tác nhân được đưa vào quá trình đánh giá, càng nhiều lần đối chiếu và lặp lại được thực hiện, bức tranh về thương hiệu số của địa phương càng trở nên chính xác, đầy đủ và có giá trị cho việc điều hành.

### 7.7.3. Phát hiện khoảng cách bản sắc

Việc chẩn đoán cần ba đỉnh chứ không phải hai. Đỉnh thứ nhất là bản sắc mà địa phương tự nhận và muốn giữ. Đỉnh thứ hai là bằng chứng thực tế, tức là những gì thị trường và cộng đồng thực sự trải nghiệm. Đỉnh thứ ba là bản sắc mà hệ thống có đủ dữ liệu để suy ra. Nếu chỉ so đỉnh thứ nhất với đỉnh thứ ba, địa phương có nguy cơ đi sửa dữ liệu để máy nói đúng điều mình muốn, trong khi trải nghiệm thật chưa hỗ trợ định vị ấy. Khi đó khoảng cách được che đi chứ không được thu hẹp.

Mọi địa phương đều có một bản sắc mà mình muốn xây dựng. Đó có thể là một thành phố di sản, một trung tâm sáng tạo, một điểm đến nghỉ dưỡng sức khỏe hay một vùng đất của văn hóa bản địa. Bản sắc này được thể hiện trong chiến lược phát triển, trong quy hoạch, trong truyền thông và trong cách chính quyền cùng cộng đồng mong muốn được nhận diện. Nhưng trong nền kinh tế tác nhân, tồn tại một bản sắc khác quan trọng không kém: bản sắc mà AI đang thực sự gán cho địa phương. Hai bản sắc này không phải lúc nào cũng trùng khớp.

Một địa phương có thể đầu tư nhiều năm để xây dựng hình ảnh "thành phố sáng tạo", nhưng phần lớn các hệ thống AI vẫn chỉ liên hệ nơi đó với công nghiệp nặng hoặc du lịch giá rẻ. Một vùng biển muốn trở thành điểm đến nghỉ dưỡng cao cấp nhưng AI vẫn liên tục mô tả bằng các từ khóa "đông đúc", "giá rẻ", "du lịch đại chúng". Một địa phương đã phát triển mạnh du lịch sức khỏe nhưng trong các câu trả lời của AI vẫn gần như không xuất hiện khi người dùng tìm kiếm sức khỏe toàn diện, spa hay chăm sóc sức khỏe. Khoảng cách giữa bản sắc mong muốn (desired identity) và bản sắc mà máy đang gán (machine-perceived identity) chính là khoảng cách bản sắc (identity gap). Đây là một trong những chỉ số quan trọng nhất của quản trị thương hiệu trong thời đại AI.

Điều đáng chú ý là khoảng cách bản sắc thường không xuất hiện vì AI "hiểu sai". Trong phần lớn trường hợp, AI chỉ đang phản ánh đúng những tín hiệu mà hệ sinh thái dữ liệu cung cấp. Nếu hàng nghìn nguồn dữ liệu vẫn mô tả địa phương theo hình ảnh cũ, nếu các thực thể mới chưa được kết nối, nếu các dịch vụ mới chưa có đủ bằng chứng số hoặc nếu dữ liệu vận hành chưa được cập nhật, AI sẽ tiếp tục duy trì mô hình nhận thức cũ. Nói cách khác, khoảng cách bản sắc là biểu hiện của khoảng cách trong hạ tầng tri thức, chứ không chỉ trong truyền thông. Chính vì vậy, việc phát hiện khoảng cách bản sắc không nên dừng ở việc ghi nhận rằng "AI đang mô tả chưa đúng". Quan trọng hơn là truy tìm nguyên nhân của khoảng cách ấy.

- *Khoảng cách xuất phát từ thiếu dữ liệu hay dữ liệu lỗi thời?*
- *Từ quan hệ thực thể chưa đầy đủ?*
- *Từ sự thiếu nhất quán giữa các nguồn?*
- *Hay từ việc địa phương đã thay đổi trong thực tế nhưng chưa để lại đủ dấu vết số?*

Mỗi khoảng cách đều tương ứng với một mắt xích cần được sửa trong hệ thống tri thức. Một địa phương có khoảng cách bản sắc nhỏ sẽ có sự thống nhất cao giữa điều mình là, điều mình muốn trở thành và điều AI đang hiểu. Ngược lại, khoảng cách bản sắc càng lớn thì chi phí để được lựa chọn trong nền kinh tế tác nhân càng cao, bởi AI luôn ra quyết định dựa trên mô hình nhận thức hiện có chứ không dựa trên những tuyên bố về tương lai. Do đó, mục tiêu cuối cùng của quản trị thương hiệu không còn là tạo ra một hình ảnh hấp dẫn hơn, mà là liên tục thu hẹp khoảng cách giữa bản sắc thực và bản sắc mà máy đang suy luận.

### 7.7.4. Từ kiểm toán sang sửa bằng chứng gốc

Trong mô hình quản trị truyền thống, khi thương hiệu yếu hoặc bị hiểu sai, phản ứng quen thuộc thường là triển khai thêm một chiến dịch truyền thông. Làm video mới, thiết kế khẩu hiệu mới, mua quảng cáo nhiều hơn, mời người nổi tiếng, tổ chức thêm sự kiện hay tăng ngân sách marketing được xem là giải pháp mặc định. Trong nền kinh tế tác nhân, đó thường không phải là điểm cần sửa. Nếu kết quả kiểm toán cho thấy AI liên tục hiểu sai một địa phương, vấn đề hiếm khi nằm ở việc quảng cáo chưa đủ hấp dẫn. Vấn đề nằm ở dữ liệu mà AI dùng để suy luận. Một đoạn quảng cáo đơn lẻ hiếm khi đủ sức thay đổi mô hình suy luận nếu các nguồn nền, dữ liệu vận hành và bằng chứng bên ngoài vẫn kể một câu chuyện khác. Mô hình chỉ dịch chuyển khi cả mạng lưới tri thức mà nó tham chiếu thay đổi theo hướng nhất quán và có đủ bằng chứng. Bởi vậy, sau mỗi cuộc kiểm toán, hành động đầu tiên không phải là họp với đơn vị truyền thông mà là rà soát toàn bộ hạ tầng dữ liệu. Nguyên tắc ở đây là ưu tiên chứ không phải loại trừ: sửa nguồn dữ liệu và bằng chứng gốc trước, sửa thông điệp sau. Có trường hợp lỗi thật sự nằm ở chỗ nội dung công khai quá nghèo nàn và khi ấy bổ sung nội dung vẫn là việc phải làm.

- *Thực thể nào còn thiếu?*
- *Thuộc tính nào chưa đầy đủ?*
- *Quan hệ nào chưa được khai báo?*
- *Nguồn chính thức nào chưa được cập nhật?*
- *Dịch vụ nào chưa có khả năng giao dịch?*
- *API nào chưa phản ánh dữ liệu thời gian thực?*
- *Đồ thị tri thức còn những mắt xích đứt gãy nào?*

Nếu AI chưa nhận diện địa phương là điểm đến du lịch sức khỏe, giải pháp không phải là thêm một khẩu hiệu "Wellness Destination". Giải pháp là xây dựng đầy đủ các thực thể về cơ sở chăm sóc sức khỏe, kết nối chúng với khách sạn, khu nghỉ dưỡng, bác sĩ, chuyên gia, chương trình trị liệu, lịch hoạt động, khả năng đặt lịch và dữ liệu vận hành theo thời gian thực. Khi cấu trúc tri thức thay đổi, mô hình suy luận của AI sẽ dần thay đổi theo.

Điều này phản ánh một sự chuyển dịch rất lớn trong đầu tư công. Ngân sách thương hiệu không còn chủ yếu dành cho việc sản xuất thông điệp, mà ngày càng dành cho việc xây dựng hạ tầng tri thức. Trang web chính thức, dữ liệu mở, đồ thị tri thức, API, danh mục dịch vụ, chuẩn định danh thực thể, khả năng cập nhật thời gian thực và khả năng giao dịch trở thành những tài sản thương hiệu quan trọng không kém logo hay chiến dịch quảng bá. Có thể xem quá trình này như một vòng lặp khép kín.

Đầu tiên, địa phương kiểm toán mô hình suy luận của AI để biết hệ thống đang hiểu gì. Tiếp theo, xác định các khoảng cách giữa bản sắc mong muốn và bản sắc được mô tả. Sau đó, không sửa câu chuyện mà sửa cấu trúc dữ liệu, mạng lưới quan hệ và khả năng hành động của toàn bộ hệ sinh thái số. Cuối cùng, tiếp tục kiểm toán để đo xem mô hình suy luận đã thay đổi hay chưa. Đó là một vòng lặp học hỏi liên tục giữa địa phương và các hệ thống trí tuệ nhân tạo. Nguyên tắc quản trị rút ra từ vòng lặp này rất ngắn, muốn thay đổi điều máy nói về mình, trước hết phải thay đổi điều máy có thể học từ mình.

### 7.7.5. Cơ chế liên thông tín hiệu địa phương

Cơ chế này có thể được tổ chức dưới hình thức một liên minh dữ liệu, nhưng không đòi hỏi các thành viên chuyển toàn bộ dữ liệu về một kho chung. Liên minh dữ liệu là một mô hình tổ chức, liên thông tín hiệu là chức năng cần đạt và trong phần lớn trường hợp chức năng ấy đạt được bằng chuẩn giao tiếp chung, xác nhận theo truy vấn và liên kết sang hệ thống của chính doanh nghiệp. Một hệ quả trực tiếp của toàn bộ chương là không một tổ chức nào có thể một mình xây dựng khả năng hành động của địa phương.

Trong kỷ nguyên truyền thông đại chúng, mỗi đơn vị có thể hoạt động tương đối độc lập và hệ thống vẫn chạy được dù dữ liệu phân tán. Trong nền kinh tế tác nhân thì không. Tác nhân không nhìn thấy từng tổ chức riêng lẻ mà nhìn thấy một hệ sinh thái có hoàn thành được nhiệm vụ hay không. Một khách sạn có dữ liệu hoàn hảo cũng không cứu được hành trình nếu giao thông công cộng không có API. Một điểm tham quan đã số hóa đầy đủ cũng không tối ưu được trải nghiệm nếu lịch sự kiện không cập nhật. Một nhà hàng đặt bàn được nhưng thanh toán không tích hợp thì giao dịch vẫn dừng. Trong một chuỗi hành động cụ thể, năng lực của địa phương bị giới hạn bởi mắt xích yếu nhất cần thiết để hoàn tất chuỗi ấy. Một mắt xích yếu không làm yếu mọi nhiệm vụ, chỉ làm yếu những nhiệm vụ phụ thuộc vào nó.

Đó là lý do mô hình cạnh tranh cũng thay đổi. Địa phương không còn cạnh tranh bằng từng doanh nghiệp riêng lẻ mà bằng khả năng phối hợp của toàn bộ hệ sinh thái. Liên minh mới ấy gồm sáu chủ thể nối với nhau qua một lõi chung, không xếp thành tầng trên tầng dưới. Chủ thể thứ nhất là chính quyền.

Vai trò của chính quyền không còn chủ yếu là quảng bá hình ảnh, mà là xây dựng các tiêu chuẩn chung: chuẩn thực thể, chuẩn định danh, chuẩn dữ liệu mở, quy tắc chia sẻ thông tin, cơ chế cập nhật và khung quản trị chất lượng dữ liệu. Chính quyền trở thành kiến trúc sư của hạ tầng số địa phương, bảo đảm mọi thành phần có thể kết nối với nhau. Chủ thể thứ hai là doanh nghiệp.

Doanh nghiệp không chỉ cung cấp dịch vụ cho con người mà còn phải cung cấp dịch vụ theo cách máy có thể hiểu và sử dụng. Mỗi khách sạn, nhà hàng, bảo tàng, bệnh viện, khu vui chơi hay đơn vị vận tải đều trở thành một nút trong mạng lưới tri thức của địa phương. Chất lượng dữ liệu và khả năng kết nối của từng doanh nghiệp quyết định trực tiếp khả năng được tác nhân lựa chọn. Chủ thể thứ ba là các nền tảng.

OTA, hệ thống thanh toán, bản đồ số, mạng xã hội, nền tảng đặt vé, hệ thống quản lý sự kiện hay các nhà cung cấp AI không còn chỉ là kênh phân phối. Họ là lớp điều phối giúp dữ liệu từ nhiều nguồn khác nhau có thể lưu thông, đồng bộ và trở thành các chuỗi hành động hoàn chỉnh. Mối quan hệ giữa địa phương và nền tảng vì vậy không nên chỉ là mua lưu lượng truy cập, mà phải là hợp tác để hình thành các tiêu chuẩn kết nối lâu dài. Chủ thể thứ tư là hạ tầng dữ liệu.

Đây là lớp nền mà người dùng gần như không nhìn thấy nhưng lại quyết định toàn bộ khả năng vận hành của tác nhân. Đồ thị tri thức, định danh thực thể, API, dữ liệu thời gian thực, chuẩn ngữ nghĩa, cơ chế xác thực, hệ thống đồng bộ và các bảng điều khiển chất lượng dữ liệu tạo nên "bộ thần kinh" của địa phương trong không gian số. Nếu hạ tầng này không tồn tại hoặc vận hành kém, mọi nỗ lực truyền thông phía trên sẽ chỉ tạo ra sự chú ý ngắn hạn chứ không chuyển hóa thành hành động. Còn lại là tác nhân và người dùng cuối.

Tác nhân không phải là đối tượng ở cuối chuỗi giá trị mà là điểm kiểm tra cuối cùng của toàn bộ hệ thống. Mỗi lần một tác nhân thành công trong việc tìm hiểu, lập kế hoạch, đặt dịch vụ, thanh toán và hoàn tất hành trình là một lần toàn bộ liên minh được xác nhận là đang hoạt động hiệu quả. Ngược lại, mỗi lần tác nhân thất bại cũng là tín hiệu cho thấy hệ sinh thái còn tồn tại một mắt xích chưa được kết nối.

Bốn chủ thể vừa nêu chưa đủ. Thiếu một bên mà nếu vắng mặt thì cả cấu trúc sẽ lệch: cộng đồng dân cư, các tổ chức nghề nghiệp và những người nắm giữ tri thức văn hóa. Đây là bên duy nhất có tư cách quyết định phần tri thức nào được đưa vào lược đồ và phần nào giữ lại, tức là bên thực thi quyền mờ. Nếu liên minh chỉ gồm chính quyền, doanh nghiệp, nền tảng và hạ tầng kỹ thuật thì mọi quyết định về dữ liệu văn hóa sẽ được đưa ra nhân danh cộng đồng mà không có cộng đồng. Chủ thể thứ sáu là chính các tác nhân cùng người dùng cuối. Quan hệ giữa sáu nhóm này là mạng lưới có phản hồi chứ không phải chuỗi truyền dẫn một chiều: mỗi thất bại của tác nhân là một tín hiệu quay ngược lên tất cả các bên.

Có một chi tiết dễ bị bỏ qua. Nó được hình thành thông qua chuẩn dữ liệu chung, khả năng liên thông và trách nhiệm duy trì chất lượng thông tin. Khi mọi thành viên cùng đóng góp vào một hạ tầng tri thức thống nhất, giá trị tạo ra sẽ lớn hơn rất nhiều so với tổng giá trị của từng tổ chức riêng lẻ.

Đó cũng là thay đổi sâu nhất trong quản trị địa phương. Chính quyền không chỉ điều phối dự án mà còn điều phối dòng chảy dữ liệu. Doanh nghiệp không chỉ cạnh tranh bằng chất lượng dịch vụ mà còn bằng chất lượng dữ liệu. Nền tảng không chỉ phân phối khách mà đồng kiến tạo hạ tầng. Và tác nhân trở thành một lớp khách hàng trung gian mới của cả hệ sinh thái, đứng giữa địa phương và người dùng cuối.

Vì vậy, trong nền kinh tế tác nhân, đơn vị cạnh tranh thực sự không còn là từng khách sạn, từng khu du lịch hay từng doanh nghiệp riêng lẻ. Đơn vị cạnh tranh là toàn bộ địa phương như một hệ sinh thái dữ liệu có khả năng được máy đọc, suy luận và hành động. Đó chính là liên minh mới quyết định năng lực cạnh tranh của địa phương trong kỷ nguyên AI.

Về thể chế, việc ban hành các văn bản luật đã tạo khung pháp lý cho cơ sở dữ liệu tổng hợp quốc gia, dữ liệu mở và các sản phẩm dịch vụ dữ liệu tại địa phương. Việc xây dựng khung kiến trúc dữ liệu cần được đặt ở cấp kiến trúc chứ không phải cấp ứng dụng. Về quy mô thị trường, sự gia tăng về lượt khách và tổng doanh thu chính là thước đo phản ánh khối lượng giao dịch đủ lớn để một điểm phần trăm cải thiện ở bộ lọc thứ tư trở thành con số có nghĩa.

Mắt xích yếu nhất của hệ sinh thái. Với mỗi chuỗi hành động cụ thể, khả hành động của một địa phương bị giới hạn bởi khâu yếu nhất trong chuỗi từ khám phá, xác minh, giữ chỗ, xác nhận đến thanh toán. Một khách sạn có dữ liệu hoàn hảo không bù được cho một hệ thống vận tải không có API, bởi tác nhân đánh giá khả năng hoàn thành cả nhiệm vụ chứ không chấm điểm từng nhà cung cấp.

Về chỉ báo sớm. Tỷ lệ tài sản đã trở thành tồn kho giao dịch được sẽ dự đoán thị phần khách của một địa phương tốt hơn quy mô ngân sách xúc tiến. Đây là giả thuyết chứ chưa phải quy luật, vì hiện chưa có bộ dữ liệu so sánh giữa các địa phương đủ dài để kiểm định và vì hai biến số này thường cùng tăng ở những nơi có năng lực quản trị tốt, khiến việc tách nhân quả khỏi tương quan rất khó.

Trung gian hóa lần hai. Địa phương chậm chuẩn hóa sẽ bị trung gian hóa một lần nữa, lần này không phải bởi các đại lý du lịch trực tuyến mà bởi lớp tác nhân nằm phía trên chúng. Khi ấy phần chênh lệch không mất đi mà chỉ chuyển chủ và địa phương phải trả hai lớp phí thay vì một. Điều kiện hiện có mới ở mức chỉ dấu, chưa đủ để khẳng định.

Đây là bài toán công nghệ, nên để doanh nghiệp tự lo. Nếu lợi ích của việc chuẩn hóa rơi trọn vào người bỏ chi phí thì lập luận này đúng. Thực tế thì không. Giá trị của một quan hệ trong đồ thị chỉ hiện ra khi cả hai đầu quan hệ cùng tồn tại, nghĩa là phần lớn lợi ích của khoản đầu tư đầu tiên rơi vào người đầu tư thứ hai. Đó là cấu trúc kinh điển của một thất bại điều phối và cũng là lý do lớp định danh cùng chuẩn kết nối mang tính chất hàng hóa công chứ không phải sản phẩm thương mại.

Không ký hợp đồng nền tảng mà không giữ quyền khả chuyển dữ liệu và quyền truy xuất bản sao đầy đủ theo định dạng máy đọc được. Không đo hiệu quả xúc tiến bằng lượt hiển thị khi bộ lọc thứ tư của chính địa phương chưa thông. Trong tình trạng ấy, chỉ số hiển thị đo nỗ lực chứ không đo kết quả.

## 7.8. KẾT LUẬN CHƯƠNG

Nếu các chương trước của cuốn sách trả lời câu hỏi địa phương cần được máy nhìn thấy như thế nào, thì Chương 7 đi thêm một bước quan trọng hơn, địa phương phải được máy hành động cùng như thế nào.

Đây là thay đổi căn bản về tiêu chuẩn cạnh tranh. Trong nền kinh tế trải nghiệm, một địa phương thành công khi tạo ra những không gian đáng nhớ, những lễ hội hấp dẫn, những di sản được gìn giữ và những câu chuyện chạm tới cảm xúc. Giá trị tích lũy qua trải nghiệm, ký ức và danh tiếng, với con người ở trung tâm quá trình lựa chọn.

Trong nền kinh tế tác nhân, tất cả những yếu tố ấy vẫn giữ nguyên giá trị. Không dữ liệu nào thay được một lễ hội đặc sắc, một di sản sống hay một bản sắc chân thực. AI không tạo ra giá trị của địa phương. AI chỉ quyết định giá trị ấy có được nhìn thấy, được hiểu đúng và được đưa vào quá trình lựa chọn hay không. Chính vì vậy, địa phương của thời đại này phải được xây trên hai lớp hạ tầng song song.

Lớp thứ nhất là hạ tầng trải nghiệm: quảng trường, công viên, bảo tàng, nhà hát, lễ hội, cảnh quan, cộng đồng, doanh nghiệp và toàn bộ những gì tạo nên một nơi chốn đáng sống, đáng đến và đáng nhớ.

Lớp thứ hai là hạ tầng khả hành động: mô hình dữ liệu, định danh thực thể, đồ thị tri thức, quan hệ ngữ nghĩa, tồn kho thời gian thực, API, chuẩn giao dịch, hệ thống xác nhận, thanh toán và các cơ chế để tác nhân có thể hoàn thành một chuỗi hành động mà không gặp đứt gãy.

Hai lớp này không cạnh tranh mà bổ sung cho nhau. Chỉ có hạ tầng trải nghiệm thì tạo ra giá trị nhưng có nguy cơ không được chọn. Chỉ có hạ tầng số mà thiếu trải nghiệm thật thì dễ được tìm thấy nhưng khó tạo ký ức và lòng trung thành. Trọng tâm của quản trị vì thế đang dịch chuyển: câu hỏi không còn là làm thế nào để quảng bá nhiều hơn, mà là làm thế nào để cả hệ sinh thái trở thành một hệ thống mà máy có thể đọc, suy luận và hành động một cách đáng tin cậy.

Điều này đòi hỏi một cách tiếp cận khác. Thương hiệu không còn chỉ được xây bằng truyền thông mà còn bằng chất lượng của hạ tầng tri thức. Quản trị không còn chỉ đo lượng khách hay lượng hiển thị mà còn đo độ đầy đủ của dữ liệu, tính nhất quán của thực thể và tỷ lệ tác nhân hoàn thành được nhiệm vụ. Đầu tư không còn chỉ dành cho quảng bá mà phải mở sang hạ tầng dữ liệu và năng lực liên thông.

Nói cách khác, thiết kế địa phương cho máy không phải là số hóa một điểm đến mà là tái thiết kế năng lực vận hành theo ngôn ngữ mà các hệ thống suy luận có thể dùng. Đây không phải là một dự án công nghệ mà là một chương trình chuyển đổi mô hình phát triển, nơi dữ liệu và khả năng hành động trở thành hạ tầng kinh tế giống như đường sá, điện hay viễn thông ở các giai đoạn trước.

Có thể rút gọn toàn bộ chương thành một mệnh đề: địa phương trong nền kinh tế tác nhân không được định nghĩa chỉ bởi những gì nó sở hữu, mà còn bởi những gì các tác nhân có thể thực hiện với nó. Tuy nhiên, năng lực đó không được xây trong khoảng trống. Mỗi quốc gia có điều kiện thể chế, dữ liệu, công nghệ và cấu trúc thị trường khác nhau và những gì triển khai nhanh ở một nền kinh tế phát triển chưa hẳn đã khả thi với một nền kinh tế khác. Vì vậy, sau khi xác lập mô hình lý tưởng, cuốn sách chuyển sang một câu hỏi thực tiễn hơn.

Đó chính là nội dung của Chương 8. Không còn bàn về việc địa phương nên được thiết kế như thế nào, mà đi sâu vào việc làm thế nào để một quốc gia đang phát triển có thể xây dựng được hạ tầng khả hành động cho toàn bộ hệ thống địa phương của mình trong kỷ nguyên AI. Trong nền kinh tế tác nhân, suy giảm không nhất thiết bắt đầu khi khách ngừng đến. Nó có thể bắt đầu sớm hơn, khi máy ngừng coi địa phương là một lựa chọn có thể thực hiện.

Chương này khép lại bằng Bảng 7.3, một bảng đối chiếu để mỗi khái niệm đã dựng ở các chương trước đều dẫn được tới một công cụ cụ thể. Bảng cũng là bản đồ đọc cho danh mục việc cần làm ở Phụ lục 8A.

**Bảng 7.3.** Từ khái niệm đến công cụ

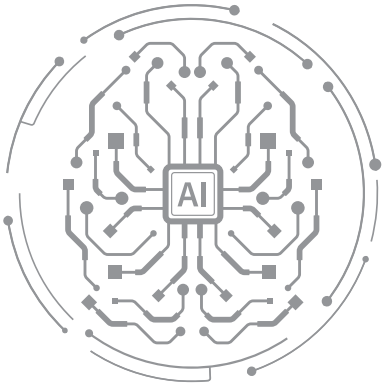
| Khái niệm | Nơi trình bày | Công cụ tương ứng |
|---|---|---|
| **Phân giải** | Chương 3 và Chương 4 | Quy chuẩn định danh, bảng ánh xạ tên, mã định danh bền qua đổi tên và sáp nhập |
| **Chứng thực** | Chương 4 | Phân công nguồn có thẩm quyền theo loại câu hỏi, ghi xuất xứ, lịch xác nhận định kỳ |
| **Xếp hạng** | Chương 3 và Chương 5 | Chỉ số hiện diện, tỷ lệ lọt tập ứng viên trên một bộ kịch bản truy vấn cố định |
| **Khả hành động** | Chương 3 và Chương 6 | Giao diện dữ liệu, tồn kho giao dịch được, điều kiện giao dịch và cam kết có hiệu lực |
| **Vòng ký ức** | Chương 4 | Cơ chế phát hiện sai lệch và quy trình sửa sai theo năm nấc chi phí |
| **Van quyền mờ** | Chương 2 và Chương 9 | Danh mục dữ liệu không công khai kèm căn cứ giữ mờ và chủ thể quyết định |

*Nguồn: Tác giả xây dựng.*

CHƯƠNG 8

# VIỆT NAM: DẤU THANH, ĐỊA DANH VÀ NĂNG LỰC CẤP XÃ

*Bộ lọc thứ nhất trong điều kiện Việt Nam. Phân mảnh thực thể do dấu thanh và viết tắt, địa danh sau sáp nhập, khoảng cách giữa dữ liệu hành chính và dữ liệu quyết định và năng lực số ở tuyến đầu là cấp xã.*


AI

Tám mục của chương này đều nằm ở bộ lọc thứ nhất. Đó là một lựa chọn có chủ ý, không phải một sự thu hẹp. Với phần lớn địa phương ở Việt Nam, điểm nghẽn chưa phải xếp hạng hay giao dịch mà là phân giải: một cái tên có quá nhiều cách viết, một địa giới vừa đổi mà dữ liệu chưa đổi theo. Cơ chế tái cấu trúc ở đây hiện ra theo chiều ngược lại và đó là chiều đáng lo hơn: khi lược đồ chưa nhận ra một thực thể, thực thể ấy không bị mô tả sai, nó đơn giản không tồn tại để mô tả. Một địa phương không biến mất vì bị lãng quên. Nó biến mất vì được ghi lại theo quá nhiều cách khác nhau.

Việt Nam bước vào nền kinh tế tác nhân với một lợi thế hiếm có nhưng cũng mang theo một nhóm trở ngại rất đặc thù. Lợi thế nằm ở chỗ đất nước sở hữu mật độ di sản, cảnh quan, văn hóa sống và trải nghiệm địa phương rất phong phú trên một không gian địa lý không quá lớn. Từ những đô thị di sản như Huế, Hội An, những kỳ quan thiên nhiên như Tràng An hay Phong Nha – Kẻ Bàng, đến các vùng văn hóa đặc sắc như Cao Bằng, Tây Bắc hay Đồng bằng sông Cửu Long, Việt Nam có một nguồn tài sản trải nghiệm mà không nhiều quốc gia có được. Tuy nhiên, chính sự phong phú ấy lại chưa được chuyển hóa đầy đủ thành một hạ tầng dữ liệu mà máy có thể đọc, hiểu và hành động.

Nếu các chương trước phân tích những nguyên lý phổ quát của nền kinh tế tác nhân, chương này đi vào những vấn đề chỉ Việt Nam hoặc một số rất ít quốc gia gặp phải: dấu thanh và bảng mã, địa danh có nhiều cách viết, đơn vị hành chính vừa thay đổi trên quy mô lớn, dữ liệu phân tán giữa nhiều cơ quan và khoảng cách giữa dữ liệu phục vụ quản lý với dữ liệu phục vụ giao dịch.

Một nghịch lý đáng chú ý là phần lớn những rào cản này không xuất phát từ trí tuệ nhân tạo. AI chỉ bộc lộ những điểm yếu vốn đã tồn tại trong cách địa phương được tổ chức và biểu diễn trên môi trường số. Khi một địa danh tồn tại dưới nhiều tên gọi khác nhau, khi dữ liệu giữa các cơ quan không liên thông, khi giờ mở cửa, giá vé, tình trạng hoạt động hay thông tin liên hệ không được cập nhật thường xuyên, vấn đề không nằm ở khả năng suy luận của mô hình ngôn ngữ mà nằm ở chất lượng của hệ sinh thái dữ liệu đầu vào. Máy có thể suy luận qua một phần mâu thuẫn, nhưng không thể tự tạo ra một sự nhất quán có thẩm quyền khi hệ sinh thái dữ liệu đầu vào chưa nhất quán.

Phần lớn trở ngại ấy không đòi hỏi công nghệ đột phá. Chúng nằm ở những công việc cơ bản nhưng lâu nay ít được coi là hạ tầng: chuẩn hóa tên gọi, hợp nhất thực thể, xác lập quyền cập nhật, xây dựng chuẩn liên thông và duy trì dữ liệu ở trạng thái sống.

Chương này cũng muốn thay đổi một cách nhìn khá phổ biến về chuyển đổi số địa phương. Trong nhiều năm, chuyển đổi số thường được hiểu là xây dựng cổng thông tin điện tử, số hóa hồ sơ, triển khai dịch vụ công trực tuyến hoặc đầu tư các nền tảng quản trị. Những việc đó đều cần thiết, nhưng chưa đủ để địa phương hiện diện trong nền kinh tế tác nhân. Một địa phương có thể vận hành chính quyền điện tử rất tốt nhưng vẫn gần như vô hình trước các hệ thống AI nếu các thực thể du lịch, dịch vụ, sự kiện, hạ tầng và khả năng giao dịch không được biểu diễn theo cách mà máy có thể nhận diện và sử dụng.

Chính vì vậy, trọng tâm của chương này không phải là công nghệ, mà là năng lực biểu diễn địa phương trước máy. Từ bài toán dấu thanh, địa danh sau sáp nhập, sự khác biệt giữa dữ liệu hành chính và dữ liệu điểm đến, đến năng lực số ở cấp xã và cấp tỉnh, rồi tới cơ chế phối hợp giữa chính quyền với doanh nghiệp, tất cả đều dẫn tới cùng một câu hỏi: địa phương có đang tồn tại dưới dạng mà các hệ thống AI có thể đọc, hiểu và giao dịch hay không?

Đây cũng là chương thực hành nhất của cuốn sách. Nếu các chương trước trả lời câu hỏi vì sao nền kinh tế tác nhân vận hành như vậy, chương này trả lời câu hỏi Việt Nam cần làm gì. Điều đáng lạc quan là phần lớn rào cản khả đọc của Việt Nam không phải giới hạn bất biến, mà là những vấn đề có thể khắc phục bằng tiêu chuẩn dữ liệu tốt hơn và sự phối hợp rõ hơn giữa các cấp chính quyền, doanh nghiệp và cộng đồng.

Rào cản lớn nhất giữa Việt Nam và nền kinh tế tác nhân không phải là thiếu công nghệ, mà là thiếu năng lực tự biểu diễn. Một địa phương chỉ trở nên khả đọc khi danh tính, lịch sử, trạng thái vận hành và quyền cập nhật của nó được tổ chức thành một hệ thống dữ liệu nhất quán.

Quy mô cuộc sắp xếp hành chính năm 2025: Từ ngày 1 tháng 7 năm 2025, chính quyền địa phương Việt Nam vận hành theo hai cấp là cấp tỉnh và cấp xã. Đơn vị hành chính cấp huyện kết thúc hoạt động. Số đơn vị hành chính cấp tỉnh giảm từ 63 xuống 34, gồm 28 tỉnh và 6 thành phố trực thuộc trung ương. Số đơn vị hành chính cấp xã còn 3.321, gồm 2.621 xã, 687 phường và 13 đặc khu, giảm 6.714 đơn vị so với trước khi sắp xếp, tương đương giảm 66,91%.[1]

1. Nghị quyết 60-NQ/TW ngày 12 tháng 4 năm 2025 và Nghị quyết 202/2025/QH15 của Quốc hội. Số đơn vị hành chính cấp xã sau sắp xếp là 3.321, giảm 6.714 đơn vị.

## 8.1. BÀI TOÁN DẤU THANH VÀ PHÂN MẢNH THỰC THỂ

Nếu Chương 7 đã chỉ ra rằng điều kiện đầu tiên để một địa phương tham gia nền kinh tế tác nhân là phải trở thành một thực thể mà máy có thể nhận diện và hành động, thì ở Việt Nam, trở ngại đầu tiên lại xuất hiện ngay từ chính tên gọi. Trước khi AI có thể hiểu một điểm đến đẹp đến đâu, giàu di sản đến mức nào hay phù hợp với ai, nó phải trả lời được một câu hỏi tưởng như rất đơn giản: đây có phải là cùng một địa điểm hay không?

Con người hợp nhất các cách gọi gần như tự động. Chúng ta đọc "Huế" và "Hue", "Hội An" và "Hoi An", "TP.HCM" và "Ho Chi Minh City", rồi hiểu ngay rằng tất cả đang nói về cùng một nơi, dựa vào ngữ cảnh, kinh nghiệm và tri thức xã hội.

Máy cũng làm được điều đó, nhưng chỉ ở một tầng và chỉ với một mức xác suất. Một mô hình ngôn ngữ hiện nay thường nhận ra rằng "Hue" và "Huế" tương đương trong lúc đọc. Điều nó không làm được là hợp nhất các bản ghi trong lúc lưu. Nếu năm biến thể tên nằm ở năm bản ghi khác nhau trong chỉ mục truy hồi, thì đánh giá, liên kết và lịch sử giao dịch vẫn được đếm riêng trên từng bản ghi. Khả năng hiểu tên tương đương không đồng nghĩa với khả năng hợp nhất lịch sử dữ liệu và toàn bộ mục này nằm ở khoảng cách giữa hai khả năng ấy.

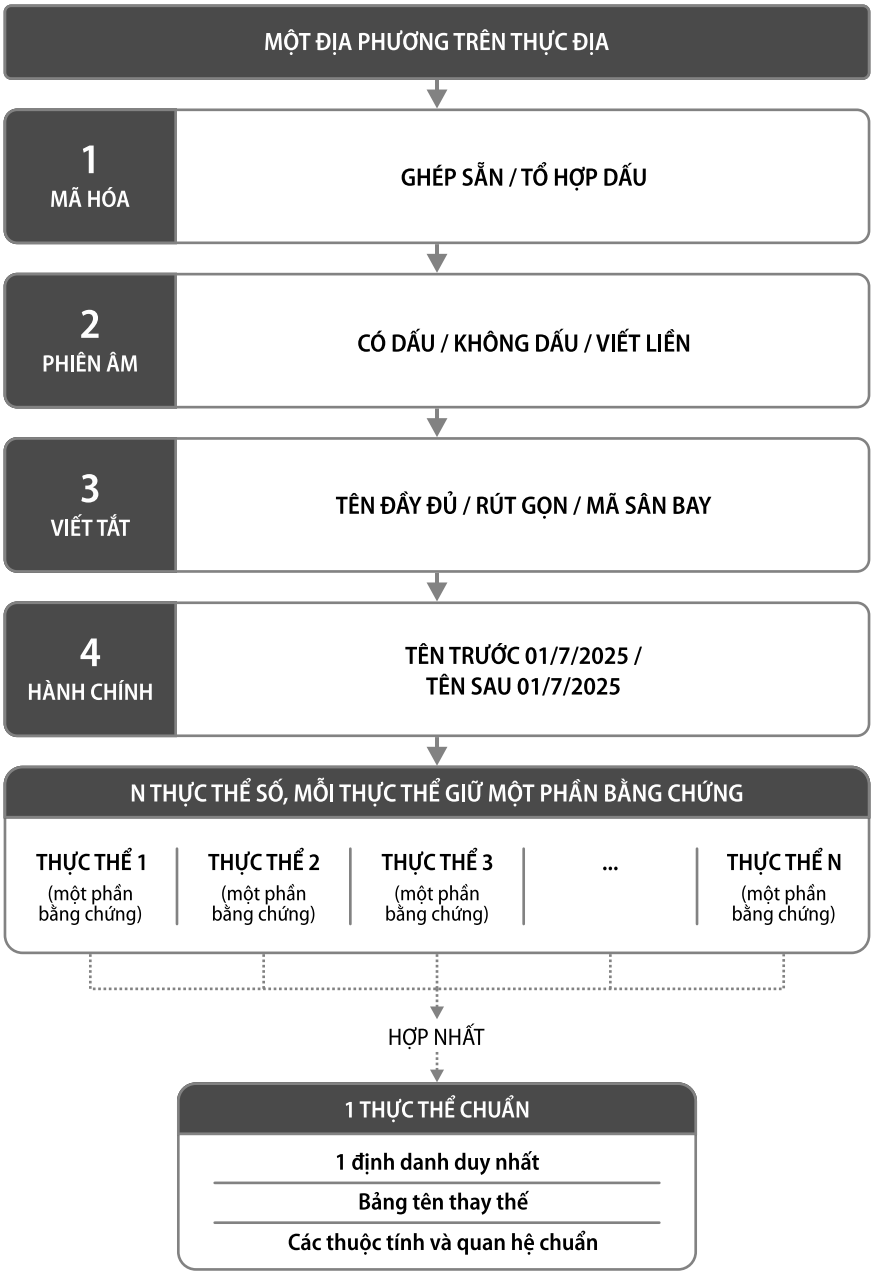
Ở Việt Nam, vấn đề này mang một tổ hợp đặc biệt. Ngôn ngữ có dấu thanh, nhiều quy ước phiên âm, nhiều cách viết tắt và sự song song giữa tiếng Việt với tiếng Anh khiến một địa danh có thể tồn tại dưới nhiều "danh tính số" khác nhau. Quá trình đô thị hóa nhanh, thay đổi địa giới hành chính và sự phát triển của nhiều nền tảng số trong các giai đoạn khác nhau càng làm số lượng biến thể tăng lên theo thời gian. Không có quốc gia nào

hoàn toàn tránh khỏi vấn đề này, nhưng ở Việt Nam, mức độ phức tạp lớn hơn đáng kể vì sự kết hợp đồng thời giữa ngôn ngữ, lịch sử hành chính và quá trình số hóa chưa đồng bộ.

Đây không phải vấn đề của riêng các hệ thống AI. Công cụ tìm kiếm, bản đồ số, nền tảng đặt phòng và cơ sở dữ liệu du lịch đều gặp cùng bài toán. AI chỉ làm nó lộ ra rõ hơn, vì phải tổng hợp nhiều nguồn để tạo một câu trả lời duy nhất.

Vì vậy, bài toán của mục này không phải là dấu thanh, Unicode hay cách gõ tiếng Việt theo nghĩa kỹ thuật. Đó chỉ là biểu hiện bề mặt của một vấn đề sâu hơn: làm thế nào để toàn bộ hệ sinh thái số cùng nhận diện một địa phương như một thực thể duy nhất. Chỉ khi câu hỏi đó được giải quyết, mọi lớp dữ liệu phía sau, từ tri thức, danh tiếng, tồn kho đến khả năng giao dịch, mới có thể tích lũy trên cùng một nền tảng. Trong nền kinh tế tác nhân, khả năng được máy nhận diện đúng luôn đến trước khả năng được máy đề xuất và được đề xuất luôn đến trước khả năng được máy giao dịch. Chính vì vậy, bài toán dấu thanh thực chất là bài toán đầu tiên của chủ quyền dữ liệu địa phương trong kỷ nguyên AI.

Hội tụ danh tính. Tín hiệu về một địa phương chỉ tích lũy được trên thực thể mà máy coi là một. Khi tín hiệu về cùng một địa phương bị phân bổ giữa nhiều bản ghi chưa hợp nhất, lượng bằng chứng hội tụ trên mỗi thực thể giảm. Các hệ thống thường tích lũy và đánh giá bằng chứng theo từng thực thể hoặc từng bản ghi, chứ không mặc nhiên cộng trên tổng các biến thể. Mệnh đề này độc lập với năng lực của mô hình đọc: một hệ thống nhận ra hai tên là tương đương vẫn không vì thế cộng dồn bằng chứng của hai bản ghi.




**Hình 8.1.** Bốn nguồn phân mảnh giữa một địa phương và thực thể mà máy nhìn thấy

*Ghi chú: Bốn nguồn hoạt động độc lập, nên về lý thuyết số biến thể có thể tăng theo tích của số biến thể ở từng tầng chứ không theo tổng. Trên thực tế không phải mọi tổ hợp đều xuất hiện.*

Hai trường hợp ngoài các trung tâm lớn cho thấy vấn đề không nằm ở quy mô. Sa Pa tồn tại trong dữ liệu dưới ít nhất ba dạng viết, gồm Sa Pa, Sapa và SaPa, cộng thêm sự thay đổi từ huyện sang thị xã rồi vào tỉnh Lào Cai sau sắp xếp. Một tác nhân đối chiếu ba dạng ấy có thể trả về ba tập kết quả khác nhau cho cùng một câu hỏi. Hà Giang thì gặp bài toán ngược: tên vùng nổi tiếng hơn tên đơn vị hành chính, nên phần lớn nội dung trên mạng nói về một cao nguyên mà máy không nối được vào thực thể hành chính nào có dữ liệu vận hành. Trong cả hai trường hợp, chi phí sửa gần bằng không về mặt kỹ thuật và gần bằng vô cùng về mặt thẩm quyền, vì không ai được phân công làm việc đó.

### 8.1.1. Một địa danh có bao nhiêu tên

Máy phải quyết định trước khi suy luận: hai chuỗi ký tự này có đại diện cho cùng một thực thể hay không. Thiếu bằng chứng để hợp nhất, mỗi cách viết trở thành một ứng viên riêng.

Chỉ riêng Hội An cũng có thể xuất hiện dưới rất nhiều dạng:

- *Hội An.*
- *Hoi An.*
- *HoiAn.*
- *Hội-an.*
- *Hoi-An.*
- *Ancient Town of Hoi An.*
- *Hoi An Ancient Town.*
- *Phố cổ Hội An.*
- *Hoi An Old Town.*

Đối với khách du lịch, đây chỉ là những cách diễn đạt khác nhau. Nhưng trong nhiều hệ thống dữ liệu, chúng có thể được

lưu trữ thành những bản ghi khác nhau, nhận những liên kết khác nhau và tích lũy những tín hiệu khác nhau. Nếu không có cơ chế phân giải thực thể (entity resolution), lịch sử đánh giá, hình ảnh, bài viết, tọa độ, sự kiện và dữ liệu giao dịch có thể bị chia nhỏ thay vì tập trung vào một hồ sơ thống nhất.

***Hội An.*** Biến thể tiếng Việt: Hội An, phố cổ Hội An, Hội-An. Biến thể tiếng Anh và viết tắt: Hoi An, Hoi An, Hoi An Ancient Town, Hoi An Old Town. Rủi ro chính: Lẫn phạm vi giữa thành phố và khu phố cổ.

***Huế.*** Biến thể tiếng Việt: Huế, Cố đô Huế, thành phố Huế. Biến thể tiếng Anh và viết tắt: Hue, Hue City, Imperial City of Hue, Complex of Hue Monuments. Rủi ro chính: Lẫn giữa đơn vị hành chính và quần thể di tích.

***Thành phố Hồ Chí Minh.*** Biến thể tiếng Việt: Thành phố Hồ Chí Minh, TP.HCM, Sài Gòn. Biến thể tiếng Anh và viết tắt: Ho Chi Minh City, HCMC, Saigon, SGN. Rủi ro chính: Tín hiệu bị chia giữa tên hành chính và tên thông dụng.

Danh sách trên chỉ nhằm minh họa, không nhằm liệt kê đầy đủ. Điều đáng chú ý là mọi biến thể đều đúng trong ngữ cảnh của nó.

Vấn đề phức tạp thêm vì các hệ thống được xây ở những thời điểm khác nhau với quy tắc khác nhau. Một cơ sở dữ liệu cũ lưu "Hoi An" để tránh lỗi dấu. Một nền tảng quốc tế dùng "Hoi An Ancient Town" theo tên trong danh mục di sản. Một trang web địa phương dùng "Phố cổ Hội An". Mỗi hệ thống đều hợp lý theo cách riêng, nhưng khi phải tổng hợp các nguồn đó, việc xác định chúng nói về cùng một thực thể không còn đơn giản.

Số lượng tên gọi tăng theo cả ngôn ngữ và thời gian. Một địa phương có thể thay đổi tên chính thức, thay đổi cấp hành chính, đổi cách phiên âm hoặc phát sinh thêm những tên gọi phổ biến trong du lịch. Mỗi lần thay đổi đều để lại một "dấu vết dữ liệu" trên internet. Internet gần như không quên những tên cũ, trong khi các hệ thống AI lại phải học cách nối tất cả các dấu vết đó thành một lịch sử liên tục.

Vậy câu hỏi "một địa danh có bao nhiêu tên" không phải câu hỏi ngôn ngữ học. Nó là câu hỏi về số lượng thực thể mà hệ sinh thái số đang giữ cho cùng một nơi.

Phong Nha: một cái tên, bốn phạm vi khác nhau. Sau tháng 7 năm 2025, cụm từ Phong Nha trong dữ liệu có thể chỉ ít nhất bốn đối tượng khác nhau. Nó có thể chỉ động Phong Nha, một hang động cụ thể. Nó có thể chỉ xã Phong Nha, một đơn vị hành chính. Nó có thể chỉ Vườn quốc gia Phong Nha - Kẻ Bàng, một khu bảo tồn. Nó cũng có thể chỉ di sản thế giới liên biên giới hiện có tên gồm cả phần Hin Nam Nô của Lào. Bốn đối tượng này lồng nhau nhưng không trùng nhau. Diện tích khác nhau, cơ quan quản lý khác nhau, quy chế pháp lý khác nhau và điều quan trọng với một tác nhân là điều kiện tiếp cận cũng khác nhau. Khi một người nói rằng họ muốn đi Phong Nha hai ngày, con người hiểu ngay đó là khu du lịch hang động. Máy phải chọn giữa bốn thực thể. Lựa chọn ấy quyết định toàn bộ phần còn lại của hành trình: vé nào, giờ nào, đi bằng đường nào, ở đâu. Nguồn và mức xác thực: Bốn phạm vi nêu trên đều là thực thể có thật. Cách trình bày dưới dạng bài toán phân giải thực thể là của tác giả.

Mô hình ngôn ngữ ngày nay đã đủ giỏi để tự hiểu mọi biến thể tên, nên bài toán dấu thanh sẽ tự biến mất. Trả lời: Phần đúng của phản bác này khá lớn. Các mô hình hiện nay xử lý biến thể

chính tả, phiên âm và viết tắt tốt hơn hẳn các hệ thống tìm kiếm theo từ khóa của thập niên trước và khoảng cách đó vẫn đang thu hẹp. Nhưng nó trả lời sai tầng của vấn đề. Khả năng nhận diện linh hoạt của mô hình nằm ở bước đọc, trong khi phân mảnh xảy ra ở bước lưu. Nếu năm biến thể tên của một điểm đến nằm ở năm bản ghi khác nhau trong chỉ mục truy hồi, thì một mô hình có độ dung sai cao vẫn truy hồi được từng bản ghi, nhưng số đánh giá, số liên kết và lịch sử giao dịch vẫn được đếm riêng trên từng bản ghi. Mô hình có thể suy đoán rằng chúng cùng chỉ một nơi, nhưng không vì thế mà tự hợp nhất và cộng dồn bằng chứng của các bản ghi. Khả năng nhận diện linh hoạt giúp giảm tỷ lệ bỏ sót. Nó không tạo ra một thực thể chuẩn để bằng chứng tự động hội tụ.

### 8.1.2. Dấu thanh không chỉ là dấu thanh

Nhiều người cho rằng dấu tiếng Việt chỉ là một vấn đề hiển thị. Nếu máy vẫn đọc được "Hue" thay vì "Huế" hay "Da Nang" thay vì "Đà Nẵng", thì việc có dấu hay không có vẻ không quan trọng. Thực tế phức tạp hơn nhiều. Trong thế giới số, dấu thanh không chỉ liên quan đến cách một từ được nhìn thấy, mà còn liên quan đến cách dữ liệu được lưu trữ, trao đổi, đối chiếu và hợp nhất. Một khác biệt rất nhỏ ở tầng mã hóa có thể khiến hai chuỗi ký tự giống hệt nhau đối với con người trở thành hai giá trị hoàn toàn khác nhau đối với máy.

Chuẩn hóa Unicode. Một chữ tiếng Việt có dấu có thể được lưu theo hai cách. Cách thứ nhất dùng một điểm mã đã ghép sẵn cả chữ và dấu. Cách thứ hai dùng chữ cơ sở rồi ghép thêm các dấu như những điểm mã riêng. Hai cách này hiển thị giống nhau tuyệt đối trên màn hình nhưng là hai chuỗi byte khác nhau. Một

phép so sánh chuỗi thông thường sẽ kết luận rằng đây là hai giá trị khác nhau. Chuẩn hóa Unicode là bước quy mọi chuỗi về cùng một dạng trước khi so sánh, lưu trữ hoặc đối chiếu. Chi phí kỹ thuật của bước này rất thấp nếu được đặt ngay ở cửa vào của hệ thống, nhưng tăng mạnh khi phải làm sạch một kho dữ liệu lịch sử đã tích lũy.

Khác biệt ấy trở thành vấn đề khi các hệ thống không chuẩn hóa chuỗi trước khi so sánh hoặc hợp nhất dữ liệu.

Nếu hai hệ thống không sử dụng cùng một quy tắc chuẩn hóa Unicode, chúng có thể coi hai chuỗi ấy là hai giá trị khác nhau. Một hệ thống quản trị địa phương, một cơ sở dữ liệu du lịch và một nền tảng AI hoàn toàn có thể đang lưu cùng một địa danh theo ba cách mã hóa khác nhau mà không hề nhận ra điều đó. Khi dữ liệu được hợp nhất, tìm kiếm hoặc đối chiếu, những khác biệt vô hình này bắt đầu tạo ra lỗi.

Khó khăn chưa dừng lại ở Unicode. Trong nhiều năm, Việt Nam đã trải qua nhiều giai đoạn sử dụng các bảng mã khác nhau như TCVN3, VNI, Unicode hay các hệ thống chuyển mã riêng của từng phần mềm. Dù ngày nay Unicode đã trở thành tiêu chuẩn áp đảo, một lượng lớn dữ liệu lịch sử vẫn được tạo ra từ những hệ thống cũ. Mỗi lần dữ liệu được chuyển đổi giữa các bảng mã, nguy cơ mất dấu, lỗi ký tự hoặc sai lệch tên địa danh lại xuất hiện. Không ít cơ sở dữ liệu vẫn còn những địa danh bị biến thành các chuỗi ký tự khó hiểu hoặc mất hoàn toàn dấu tiếng Việt sau nhiều lần chuyển đổi.

Trong phạm vi một hệ thống riêng lẻ, những sai khác này có thể không gây hậu quả lớn. Nhưng một hệ thống AI thường phải tổng hợp thông tin từ nhiều nguồn: cổng thông tin địa phương,

bản đồ số, nền tảng đặt phòng, kho tri thức, báo chí và cơ sở dữ liệu doanh nghiệp. Chỉ cần một phần trong số đó không chuẩn hóa Unicode, việc hợp nhất thực thể khó lên đáng kể.

Vì vậy bài toán Unicode không dành riêng cho kỹ sư phần mềm. Một địa phương có thể đầu tư rất nhiều cho truyền thông, nhưng nếu dữ liệu về chính mình không được chuẩn hóa theo cùng một quy tắc mã hóa, phần giá trị tích lũy vẫn tiếp tục bị chia.

Đây cũng là một trong những điểm đáng lạc quan nhất của Việt Nam. So với việc xây dựng các mô hình AI mới hay đầu tư vào hạ tầng tính toán đắt đỏ, việc chuẩn hóa Unicode và áp dụng một chuẩn mã hóa ký tự thống nhất có chi phí rất thấp nhưng mang lại hiệu quả rất lớn. Đó là một cải thiện ở tầng hạ tầng dữ liệu, gần như vô hình với người sử dụng, nhưng lại quyết định việc các hệ thống AI có thực sự nhìn thấy địa phương như một thực thể thống nhất hay không. Trong nền kinh tế tác nhân, đôi khi một tiêu chuẩn dữ liệu đúng còn có giá trị lớn hơn nhiều một chiến dịch quảng bá tốn kém.

Tên tiếng Anh, tiếng Pháp hay tiếng Nhật cũng có nhiều biến thể. Trả lời: Đúng. Chương này không lập luận rằng Việt Nam là ngoại lệ. Munich và München, Firenze và Florence, Beijing và Peking đều là những trường hợp phân mảnh danh tính kinh điển trong dữ liệu địa lý. Điểm khác biệt không nằm ở việc có hay không, mà ở việc ba nguồn phân mảnh cùng hoạt động một lúc: hệ chữ có dấu với nhiều cách mã hóa, một truyền thống phiên âm không chuẩn hóa và một chu kỳ cải cách hành chính tương đối dày. Một quốc gia chỉ gặp yếu tố thứ nhất có thể xử lý bằng chuẩn hóa mã hóa. Một quốc gia gặp cả ba cần một chương trình quản trị thực thể.

### 8.1.3. Từ viết tắt tạo ra thực thể mới

Nếu dấu thanh làm một địa danh xuất hiện dưới nhiều dạng khác nhau thì các từ viết tắt còn tạo ra một lớp phân mảnh phức tạp hơn. Đối với con người, "Thành phố Hồ Chí Minh", "TP.HCM", "Ho Chi Minh City", "HCMC", "Saigon" hay thậm chí "SG" đều có thể được hiểu là cùng một nơi, tùy theo ngữ cảnh. Nhưng đối với máy, mỗi cách gọi trước hết là một chuỗi ký tự độc lập. Chỉ sau khi được đối chiếu với nhiều nguồn dữ liệu khác nhau, chúng mới có thể được hợp nhất thành một thực thể duy nhất.

Lấy Thành phố Hồ Chí Minh làm ví dụ. Chỉ riêng trên internet đã tồn tại đồng thời rất nhiều cách biểu diễn:

- *Thành phố Hồ Chí Minh.*
- *TP. Hồ Chí Minh.*
- *TP.HCM.*
- *TP HCM.*
- *Ho Chi Minh City.*
- *HCMC.*
- *Saigon.*
- *SG.*

Mỗi tên gọi đều có một lịch sử và một ngữ cảnh riêng. "Saigon" vẫn được sử dụng rộng rãi trong ngành du lịch, khách sạn, nhà hàng và hàng không. "HCMC" phổ biến trong cộng đồng doanh nghiệp quốc tế và nhiều cơ sở dữ liệu. "TP.HCM" xuất hiện trong văn bản hành chính. Trong khi đó, "SG" lại được sử dụng rất nhiều trên mạng xã hội, trong thương mại điện tử và giao tiếp hằng ngày.

Khó khăn thật sự nằm ở chiều thời gian. "Saigon" trong ngữ cảnh du lịch hiện nay thường tương ứng với Thành phố Hồ Chí

Minh, còn trong nhiều tài liệu lịch sử, "Sài Gòn" lại chỉ một phạm vi hành chính khác trước năm 1975. Cùng một tên gọi mang những phạm vi khác nhau theo từng thời kỳ, nên hợp nhất thực thể là bài toán lịch sử chứ không chỉ là bài toán chính tả.

Hệ quả là mỗi từ viết tắt có thể trở thành một cửa vào riêng của cùng một địa phương. Đánh giá tích lũy dưới "Ho Chi Minh City", bài viết dưới "Saigon", sự kiện dưới "TP.HCM", dữ liệu doanh nghiệp dưới "HCMC". Không có liên kết giữa các cửa vào, hệ thống xử lý nhiều hồ sơ nhỏ thay vì một hồ sơ đủ mạnh.

Hiện tượng này không riêng của Thành phố Hồ Chí Minh: Hà Nội, Bà Rịa - Vũng Tàu, Thừa Thiên Huế trước đây và thành phố Huế hiện nay đều tồn tại song song nhiều cách gọi trong các cơ sở dữ liệu khác nhau.

Bài toán không phải là loại bỏ từ viết tắt, mà là ghi nhận chúng như những tên thay thế của một thực thể chuẩn. Sự đa dạng nằm ở tầng tên gọi, tính thống nhất phải nằm ở tầng định danh.

### 8.1.4. Khi tiếng Anh và tiếng Việt không gặp nhau

Nếu dấu thanh và từ viết tắt làm một địa danh xuất hiện dưới nhiều biến thể trong cùng một ngôn ngữ, thì sự song song giữa tiếng Việt và tiếng Anh lại tạo ra một lớp phân mảnh sâu hơn. Đây là đặc điểm rất phổ biến ở các quốc gia có ngành du lịch phát triển, nhưng ở Việt Nam vấn đề trở nên phức tạp hơn vì quá trình quốc tế hóa diễn ra không đồng đều. Nhiều địa phương chưa bao giờ thống nhất cách phiên âm chính thức, trong khi các doanh nghiệp, nền tảng quốc tế và cơ quan quản lý lại sử dụng những quy ước khác nhau.

Lấy Huế làm ví dụ. Đối với người Việt, "Huế" là một địa danh duy nhất. Nhưng trên môi trường số, địa phương này có thể xuất hiện dưới nhiều dạng:

- *Huế.*
- *Hue.*
- *Hue City.*
- *Thành phố Huế.*
- *Imperial City of Hue.*
- *Hue Imperial City.*

Mỗi tên gọi phản ánh một góc nhìn: có tên chỉ thành phố, có tên nhấn mạnh quần thể di tích, có tên phục vụ khách quốc tế, có tên dùng trong văn bản hành chính. Con người đọc bốn cách gọi ấy và hiểu ngay quan hệ giữa chúng. Với máy, việc xác định chúng là một thực thể hay nhiều thực thể là một bước suy luận riêng.

Khó khăn lớn hơn nằm ở chỗ tiếng Anh không chỉ là bản dịch của tiếng Việt. Trong nhiều trường hợp, tên tiếng Anh phản ánh một khái niệm khác. "Hue" có thể chỉ thành phố Huế, nhưng "Complex of Hue Monuments" lại chỉ quần thể di tích được UNESCO công nhận. "Hoi An" có thể chỉ toàn bộ thành phố, trong khi "Hoi An Ancient Town" chỉ khu phố cổ. "Phong Nha" có thể chỉ thị trấn, còn "Phong Nha - Ke Bang National Park" lại là một vườn quốc gia. Nếu các thực thể này không được mô tả rõ ràng trong đồ thị tri thức, AI rất dễ trộn lẫn phạm vi địa lý và phạm vi chức năng của từng địa danh.

Mỗi nền tảng quốc tế lại dùng một quy ước riêng: bản đồ theo tên địa lý chính thức, nền tảng đặt phòng theo tên quen với khách quốc tế, hãng hàng không theo mã sân bay, tổ chức quốc tế theo tên đã đăng ký trong hồ sơ di sản. Cùng một nơi được biểu diễn theo nhiều tiêu chuẩn, không có điểm quy chiếu chung.

Người dùng một công cụ tìm kiếm có thể tự thử nhiều từ khóa. Một tác nhân phải tự làm việc đó. Khi người dùng chỉ nói "tôi muốn một chuyến đi ba ngày ở Huế", hệ thống phải biết rằng thông tin liên quan có thể nằm dưới "Huế", "Hue", "Hue City" hay "Imperial City of Hue".

Bài toán vì thế không nằm ở việc lựa chọn tiếng Việt hay tiếng Anh làm chuẩn. Một địa phương trong nền kinh tế tác nhân phải đa ngôn ngữ nhưng đơn thực thể. Nghĩa là mọi cách gọi bằng tiếng Việt, tiếng Anh và các ngôn ngữ khác đều phải hội tụ về cùng một hồ sơ số trung tâm. Tên gọi có thể khác nhau để phục vụ những nhóm người dùng khác nhau, nhưng phía sau tất cả phải là một định danh duy nhất, nơi toàn bộ dữ liệu, quan hệ và lịch sử tương tác được tích lũy liên tục.

Đó cũng là khác biệt giữa dịch tên và quản trị thực thể. Dịch tên chỉ giúp con người đọc được địa danh bằng một ngôn ngữ khác. Quản trị thực thể giúp các hệ thống AI hiểu rằng tất cả những tên gọi ấy đang chỉ về cùng một địa phương. Trong nền kinh tế tác nhân, giá trị không nằm ở số lượng ngôn ngữ được hỗ trợ, mà nằm ở khả năng duy trì một danh tính số thống nhất xuyên suốt mọi ngôn ngữ và mọi nền tảng.

### 8.1.5. Kiến trúc tối thiểu của danh tính địa phương

Bốn nguồn phân mảnh ở các tiết trên dẫn tới cùng một yêu cầu và yêu cầu ấy có thể phát biểu gọn. Hệ quả nặng nhất của phân mảnh không phải là khó tìm kiếm mà là niềm tin của máy không tích lũy được. Chương 4 đã cho thấy ký ức thuật toán hình thành từ chuỗi bằng chứng gắn vào một thực thể, nên khi bằng chứng bị chia, không thực thể nào đủ mạnh để trở thành một hồ sơ đáng tin cậy.

Chính vì vậy, phần lớn hệ thống AI đều cố gắng xây dựng một thực thể chuẩn (canonical entity) làm trung tâm của toàn bộ dữ liệu. Điều quan trọng không phải là bắt buộc mọi người chỉ sử dụng một tên duy nhất. Một người vẫn có thể tìm "Hue", "Huế" hay "Cố đô Huế", một doanh nghiệp vẫn có thể ghi "HCMC" hoặc "TP.HCM". Điều quan trọng là phía sau tất cả những tên gọi đó phải tồn tại một định danh chuẩn, nơi mọi dữ liệu được hợp nhất và tiếp tục tích lũy theo thời gian.

Thực thể chuẩn và bảng tên thay thế. Thực thể chuẩn là bản ghi trung tâm mà mọi dữ liệu về một đối tượng được gắn vào. Nó có một định danh ổn định trong suốt vòng đời của thực thể, không thay đổi khi tên gọi hoặc đơn vị quản lý thay đổi. Khi một thay đổi tạo ra thực thể pháp lý hoặc địa lý mới, định danh mới được cấp và nối với định danh tiền nhiệm bằng quan hệ kế thừa, thay vì tiếp tục dùng định danh cũ như thể vẫn là một thực thể. Bảng tên thay thế là danh sách mọi cách gọi khác đang trỏ về thực thể chuẩn đó, kèm ngôn ngữ, ngữ cảnh sử dụng và khoảng thời gian có hiệu lực. Tên cũ không bị xóa, nó được ghi kèm ngày kết thúc hiệu lực. Hai thành phần này giải quyết một mâu thuẫn tưởng như không thể dung hòa: hệ thống cần một điểm quy chiếu duy nhất để tích lũy bằng chứng, trong khi đời sống cần nhiều cách gọi. Điểm quy chiếu duy nhất nằm ở tầng định danh, sự đa dạng nằm ở tầng tên gọi.

Chuẩn hóa tên gọi là áp đặt một cách gọi duy nhất, làm mất sự phong phú của địa danh Việt Nam. Trả lời: Đây là một lo ngại chính đáng và cần được phân biệt rõ. Tên chuẩn trong kỹ thuật dữ liệu không phải tên duy nhất được phép dùng. Nó là điểm quy chiếu nội bộ để hệ thống biết mọi cách gọi đang trỏ về đâu. Trong một lược đồ dữ liệu thiết kế đúng, bảng tên thay thế là nơi sự phong phú được bảo tồn, không phải nơi nó bị xóa. Một

địa danh có mười cách gọi thì cả mười đều được ghi lại kèm ngữ cảnh sử dụng, thay vì chín cách bị loại và một cách được giữ. Phần chưa giải quyết được nằm ở chỗ khác. Việc quyết định cách gọi nào được đặt ở vị trí tên chuẩn vẫn là một quyết định có tính quyền lực và nó thường được đưa ra bởi cơ quan hành chính chứ không phải bởi cộng đồng đang sống với địa danh đó. Quyền mờ nêu ở Chương 1 áp dụng cho phần tri thức không đưa vào lược đồ, nó không giải quyết được câu hỏi ai chọn tên chuẩn.

Khi thực thể chuẩn được xác lập, lớp hạ tầng tiếp theo mới dựng được. Đó là đồ thị tri thức (knowledge graph), nơi địa phương có quan hệ với các thực thể khác: Huế nối với Đại Nội, sông Hương, Festival Huế, ga Huế, sân bay Phú Bài và hàng nghìn cơ sở dịch vụ. Khi các quan hệ ấy hội tụ về một thực thể trung tâm, máy có thể suy luận đầy đủ hơn thay vì chỉ dựa vào những kết quả tra cứu rời rạc.

Điều này cũng lý giải vì sao hai địa phương có lượng thông tin tương đương lại có khả năng được đề xuất rất khác nhau. Nơi có nhiều dữ liệu hơn không đương nhiên thắng. Một lượng dữ liệu lớn nhưng phân tán có thể kém giá trị hơn một tập dữ liệu nhỏ hơn nhưng cùng hội tụ về một thực thể thống nhất.

> ▪ ***Ở mức tối thiểu, một danh tính địa phương đủ dùng cho máy cần sáu thành phần.***
> • *Định danh ổn định: một mã không thay đổi khi tên gọi hoặc đơn vị chủ quản thay đổi và khi thay đổi tạo ra thực thể mới thì mã mới được nối với mã tiền nhiệm.*
> • *Tên chuẩn: cách gọi được hệ thống dùng làm điểm quy chiếu, không phải cách gọi duy nhất được phép dùng.*
> • *Bảng tên thay thế: mọi cách gọi khác, kèm ngôn ngữ và ngữ cảnh sử dụng.*

> • *Phạm vi: thực thể này là đơn vị hành chính, khu di sản, cụm trải nghiệm hay một điểm cụ thể.*
> • *Thời gian hiệu lực: mỗi tên và mỗi phạm vi có ngày bắt đầu và ngày kết thúc.*
> • *Quan hệ kế thừa: thực thể tiền nhiệm và thực thể kế nhiệm được nối bằng liên kết tường minh.*

Sáu thành phần này không đòi hỏi công nghệ mới. Chúng đòi hỏi một quyết định quản trị: ai giữ danh mục, ai được sửa và sửa theo quy trình nào. Toàn bộ phần còn lại của chương là hệ quả của việc danh mục ấy có tồn tại hay không.

Đối với Việt Nam, đây là điểm khởi đầu của toàn bộ chương trình xây dựng khả đọc. Trước khi nói đến API, song sinh số, dữ liệu thời gian thực hay giao dịch bằng tác nhân, địa phương phải trả lời được câu hỏi cơ bản nhất: máy có biết tất cả những tên gọi khác nhau ấy đang chỉ cùng một nơi hay không? Nếu câu trả lời vẫn là "không chắc", thì mọi lớp hạ tầng phía sau đều sẽ được xây dựng trên một nền móng chưa ổn định.

Đó cũng là kết luận quan trọng nhất của mục này. Vấn đề đầu tiên của Việt Nam không phải AI hiểu sai địa phương, mà là AI chưa chắc đang đọc đúng một thực thể. Khi danh tính số còn bị phân mảnh, niềm tin thuật toán cũng bị phân mảnh. Và khi niềm tin chưa thể tích lũy trên một thực thể thống nhất, địa phương rất khó trở thành một chủ thể đủ mạnh trong nền kinh tế tác nhân. Đây là lý do vì sao phân giải thực thể không phải là một chi tiết kỹ thuật của ngành dữ liệu, mà là nền móng của năng lực cạnh tranh địa phương trong kỷ nguyên AI. Máy không hỏi một địa phương đẹp đến đâu. Nó hỏi địa phương ấy có phải là một hay không.

## 8.2. ĐỊA DANH SAU SÁP NHẬP: KHI BẢN ĐỒ ĐỔI MÀ DỮ LIỆU KHÔNG ĐỔI

Đối với con người, việc thay đổi tên một đơn vị hành chính thường chỉ là thay đổi trên giấy tờ hoặc biển chỉ dẫn. Sau một thời gian, người dân quen với tên mới và cuộc sống tiếp tục diễn ra bình thường. Nhưng đối với máy, mỗi lần địa giới thay đổi đều có thể tạo ra một đứt gãy trong chuỗi nhận diện thực thể. Một địa danh vừa có tên cũ, vừa có tên mới, vừa tồn tại trong nhiều cơ sở dữ liệu chưa được cập nhật. Điều này khiến cùng một nơi có thể xuất hiện dưới nhiều danh tính khác nhau.

Về mặt quản lý nhà nước, đây là quá trình bình thường nhằm nâng cao hiệu quả hành chính. Tuy nhiên, trong thế giới số, mỗi lần thay đổi lại tạo ra một khoảng trễ giữa bản đồ pháp lý và bản đồ dữ liệu. Trong khoảng thời gian đó, các hệ thống khác nhau có thể cùng lúc sử dụng những phiên bản địa danh khác nhau.

Khoảng trễ này không chỉ xuất hiện trên các trang web địa phương. Nó lan sang bản đồ số, nền tảng đặt dịch vụ, hệ thống logistics, cơ sở dữ liệu doanh nghiệp, mạng xã hội, kho tri thức mở và cả tập dữ liệu mà các mô hình AI sử dụng để huấn luyện hoặc truy hồi thông tin. Khi các nguồn không còn đồng bộ, máy phải đối mặt với một câu hỏi tưởng như rất đơn giản nhưng lại khó trả lời: đây là một địa phương mới hay chỉ là tên mới của cùng một thực thể?

Hệ quả không dừng ở việc hiển thị sai tên. Quan trọng hơn, lịch sử của một địa phương có thể bị chia cắt thành nhiều phần. Đánh giá của du khách nằm dưới tên cũ, dữ liệu lưu trú nằm dưới mã hành chính cũ, còn thông tin chính thức lại xuất hiện dưới tên mới. Nếu hệ thống không hợp nhất được các phiên bản này,

AI sẽ coi đó là những thực thể khác nhau. Niềm tin vốn được tích lũy trong nhiều năm bị phân mảnh chỉ vì dữ liệu không theo kịp sự thay đổi của địa giới.

Đối với con người, việc biết rằng "xã A nay thuộc xã B" khá dễ dàng. Đối với máy, điều đó chỉ đúng khi có một chuỗi ánh xạ (mapping) rõ ràng giữa tên cũ, tên mới, mã định danh cũ và mã định danh mới. Nếu chuỗi liên kết này bị thiếu, mọi thay đổi hành chính đều có nguy cơ trở thành một điểm đứt trong đồ thị tri thức của địa phương.

Vì vậy, thách thức của Việt Nam không chỉ là cập nhật bản đồ sau mỗi lần sáp nhập. Thách thức lớn hơn là duy trì tính liên tục của thực thể số (entity continuity) cho từng địa danh. Trong nền kinh tế tác nhân, một địa phương không chỉ cần tồn tại trên bản đồ hành chính hiện tại mà còn phải duy trì được mối liên hệ với toàn bộ lịch sử dữ liệu của chính mình. Nếu không, mỗi lần định danh trên bản đồ thay đổi cũng có thể đồng nghĩa với việc máy phải "làm quen lại từ đầu" với một địa phương mà trên thực tế chưa từng biến mất.

Độ trễ giữa hiệu lực pháp lý và hiệu lực dữ liệu. Một quyết định hành chính có hiệu lực tức thời trong hệ thống pháp luật nhưng lan truyền dần trong hệ thống dữ liệu. Độ trễ tăng theo số hệ thống độc lập đang lưu thực thể đó và giảm khi có một nguồn công bố duy nhất mà máy đọc được. Trong khoảng trễ, cùng một nơi tồn tại đồng thời dưới nhiều danh tính đều hợp lý theo dữ liệu của hệ thống lưu nó.

Một di sản đổi tên gọi quốc tế trong cùng tháng. Ngày 13 tháng 7 năm 2025, tại kỳ họp lần thứ 47 của Ủy ban Di sản Thế giới, UNESCO phê duyệt điều chỉnh ranh giới Di sản Thiên nhiên

Thế giới Vườn quốc gia Phong Nha - Kẻ Bàng, mở rộng sang Vườn quốc gia Hin Nam Nô của tỉnh Khăm Muộn, Lào. Tên gọi của di sản sau điều chỉnh là Vườn quốc gia Phong Nha - Kẻ Bàng và Vườn quốc gia Hin Nam Nô. Đây là di sản thiên nhiên liên biên giới đầu tiên giữa Việt Nam và Lào. Trong các văn bản công bố sau ngày 01 tháng 7 năm 2025, di sản trên biên giới này được ghi thuộc tỉnh Quảng Trị (Việt Nam) và tỉnh Khăm Muộn (Lào), không còn ghi thuộc tỉnh Quảng Bình.[1]

Phong Nha, tháng 7 năm 2025: ba sự kiện, bốn thuộc tính định danh thay đổi. Ngày 01 tháng 7 năm 2025, ba thay đổi có hiệu lực cùng lúc với một địa điểm duy nhất. Tỉnh Quảng Bình, đơn vị chủ quản của Vườn quốc gia Phong Nha - Kẻ Bàng, hợp nhất với tỉnh Quảng Trị và mang tên tỉnh Quảng Trị. Đơn vị cấp huyện quản lý địa bàn kết thúc hoạt động. Trung tâm du lịch hang động trở thành một đơn vị hành chính cấp xã mang tên xã Phong Nha. Mười hai ngày sau, ngày 13 tháng 7 năm 2025, UNESCO phê duyệt điều chỉnh ranh giới di sản, mở rộng sang Vườn quốc gia Hin Nam Nô của Lào và tên gọi quốc tế của di sản thay đổi theo. Trong mười ba ngày, ba sự kiện đã làm thay đổi bốn thuộc tính định danh của cùng một khối núi đá vôi: tỉnh chủ quản, cấp hành chính trung gian, tên đơn vị cơ sở và tên gọi quốc tế của di sản. Không thuộc tính nào trong số đó liên quan tới bản thân tài sản. Hang động không dịch chuyển, hệ sinh thái không đổi, đường vào không thay đổi. Nhưng với những tập dữ liệu đang mô tả nơi này bằng cụm từ Phong Nha - Kẻ Bàng, tỉnh Quảng Bình, thì kể từ ngày 01 tháng 7 năm 2025 mô tả ấy không còn khớp với thực tế hành chính. Có thể giả định số tập dữ liệu và nội dung vẫn dùng mô tả cũ là đáng kể và không có cơ chế nào tự cập nhật chúng.

1. Cổng Thông tin điện tử Chính phủ và Nghị quyết 202/2025/QH15. Tỉnh Quảng Bình hợp nhất vào tỉnh Quảng Trị từ ngày 1 tháng 7 năm 2025.

*Nguồn và mức xác thực:* Tình huống dựng lại từ các văn bản và công bố chính thức đã dẫn ở các hộp dữ liệu của chương này. Các mốc thời gian và tên gọi là có thật. Phần suy luận về hệ quả đối với dữ liệu đang lưu hành là phân tích của tác giả, không dựa trên đo lường.

### 8.2.1. Địa giới đổi nhanh hơn internet

Trong nhiều thập niên, internet được xem là nơi lưu giữ thông tin nhanh nhất. Một văn bản được công bố hôm nay có thể xuất hiện trên hàng nghìn trang web chỉ sau vài giờ. Nhưng nghịch lý của thời đại AI là thông tin có thể lan rất nhanh, còn việc sửa thông tin đã lan đi lại diễn ra rất chậm. Điều này đặc biệt rõ với những thay đổi về địa giới hành chính.

Khi một xã được sáp nhập, một đơn vị cấp huyện kết thúc hoạt động hay một tỉnh hình thành sau khi hợp nhất, bản đồ pháp lý thay đổi ngay từ thời điểm văn bản có hiệu lực. Trong khi đó, thế giới số không có một nút "cập nhật đồng loạt". Nhiều cơ sở dữ liệu độc lập vẫn tiếp tục sử dụng tên cũ cho đến khi từng hệ thống tự cập nhật theo chu kỳ riêng của mình.

Một trang web du lịch có thể đã đổi sang tên mới. Nhưng nền tảng đặt phòng vẫn dùng địa chỉ cũ. Bản đồ số cập nhật theo một chu kỳ, hệ thống logistics theo một chu kỳ khác. Một doanh nghiệp chưa kịp sửa giấy phép điện tử. Các bài báo cũ vẫn được Google lập chỉ mục. Những bài viết trên mạng xã hội vẫn nhắc đến tên cũ. Trong kho dữ liệu mà các mô hình AI truy hồi, cả hai phiên bản cùng tồn tại song song.

Kết quả là trong một khoảng thời gian khá dài, internet tồn tại hai bản đồ của cùng một quốc gia: một bản đồ pháp lý và một bản đồ dữ liệu. Con người thường dễ dàng hiểu rằng đây chỉ là

hai cách gọi của cùng một nơi. Máy cũng có thể suy luận như vậy, nhưng chỉ khi có đủ ngữ cảnh và bằng chứng liên kết và nó không mặc nhiên duy trì được quan hệ ấy trong dữ liệu nếu không được cung cấp mối liên kết giữa chúng.

Đó là lý do một truy vấn rất đơn giản như tìm khách sạn, nhà hàng hay điểm du lịch cũng có thể cho ra những kết quả không nhất quán. Một phần dữ liệu được trả về theo tên mới, phần còn lại theo tên cũ. Đánh giá của người dùng, lịch sử giao dịch, tọa độ bản đồ và thông tin chính thức bị phân tán sang nhiều thực thể khác nhau. Rủi ro lớn hơn thường không nằm ở việc xuất hiện dữ liệu sai mà ở dữ liệu đúng nhưng không còn được kết nối.

Hiện tượng này phản ánh một quy luật quan trọng của hạ tầng số: internet không vận hành như một cơ sở dữ liệu tập trung mà như một hệ sinh thái gồm vô số bản sao độc lập. Mỗi hệ thống có lịch cập nhật, tiêu chuẩn dữ liệu và quy trình đồng bộ riêng. Vì vậy, một quyết định hành chính có thể đổi bản đồ pháp lý nhanh hơn nhiều so với tốc độ đồng bộ của toàn bộ internet.

Trong kỷ nguyên tìm kiếm truyền thống, khoảng trễ này chủ yếu gây bất tiện cho người dùng. Nhưng trong nền kinh tế tác nhân, hậu quả lớn hơn nhiều. AI không chỉ đọc thông tin mà còn dựa vào thông tin đó để lập kế hoạch, xếp hạng, đề xuất và thực hiện giao dịch. Nếu các nguồn dữ liệu mâu thuẫn về chính danh của một địa phương, mức độ chắc chắn của mô hình sẽ giảm xuống. Thay vì nhận diện một thực thể thống nhất, AI nhìn thấy nhiều thực thể có mức độ tương đồng cao nhưng không đủ bằng chứng để khẳng định chúng là một.

Có thể nói, địa giới hành chính thay đổi trong một quyết định của Nhà nước, nhưng địa giới dữ liệu chỉ thay đổi khi các

hệ thống phân tán lần lượt hoàn tất việc cập nhật. Khoảng cách giữa hai quá trình ấy chính là vùng rủi ro mà Việt Nam phải đối mặt trong quá trình chuyển đổi sang nền kinh tế tác nhân.

Bản đồ pháp lý và bản đồ dữ liệu. Bản đồ pháp lý đổi tại đúng ngày văn bản có hiệu lực. Nó là một sự kiện, có thời điểm xác định và không có trạng thái trung gian. Bản đồ dữ liệu đổi theo tốc độ đồng bộ của từng hệ thống. Nó là một quá trình, kéo dài, không đồng đều và trong suốt quá trình đó nhiều phiên bản của cùng một nơi cùng tồn tại. Phần lớn khó khăn được mô tả trong mục này nằm ở khoảng cách giữa một sự kiện và một quá trình. Khoảng cách ấy không nhất thiết bắt nguồn từ lỗi của một cơ quan riêng lẻ, nhưng cũng không cơ quan nào tự nhiên chịu trách nhiệm thu hẹp nó.

Đây cũng là lý do sau mỗi đợt điều chỉnh địa giới, công việc quan trọng không chỉ là đổi biển tên hay cập nhật văn bản hành chính. Quan trọng hơn là nhanh chóng xây dựng lớp ánh xạ thực thể (entity mapping layer), liên kết toàn bộ tên cũ, tên mới, mã định danh cũ, mã định danh mới, tọa độ và lịch sử của từng địa phương. Chỉ khi lớp liên kết này tồn tại, các hệ thống mới có thể nối đúng thực thể hiện tại với lịch sử của các thực thể tiền nhiệm.

### 8.2.2. Ba lớp dữ liệu của một tỉnh mới

Mỗi lần sáp nhập tỉnh, điều được tạo ra không chỉ là một đơn vị hành chính mới mà còn là một bài toán dữ liệu mới. Trên giấy tờ, tỉnh mới có thể chính thức tồn tại từ một ngày xác định. Nhưng trên internet, "tỉnh mới" không bắt đầu từ con số không. Nó phải kế thừa và hợp nhất nhiều lớp dữ liệu đã được tích lũy trong nhiều năm trước đó.

Đối với AI, một tỉnh sau sáp nhập phải quản lý đồng thời ba lớp dữ liệu: hồ sơ của các đơn vị tiền nhiệm, hồ sơ của thực thể kế nhiệm và mạng lưới tham chiếu cũ gồm địa chỉ web, mã định danh, liên kết và trích dẫn. Lớp thứ nhất và lớp thứ ba là di sản theo đúng nghĩa, lớp thứ hai là phần mới hình thành. Lớp thứ ba không phải một lịch sử ngang cấp với hai lớp đầu. Nó là hạ tầng trỏ đường và thường có độ bền cao vì các liên kết và trích dẫn đã phát tán không tự biến mất khi đơn vị hành chính thay đổi.

Lớp thứ nhất là hồ sơ của các đơn vị tiền nhiệm. Mỗi tỉnh cũ đã có hồ sơ riêng trong hàng nghìn hệ thống: đồ thị tri thức, bản đồ số, cơ sở dữ liệu doanh nghiệp, nền tảng du lịch, báo chí và kho dữ liệu mở. Mỗi hồ sơ tích lũy lượt tìm kiếm, đánh giá, hình ảnh, bài viết, liên kết và dữ liệu giao dịch. Đó là vốn niềm tin mà máy đã học trong nhiều năm.

Lớp thứ hai là hồ sơ của thực thể kế nhiệm, tức thực thể hành chính mới hình thành sau quyết định sáp nhập (merged entity). Tỉnh mới có tên mới, mã hành chính mới và phạm vi địa lý mới. Tuy nhiên, trong giai đoạn đầu, thực thể này gần như chưa có lịch sử số của riêng mình. Nếu các hệ thống không kế thừa dữ liệu từ các tỉnh cũ, một số hệ thống có thể xử lý đây như một thực thể mới với rất ít bằng chứng để đánh giá. Điều đó tạo ra nghịch lý: một địa phương có hàng trăm nghìn dân và nhiều thập niên lịch sử lại xuất hiện trước máy như một thực thể vừa được sinh ra.

Lớp thứ ba là mạng lưới tham chiếu cũ (legacy references), di sản của internet. Hàng nghìn địa chỉ trang web, đường dẫn báo điện tử, trang giới thiệu điểm đến, hồ sơ doanh nghiệp, URL bản đồ và liên kết trên mạng xã hội vẫn chứa tên của các tỉnh cũ. Chúng tiếp tục tồn tại, được lập chỉ mục, được chia sẻ và được AI truy hồi. Internet gần như không bao giờ quên. Một URL đã được

trích dẫn hàng nghìn lần có thể tiếp tục xuất hiện nhiều năm sau khi địa danh chính thức đổi tên.

Ba lớp dữ liệu này không biến mất khi có quyết định hành chính mới. Chúng cùng tồn tại và ảnh hưởng lẫn nhau. Nếu thiếu cơ chế hợp nhất, các hệ thống có thể nhìn thấy nhiều hồ sơ, định danh và tham chiếu có quan hệ rất gần nhưng không đủ dữ liệu để nối chúng thành một chuỗi kế thừa chính xác. Một bài báo dùng tên tỉnh cũ, một trang web dùng tên tỉnh mới, còn dữ liệu bản đồ lại liên kết theo mã hành chính khác. Mỗi nguồn đều đúng trong bối cảnh của nó, nhưng toàn bộ hệ thống lại thiếu một cầu nối nhận dạng.

Đây chính là bản chất của bài toán chuyển đổi và kế thừa thực thể (entity migration). Sáp nhập hành chính không chỉ là gộp diện tích hay dân số, mà còn là chuyển toàn bộ lịch sử dữ liệu của nhiều thực thể sang một thực thể mới mà không làm mất tính liên tục. Nếu quá trình này thất bại, mọi tín hiệu tích lũy trước đây sẽ bị phân mảnh.

Có thể hình dung như hai doanh nghiệp hợp nhất thành một pháp nhân mới nhưng lịch sử tín dụng, hồ sơ khách hàng và toàn bộ đánh giá không được nối sang pháp nhân kế nhiệm. Trên giấy tờ, pháp nhân mới đã tồn tại. Với hệ thống chấm điểm, đó là một chủ thể chưa có quá khứ. Địa phương sau sáp nhập đối mặt với đúng nguy cơ ấy trong thế giới số.

Vì vậy, sau mỗi lần điều chỉnh địa giới, công việc quan trọng không chỉ là công bố tên mới. Quan trọng hơn là xây dựng một chuỗi kế thừa thực thể (entity lineage): thực thể nào là tiền nhiệm, thực thể nào là kế nhiệm, địa chỉ web nào được chuyển hướng, mã định danh nào được thay thế và quan hệ nào cần

giữ trong đồ thị tri thức. Cần nói chính xác ở đây: hai tỉnh cũ hợp nhất thành một tỉnh mới không phải là ba giai đoạn của cùng một thực thể, vì phạm vi, thẩm quyền và danh tính pháp lý đều đã thay đổi. Đó là quan hệ kế thừa giữa những thực thể khác nhau. Chính vì khác nhau nên liên kết kế thừa phải được ghi tường minh, không thể để máy tự suy ra.

Tỉnh Quảng Trị sau hợp nhất. Tỉnh Quảng Trị hiện nay hình thành từ việc sắp xếp toàn bộ diện tích tự nhiên và quy mô dân số của tỉnh Quảng Bình và tỉnh Quảng Trị thành một tỉnh mới mang tên tỉnh Quảng Trị. Sau khi sắp xếp, tỉnh Quảng Trị có diện tích tự nhiên 12.700 $km^2$ và quy mô dân số 1.870.845 người. Cả hai con số này đều nằm trong chính văn bản pháp luật, không phải số liệu tổng hợp lại.

Về mặt pháp lý, việc sáp nhập đã hoàn tất. Tuy nhiên, về mặt dữ liệu, đây mới là giai đoạn bắt đầu bộc lộ những bất cập. Các tập dữ liệu được tạo trước tháng 7 năm 2025 không tự cập nhật và một phần trong số đó đã đi vào dữ liệu huấn luyện, vào chỉ mục của các nền tảng, vào bài viết của hàng nghìn trang web sẽ tồn tại nhiều năm nữa. Mỗi lần một hệ thống mới lấy dữ liệu từ những nguồn đó, tên cũ lại được tái sinh. Đó là lý do hợp nhất thực thể không phải một dự án có ngày kết thúc mà là một chức năng phải duy trì.

### 8.2.3. Bản đồ, OTA và AI không cập nhật cùng nhịp

Một thay đổi hành chính không đi vào toàn bộ hệ sinh thái số theo một con đường duy nhất. Nó phải đi qua hàng trăm nền tảng khác nhau, mỗi nền tảng có cơ chế thu thập, kiểm duyệt và cập nhật dữ liệu riêng. Vì vậy, sau mỗi lần đổi tên hoặc sáp nhập

địa giới, điều thường xảy ra là Google Maps, các OTA và các hệ thống AI không cập nhật cùng một thời điểm.

Đây không phải là lỗi của một nền tảng cụ thể mà là đặc tính của internet phân tán. Mỗi hệ thống có nguồn dữ liệu, chu kỳ đồng bộ và tiêu chuẩn xác minh khác nhau.

Google Maps có thể cập nhật khá nhanh nhờ dữ liệu từ cơ quan quản lý, đối tác bản đồ và phản hồi của cộng đồng. Một nền tảng OTA như Booking, Agoda hay Traveloka lại phụ thuộc vào dữ liệu do từng khách sạn, đối tác hoặc hệ thống quản trị điểm lưu trú cung cấp. Trong khi đó, các mô hình AI còn phức tạp hơn. Chúng không chỉ đọc dữ liệu thời gian thực mà còn dựa vào kho tri thức đã tích lũy, dữ liệu huấn luyện từ quá khứ và nhiều nguồn truy hồi khác nhau. Điều đó có nghĩa là ngay cả khi bản đồ đã đổi tên, AI vẫn có thể tiếp tục nhìn thấy rất nhiều tài liệu sử dụng tên cũ.

Khoảng thời gian này tạo ra một hiện tượng đặc biệt: cùng một địa điểm nhưng mỗi hệ thống lại kể một câu chuyện khác nhau.

Một du khách tìm kiếm trên Google Maps có thể thấy tên tỉnh mới. Khi chuyển sang OTA để đặt phòng, địa chỉ khách sạn vẫn ghi theo tỉnh cũ. Sau đó, hỏi AI về lịch trình, hệ thống lại sử dụng cả hai tên trong cùng một câu trả lời vì nó tổng hợp từ nhiều nguồn khác nhau. Không có thông tin nào hoàn toàn sai, nhưng toàn bộ trải nghiệm trở nên thiếu nhất quán.

Đối với con người, sự không nhất quán này chỉ gây nhầm lẫn. Đối với máy, nó làm giảm mức chắc chắn mà hệ thống gán cho thực thể, một đại lượng mà chương này gọi là độ tin cậy của thực thể (entity confidence). Đây là một khái niệm phân tích, không

phải khẳng định rằng mọi hệ thống đều có một biến số mang đúng tên ấy. Khi nhiều nguồn uy tín cùng mô tả một địa điểm theo các cách khác nhau, AI phải dành thêm tài nguyên để xác minh xem chúng có thực sự là một hay không. Nếu bằng chứng liên kết không đủ mạnh, hệ thống sẽ giảm mức độ chắc chắn trong quá trình suy luận và đề xuất.

Các tầng dữ liệu này còn phụ thuộc lẫn nhau. Nhiều nền tảng AI sử dụng dữ liệu từ bản đồ, từ trang web chính thức, từ OTA và từ các kho tri thức mở để kiểm chứng chéo. Nếu một tầng đã cập nhật nhưng các tầng còn lại vẫn giữ phiên bản cũ, quá trình xác nhận lẫn nhau bị suy yếu. Thay vì tạo ra nhiều bằng chứng thống nhất, internet lại tạo ra nhiều bằng chứng mâu thuẫn.

Trong nền kinh tế tìm kiếm, sự chậm trễ này chủ yếu ảnh hưởng đến trải nghiệm tra cứu. Trong nền kinh tế tác nhân, cái giá của độ trễ ấy đắt hơn. Một AI không chỉ trả lời câu hỏi mà còn phải lựa chọn khách sạn, lên hành trình, đặt dịch vụ và thực hiện giao dịch. Mọi hành động đều dựa trên giả định rằng các nguồn dữ liệu đang nói về cùng một thực thể. Khi giả định đó không còn chắc chắn, hệ thống có thể phải xác minh thêm, giảm mức ưu tiên, hỏi lại người dùng hoặc loại thực thể khỏi một số bước có rủi ro cao.

Hệ quả là bài toán sau sáp nhập không phải việc cập nhật một nền tảng duy nhất, mà là đồng bộ toàn bộ hệ sinh thái dữ liệu. Chính quyền có thể ban hành quyết định trong một ngày, nhưng giá trị của quyết định đó trong thế giới số chỉ được hoàn thành khi bản đồ, trang web chính thức, cơ sở dữ liệu doanh nghiệp, OTA, mạng xã hội, đồ thị tri thức và các hệ thống AI đều cùng hội tụ về một thực thể thống nhất.

Vì vậy, trong kỷ nguyên AI, tốc độ cập nhật không còn được đo bằng việc nền tảng nào cập nhật trước, mà bằng khoảng thời gian cần thiết để toàn bộ hệ sinh thái đạt được sự nhất quán. Chỉ khi sự nhất quán đó được thiết lập, địa phương mới thực sự hoàn tất quá trình chuyển đổi từ địa giới hành chính mới sang danh tính số mới.

### 8.2.4. Một điểm đến sống trong nhiều bản đồ

Con người thường nghĩ một địa phương chỉ có một bản đồ. Nhưng đối với AI, một điểm đến đồng thời tồn tại trong nhiều "bản đồ" khác nhau, mỗi bản đồ được xây dựng cho một mục đích riêng. Sau mỗi lần điều chỉnh địa giới, những bản đồ này không nhất thiết thay đổi cùng lúc. Chính sự lệch pha đó tạo ra nhiều vấn đề nhận diện mà con người hiếm khi để ý.

Bản đồ đầu tiên là bản đồ hành chính (administrative map). Đây là bản đồ do Nhà nước xác lập, phản ánh ranh giới pháp lý, tên đơn vị hành chính, mã địa giới và cơ cấu quản lý. Khi một tỉnh được sáp nhập hoặc một xã đổi tên, bản đồ này thay đổi ngay theo quyết định pháp luật. Đây là "sự thật pháp lý" của địa phương.

Tuy nhiên, khách du lịch hiếm khi lựa chọn điểm đến theo bản đồ hành chính. Họ sống trong bản đồ du lịch (tourism map). Trên bản đồ này, ranh giới được quyết định bởi trải nghiệm chứ không phải bởi địa giới. Một bãi biển, một vườn quốc gia, một tuyến trekking hay một khu phố ẩm thực có thể trải dài qua nhiều đơn vị hành chính nhưng trong nhận thức của du khách vẫn là một điểm đến duy nhất. Ngược lại, một tỉnh mới sau sáp nhập có thể bao gồm nhiều vùng du lịch hoàn toàn khác biệt, mỗi vùng đã có thương hiệu riêng từ trước.

Song song với đó là bản đồ thương mại (commercial map). Đây là bản đồ mà các doanh nghiệp, nền tảng đặt phòng, hệ thống vận chuyển, logistics, thanh toán và quảng cáo sử dụng. Ranh giới của bản đồ này được hình thành theo mạng lưới kinh doanh, vùng giao hàng, thị trường tiêu dùng, khu vực bán vé hay phạm vi hoạt động của từng doanh nghiệp. Không ít doanh nghiệp vẫn duy trì cấu trúc vùng kinh doanh cũ trong nhiều năm sau khi địa giới hành chính đã thay đổi.

Ba bản đồ này phản ánh ba cách nhìn hoàn toàn khác nhau về cùng một địa phương.

- ***Bản đồ hành chính trả lời:*** địa phương này thuộc đơn vị quản lý nào?

- ***Bản đồ du lịch trả lời:*** du khách nhận diện nơi này như một trải nghiệm gì?

- ***Bản đồ thương mại trả lời:*** hệ thống giao dịch và doanh nghiệp đang vận hành nơi này như thế nào?

Trong phần lớn thời gian, ba bản đồ tương đối đồng nhất nên chúng ta ít nhận ra sự khác biệt. Nhưng sau một đợt sáp nhập hoặc đổi tên, chúng có thể tách khỏi nhau trong nhiều tháng, thậm chí nhiều năm. Chính quyền đã sử dụng tên mới, ngành du lịch vẫn quảng bá thương hiệu cũ, còn các OTA và nền tảng thương mại vẫn giữ nguyên cấu trúc dữ liệu trước đây. Một địa phương lúc này không còn có một danh tính duy nhất mà tồn tại đồng thời trong nhiều hệ quy chiếu.

Đối với AI, đây không chỉ là vấn đề tên gọi mà là vấn đề ngữ cảnh (Context). Khi người dùng hỏi "đến đâu để nghỉ dưỡng cuối

tuần", hệ thống cần hiểu bản đồ du lịch. Khi tính thuế hoặc xác minh doanh nghiệp, hệ thống phải dùng bản đồ hành chính. Khi đặt khách sạn hay giao hàng, AI lại phải dựa trên bản đồ thương mại. Nếu không phân biệt được ba lớp bản đồ này, AI rất dễ suy luận sai hoặc kết nối sai thực thể.

Các tác nhân phải liên tục chuyển đổi giữa ba bản đồ. Một tác nhân du lịch có thể bắt đầu từ bản đồ du lịch để chọn điểm đến, chuyển sang bản đồ thương mại để kiểm tra tồn kho, giá và khả năng đặt dịch vụ, rồi quay về bản đồ hành chính để xác minh địa chỉ, quy định hoặc giấy phép. Nếu ba lớp dữ liệu này không liên kết chặt chẽ, chuỗi hành động sẽ bị đứt gãy ở một trong các bước. Vì vậy, mục tiêu của chuyển đổi số địa phương không phải là làm cho mọi bản đồ trở nên giống nhau. Điều đó cũng không cần thiết. Điều quan trọng là xây dựng các liên kết ngữ nghĩa giữa các bản đồ. Một điểm đến phải có khả năng được nhận diện là cùng một thực thể dù xuất hiện dưới góc nhìn quản lý nhà nước, trải nghiệm du lịch hay giao dịch thương mại.

Địa phương không còn sống trên một bản đồ duy nhất. Địa phương tồn tại đồng thời trong nhiều lớp bản đồ số. Khả năng kết nối các lớp đó mới quyết định liệu AI có nhìn thấy một điểm đến thống nhất hay chỉ là nhiều mảnh dữ liệu rời rạc. Địa giới đổi trong một ngày. Dữ liệu đổi trong nhiều năm. Khoảng cách giữa hai mốc ấy là nơi một địa phương có thể biến mất.

### 8.2.5. Sau sáp nhập, việc đầu tiên không phải làm logo mà hợp nhất thực thể số

Sau mỗi lần sáp nhập, phản ứng đầu tiên của nhiều địa phương thường là xây dựng bộ nhận diện mới: thiết kế logo, khẩu hiệu,

chiến dịch truyền thông và video quảng bá. Đây là những việc cần thiết để hình thành một hình ảnh thống nhất trong nhận thức của con người. Tuy nhiên, trong nền kinh tế tác nhân, đó không nên là ưu tiên đầu tiên.

Đối với AI, một logo mới gần như không có giá trị nếu hệ thống vẫn chưa biết thực thể mới là ai. Thứ mà các hệ thống AI cần trước hết không phải một biểu tượng đẹp hơn mà là một thực thể số thống nhất (unified digital entity). Trước khi kể câu chuyện mới, địa phương phải bảo đảm rằng toàn bộ dữ liệu cũ vẫn được nối liền với danh tính mới.

Hãy hình dung một tỉnh A và tỉnh B sáp nhập thành tỉnh C. Nếu ngay lập tức triển khai bộ thương hiệu mới nhưng không hợp nhất dữ liệu, internet sẽ đồng thời tồn tại ba hồ sơ khác nhau: hồ sơ của A, hồ sơ của B và hồ sơ của C. Đối với người dân, đó là một tỉnh duy nhất. Đối với AI, đó có thể vẫn là ba thực thể độc lập.

Khi điều này xảy ra, mọi tài sản số tích lũy trước đây bắt đầu bị chia nhỏ.

- *Hàng chục nghìn đánh giá của du khách vẫn nằm dưới tên các tỉnh cũ.*
- *Hàng triệu liên kết từ báo chí vẫn trỏ về URL cũ.*
- *Khách sạn, nhà hàng và điểm tham quan vẫn gắn với địa chỉ cũ.*
- *Các đồ thị tri thức vẫn lưu quan hệ theo thực thể cũ.*
- *Hệ thống OTA tiếp tục phân loại theo vùng dữ liệu cũ.*
- *AI truy hồi từ nhiều nguồn khác nhau và nhận được những thực thể không hoàn toàn trùng khớp.*

Nếu không có quá trình hợp nhất, tỉnh mới giống như một thương hiệu vừa ra đời nhưng không được kế thừa bất kỳ uy tín nào của các thương hiệu tiền nhiệm.

Đó là lý do trong chuyển đổi số, việc quan trọng nhất sau sáp nhập không phải là làm lại thương hiệu mà là hợp nhất thực thể (entity consolidation).

Hợp nhất thực thể không chỉ là đổi tên trong một cơ sở dữ liệu. Đó là quá trình bảo đảm các hệ thống liên quan đều hiểu rằng:

- *Thực thể mới kế thừa toàn bộ lịch sử của các thực thể cũ.*
- *Mọi mã định danh cũ đều ánh xạ sang mã định danh mới.*
- *Các URL cũ được chuyển hướng và duy trì liên kết.*
- *Các thực thể trong đồ thị tri thức được hợp nhất thay vì tạo mới.*
- *Toàn bộ đánh giá, hình ảnh, dữ liệu giao dịch và quan hệ ngữ nghĩa tiếp tục thuộc về cùng một vòng đời dữ liệu.*

Nói cách khác, địa phương không được để AI "mất trí nhớ" chỉ vì thay đổi địa giới hành chính. Đây là điểm khác biệt căn bản giữa xây dựng thương hiệu trong thời đại truyền thông và trong thời đại AI. Trước đây, thương hiệu được bắt đầu bằng hình ảnh. Ngày nay, thương hiệu số được bắt đầu bằng danh tính dữ liệu. Chỉ khi thực thể số đã được hợp nhất, mọi hoạt động truyền thông mới thực sự kế thừa được uy tín, dữ liệu và niềm tin đã tích lũy trong nhiều năm.

Có thể xem quá trình sau sáp nhập gồm hai giai đoạn. Giai đoạn một là hợp nhất thực thể, giai đoạn hai mới là xây dựng thương hiệu. Nếu làm ngược lại, địa phương sẽ có một hình ảnh mới nhưng một lịch sử dữ liệu bị phân mảnh. Đối với AI, điều đó đồng nghĩa với việc phải xây dựng lại niềm tin gần như từ đầu.

Chi phí hợp nhất. Chi phí hợp nhất thực thể ngay sau một lần thay đổi địa giới thấp hơn nhiều lần chi phí xây lại danh tiếng số sau khi dữ liệu đã phân mảnh vài năm. Giả thuyết này hợp với logic của ký ức thuật toán ở Chương 4, nhưng cho đến nay tác giả chưa tìm được bộ dữ liệu công khai nào của Việt Nam cho phép ước lượng hai khoản chi phí ấy bằng con số. Vì vậy nó được đặt ở mức giả thuyết, không phải quy luật.

Vì vậy, trong nền kinh tế tác nhân, câu hỏi đầu tiên sau mỗi lần sáp nhập không nên là "logo mới sẽ như thế nào?" mà phải là "AI đã hiểu đây vẫn là cùng một thực thể chưa?". Khi câu hỏi thứ hai chưa có lời giải, mọi nỗ lực truyền thông của câu hỏi thứ nhất đều khó phát huy hết giá trị.

Nếu bạn là chủ tịch Ủy ban nhân dân tỉnh. Yêu cầu một danh mục thực thể điểm đến cấp tỉnh với định danh ổn định, trước khi phê duyệt bất kỳ nền tảng du lịch số mới nào. Chỉ định một đơn vị duy nhất chịu trách nhiệm công bố bảng ánh xạ tên cũ sang tên mới và duy trì nó như một dịch vụ, không như một văn bản. Đặt lớp tích hợp dữ liệu ở cấp tỉnh thành một hạng mục đầu tư riêng, không gộp vào hạng mục cổng thông tin hay ứng dụng du lịch. Yêu cầu báo cáo năm chỉ số dữ liệu theo cách đo tự động, không theo biểu mẫu kê khai.

## 8.3. DỮ LIỆU HÀNH CHÍNH VÀ DỮ LIỆU ĐIỂM ĐẾN

Một trong những nghịch lý lớn nhất của chuyển đổi số ở Việt Nam là nhiều địa phương không thiếu dữ liệu mà thiếu đúng loại dữ liệu mà máy cần. Hệ thống quản lý nhà nước đã tích lũy lượng lớn hồ sơ về dân cư, đất đai, doanh nghiệp, ngân sách, quy hoạch và thủ tục hành chính. Những dữ liệu này rất quan trọng đối với hoạt động điều hành, nhưng chúng được thiết kế để phục vụ quản lý nội bộ chứ không phải để một tác nhân AI có thể hiểu, suy luận và thực hiện hành động.

Trong khi đó, nền kinh tế tác nhân lại cần một lớp dữ liệu hoàn toàn khác. Máy không chỉ cần biết một điểm đến nằm ở đâu, mà còn cần biết nơi đó đang mở cửa hay không, còn bao nhiêu chỗ trống, giá hiện tại là bao nhiêu, có thể đặt trực tiếp hay không, liên kết với những dịch vụ nào, được đánh giá ra sao và phù hợp với nhóm khách nào. Đó là dữ liệu vận hành, dữ liệu ngữ nghĩa và dữ liệu giao dịch, những thành phần quyết định khả năng hành động của AI.

Vì vậy, vấn đề của địa phương không còn là số lượng dữ liệu, mà là sự dịch chuyển từ dữ liệu để quản lý (administrative data) sang dữ liệu để máy có thể hiểu, suy luận và giao dịch (destination data). Đây là bước chuyển từ tư duy "lưu trữ thông tin" sang tư duy "xây dựng hạ tầng tri thức", nơi mỗi thực thể đều có đủ ngữ cảnh để trở thành một phần của mạng lưới quyết định của AI.

Bất tương thích mục đích. Dữ liệu được thiết kế để quản lý không tự chuyển thành dữ liệu để giao dịch, vì hai loại trả lời hai câu hỏi khác nhau. Tăng khối lượng số hóa hành chính làm tăng khả năng tra cứu nhưng không tự làm tăng khả hành động.

### 8.3.1. Dữ liệu quản lý không phải dữ liệu quyết định

Trong nhiều năm, chuyển đổi số ở địa phương chủ yếu tập trung vào số hóa hoạt động quản lý. Hệ thống thông tin được xây dựng để phục vụ cấp phép, thống kê, thanh tra, quy hoạch, quản lý dân cư hay điều hành ngân sách. Đây là những dữ liệu thiết yếu đối với bộ máy nhà nước, nhưng mục tiêu của chúng là giúp chính quyền quản lý tốt hơn, chứ không phải giúp người tiêu dùng lựa chọn tốt hơn.

Sự khác biệt này rất quan trọng. Dữ liệu quản lý trả lời những câu hỏi như: doanh nghiệp đã được cấp phép chưa, thuộc ngành nghề nào, địa chỉ đăng ký ở đâu, ai là người đại diện pháp luật. Trong khi đó, dữ liệu điểm đến phải trả lời những câu hỏi hoàn toàn khác: nơi này có phù hợp với gia đình có trẻ nhỏ không, còn chỗ trống hôm nay không, mất bao lâu để di chuyển, giá hiện tại bao nhiêu, có thể đặt ngay hay không, trải nghiệm nổi bật là gì.

Thách thức không phải là thu thập thêm nhiều dữ liệu hành chính, mà là bổ sung một lớp dữ liệu mới phản ánh trạng thái vận hành thực tế của điểm đến. Chỉ khi dữ liệu được thiết kế từ góc nhìn của người sử dụng và của tác nhân AI, địa phương mới có thể tham gia vào chuỗi đề xuất và giao dịch tự động. Một địa phương có thể quản lý rất tốt nhưng vẫn gần như vắng mặt trong tập lựa chọn của máy.

### 8.3.2. Từ hồ sơ pháp lý đến hồ sơ trải nghiệm

Một hồ sơ hành chính có thể mô tả rất chính xác một tổ chức, nhưng lại nói rất ít về trải nghiệm mà tổ chức đó mang lại. Đây là khoảng cách lớn giữa tư duy quản lý và tư duy thiết kế dữ liệu cho nền kinh tế trải nghiệm.

Hãy hình dung một cơ sở lưu trú trong cơ sở dữ liệu của cơ quan quản lý. Hồ sơ có thể bao gồm tên doanh nghiệp, mã số thuế, địa chỉ, giấy phép hoạt động, số phòng được đăng ký, người đại diện pháp luật và thời hạn giấy phép. Tất cả đều đúng và cần thiết. Tuy nhiên, nếu một tác nhân AI muốn quyết định có nên giới thiệu nơi này cho du khách hay không, gần như mọi thông tin quan trọng đều còn thiếu.

Máy cần biết khách sạn còn bao nhiêu phòng trống trong tối nay, mức giá hiện tại, thời gian nhận phòng, chính sách hủy, chất lượng kết nối internet, khoảng cách đến các điểm tham quan, phương tiện công cộng gần nhất, đánh giá gần đây của khách, hình ảnh mới nhất và khả năng đặt phòng ngay lập tức. Đây đều là những thuộc tính của trải nghiệm, chứ không phải của hồ sơ hành chính.

Nếu địa phương chỉ xây dựng được hồ sơ hành chính mà không xây dựng được hồ sơ trải nghiệm, một tác nhân sẽ biết nơi đó tồn tại nhưng không biết lý do để lựa chọn. Một thực thể chỉ được biết đến mà không thể được lựa chọn hoặc giao dịch sẽ khó tạo ra giá trị trực tiếp trong chuỗi quyết định của tác nhân.

### 8.3.3. Máy cần gì hơn mã xã

Đối với hệ thống quản lý hành chính, một địa điểm được định danh bằng mã tỉnh, mã xã, tọa độ và địa chỉ là đủ để xác định nó thuộc đơn vị nào và chịu sự quản lý của cơ quan nào. Nhưng đối với tác nhân AI, đó mới chỉ là thông tin để xác định vị trí, chưa phải thông tin để ra quyết định.

Khi được giao nhiệm vụ lập kế hoạch cho một chuyến đi, một tác nhân không quan tâm trước tiên đến mã xã của một bảo tàng

hay khu du lịch. Điều nó cần là những dữ liệu phản ánh khả năng phục vụ ngay tại thời điểm truy vấn.

Trước hết là giờ mở cửa (opening hours), gồm cả lịch theo mùa, ngày lễ và cuối tuần. Một bảo tàng nổi tiếng nhưng đóng cửa vào thứ Hai sẽ không được đề xuất cho hành trình diễn ra đúng ngày đó.

Tiếp theo là vé (ticket), gồm giá, loại vé, ưu đãi và khả năng mua trực tuyến. Giá ở đây không chỉ để người dùng tham khảo, nó là dữ liệu đầu vào để tác nhân tối ưu ngân sách cả chuyến đi.

Quan trọng hơn là tình trạng còn chỗ hoặc còn suất (availability). Đối với một truy vấn cụ thể, một trải nghiệm đã hết chỗ gần như tương đương với một lựa chọn không tồn tại.

Cuối cùng là điều kiện tiếp cận (accessibility), gồm bãi đỗ xe, lối cho người khuyết tật, thang máy, mức phù hợp với trẻ nhỏ và người cao tuổi, khả năng đi bằng phương tiện công cộng. Nhóm thuộc tính này quyết định điểm đến phù hợp với ai.

Bốn nhóm trên đều là dữ liệu động, thay đổi theo ngày và theo giờ, trong khi mã xã và địa chỉ là dữ liệu tĩnh. Vấn đề của Việt Nam không phải là chưa ai biết những trường này cần thiết. Vấn đề là chúng chưa nằm trong hệ thống hành chính và quyền cập nhật chúng lại phân tán giữa nhiều chủ thể không có nghĩa vụ chia sẻ với nhau.

Từ đây xuất hiện một chuyển dịch trong tư duy dữ liệu. Đối với con người, vị trí trả lời câu hỏi "ở đâu". Đối với AI, khả năng hành động trả lời câu hỏi "có thể làm gì ngay bây giờ". Chỉ khi cả hai lớp thông tin cùng tồn tại, một điểm đến mới thật sự sẵn sàng cho nền kinh tế tác nhân.

### 8.3.4. Hai lớp dữ liệu, một thực thể

Một địa phương sẽ phải vận hành đồng thời hai lớp dữ liệu có mục tiêu khác nhau. Sai lầm phổ biến là dùng một lược đồ duy nhất để phục vụ cả hai. Cần nói rõ để tránh hiểu sai: hai lớp này không nhất thiết là hai hệ thống tách biệt. Chúng có thể nằm trong cùng một kiến trúc, miễn là cùng tham chiếu về một thực thể. Điều phải tách là chức năng, không nhất thiết là hạ tầng.

Lớp dữ liệu quản lý phục vụ quản trị nhà nước. Đơn vị trung tâm của nó là tổ chức, cá nhân, giấy phép, hồ sơ và đơn vị hành chính. Tiêu chí quan trọng nhất là tính pháp lý, tính chính xác và khả năng kiểm soát. Dữ liệu thay đổi chậm và phản ánh trạng thái quản lý.

Lớp dữ liệu điểm đến phục vụ người tiêu dùng và các tác nhân. Đơn vị trung tâm không còn là hồ sơ hành chính mà là thực thể trải nghiệm: điểm tham quan, khách sạn, nhà hàng, sự kiện, tuyến tham quan, phương tiện, vé, khung giờ và khả năng giao dịch. Tiêu chí quan trọng nhất không phải giấy phép, mà là khả năng được đọc, được suy luận và được đặt chỗ.

Lớp quản lý trả lời "địa phương đang quản lý những gì". Lớp điểm đến trả lời "máy có thể làm gì với địa phương này".

Trong nền kinh tế tác nhân, giá trị kinh tế sẽ ngày càng được tạo ra từ câu trả lời thứ hai. Vì vậy, địa phương không chỉ cần một chính quyền số, mà còn cần một điểm đến số, nơi mọi thực thể đều có đủ dữ liệu để AI có thể hiểu, lựa chọn và hành động.

### 8.3.5. Song sinh số phải phản ánh nhịp sống

Khái niệm song sinh số thường được hiểu là một bản sao số phản ánh trạng thái của thế giới thực. Tuy nhiên, nếu song sinh số của một địa phương chỉ bao gồm bản đồ hành chính, quy hoạch, dân số, doanh nghiệp, giấy phép và các hồ sơ quản lý, thì đó mới chỉ là bản sao của bộ máy hành chính, chưa phải bản sao của chính địa phương.

Một địa phương thực không chỉ tồn tại qua các quyết định hành chính mà còn tồn tại qua vô số hoạt động diễn ra mỗi ngày. Khách sạn nhận khách và trả phòng. Nhà hàng cập nhật thực đơn. Bảo tàng mở cửa rồi đóng cửa. Sự kiện bắt đầu rồi kết thúc. Tuyến xe thay đổi lịch trình. Giá phòng biến động theo mùa. Vé tham quan có lúc còn, có lúc hết. Chính những chuyển động liên tục này mới tạo nên "nhịp sống" của địa phương.

Nếu song sinh số không phản ánh được nhịp sống đó, máy chỉ nhìn thấy một địa phương tồn tại về mặt pháp lý, chứ không nhìn thấy một địa phương đang vận hành.

Đối với mục tiêu của chương này, một song sinh số địa phương có giá trị cần ít nhất ba lớp dữ liệu.

Lớp thứ nhất là dữ liệu hành chính, mô tả cấu trúc pháp lý, địa giới, tổ chức và tài sản công. Đây là nền móng bảo đảm tính chính xác và khả năng quản lý.

Lớp thứ hai là dữ liệu vận hành, phản ánh trạng thái động của toàn bộ hệ sinh thái điểm đến: giờ mở cửa, tình trạng hoạt động, tồn kho, giá, lịch sự kiện, lưu lượng khách, giao thông và các tín hiệu thời gian thực. Đây là lớp dữ liệu giúp AI hiểu địa phương đang diễn ra điều gì.

Lớp thứ ba là dữ liệu ngữ nghĩa và quan hệ, mô tả cách các thực thể kết nối với nhau trong một đồ thị tri thức. Một nhà hàng nằm gần bảo tàng nào, khách sạn phù hợp với loại hành trình nào, lễ hội gắn với mùa nào, tuyến xe nào kết nối các điểm tham quan, trải nghiệm nào thường được kết hợp với nhau. Chính mạng lưới quan hệ này cho phép AI suy luận thay vì chỉ tra cứu.

Đối với Việt Nam, đây là một bước chuyển trong tư duy chuyển đổi số, mục tiêu không còn là bản sao số của chính quyền mà là bản sao số của cả điểm đến. Khi ba lớp cùng hiện diện, một hệ thống không chỉ thấy nơi đó trên bản đồ mà còn có thể dự báo, đề xuất và giao dịch với nó.

Nếu bạn là giám đốc sở văn hóa, thể thao và du lịch. Chuẩn hóa lịch lễ hội toàn tỉnh thành dữ liệu có cấu trúc với thời gian bắt đầu, thời gian kết thúc, địa điểm, trạng thái và đơn vị chịu trách nhiệm, trước khi đầu tư thêm nội dung quảng bá. Yêu cầu mọi di tích và điểm tham quan thuộc quyền quản lý công bố giờ mở cửa và giá vé ở một địa chỉ máy đọc được, không chỉ trên bảng thông báo tại cửa. Coi mỗi lần lễ hội bị hoãn hoặc thay đổi là một sự kiện dữ liệu phải phát đi trong ngày.

## 8.4. NĂNG LỰC SỐ Ở CẤP XÃ VÀ CẤP TỈNH

Trong nền kinh tế tác nhân, khả năng hiện diện của một điểm đến không còn chỉ được quyết định bởi truyền thông, nó ngày càng phụ thuộc vào chất lượng dữ liệu được tạo và duy trì ngay tại nơi tài sản tồn tại. Một di tích có mở cửa đúng giờ hay không, một tuyến đường đã thay đổi, một lễ hội bị hoãn, một homestay còn phòng hay một bến tàu tạm ngừng hoạt động đều là những thông tin thường được chính quyền cấp xã và các đơn vị vận hành tại chỗ phát hiện và cập nhật sớm nhất.

Điều đó tạo ra một sự thay đổi quan trọng trong tư duy quản trị. Nếu trước đây chuyển đổi số ở cấp cơ sở chủ yếu phục vụ cải cách hành chính, thì nay nó còn trực tiếp quyết định khả năng địa phương được AI đọc, hiểu và đưa vào quá trình đề xuất. Mỗi xã, mỗi phường không chỉ quản lý địa giới hay hồ sơ dân cư, mà còn trở thành một đầu mối quan trọng trong việc phát hiện, xác nhận và duy trì dữ liệu nền cho hệ sinh thái số của địa phương.

Trong mô hình này, năng lực số không được đo bằng số lượng máy tính, phần mềm hay cổng dịch vụ công, mà bằng khả năng tạo ra dữ liệu chính xác, cập nhật, có cấu trúc và có thể được các hệ thống bên ngoài sử dụng. Một cán bộ văn hóa cập nhật lịch lễ hội đúng chuẩn dữ liệu có thể tạo ảnh hưởng lâu dài hơn một chiến dịch quảng bá ngắn hạn. Một cán bộ địa chính chuẩn hóa địa chỉ và tọa độ có thể giúp người dùng tìm đúng điểm đến mà trước đây máy không nhận diện được.

Vì vậy, trong nền kinh tế tác nhân, năng lực số của cấp xã không còn là năng lực vận hành nội bộ, mà trở thành năng lực tham gia vào hạ tầng dữ liệu quốc gia và toàn cầu. Đây chính là nơi quyết định liệu một tài sản địa phương có thực sự bước vào

chuỗi giá trị số hay tiếp tục tồn tại như một tài sản "vô hình" đối với máy. Chính vì thế, xây dựng năng lực số ở cấp cơ sở không còn là nhiệm vụ công nghệ, mà là điều kiện nền tảng để mỗi địa phương giữ được khả năng hiện diện, cạnh tranh và tạo giá trị trong nền kinh tế tác nhân.

Phong Nha cho thấy chiều ngược lại của cùng một câu chuyện. Một xã miền núi ở Quảng Trị sau sắp xếp, ngân sách xúc tiến của địa phương còn hạn chế, nhưng hệ thống hang động ở đây được nhận diện tương đối tốt, vì một lý do đơn giản: các đơn vị vận hành tour đã tự chuẩn hóa tên hang, độ dài tuyến, mức độ khó và lịch khởi hành từ nhiều năm trước, do họ cần bán cho khách nước ngoài. Khả đọc ở Phong Nha là sản phẩm phụ của việc bán hàng qua kênh quốc tế và đó cũng là con đường mà phần lớn địa phương nhỏ sẽ đi trước khi có một chương trình dữ liệu chính thức.

### 8.4.1. Vì sao cấp xã mới là tuyến đầu của AI

Khi nói đến AI, người ta thường nghĩ đến những trung tâm dữ liệu, nền tảng lớn, doanh nghiệp công nghệ hoặc cơ quan quản lý ở trung ương. Nhưng đối với dữ liệu trạng thái tại chỗ, tuyến đầu của hệ thống nằm ở cấp gần tài sản nhất. Từ ngày 01 tháng 7 năm 2025, đó là cấp xã. Sau sắp xếp đơn vị hành chính, quy mô trung bình của một đơn vị cấp xã tăng khoảng ba lần.

Thay đổi ấy tác động theo hai chiều ngược nhau và hai chiều đó cần được tách bạch. Chiều thuận lợi là về nguyên tắc, đường truyền dữ liệu có thể ngắn hơn vì giảm một cấp hành chính trung gian. Chiều bất lợi ít được nói tới hơn: chức năng điều phối vốn đặt ở cấp huyện không mất đi cùng với cấp huyện, nó chỉ đổi chỗ. Cần nói rõ mức độ của phát biểu này. Tác động ròng của hai chiều ấy lên độ tươi của dữ liệu là một giả thuyết chưa kiểm

chứng được chứ không phải một kết luận và nó được ghi lại đầy đủ như đã nêu dưới đây.

Vì sao cần chức năng đó, có thể thấy rõ qua hành trình của một người đi du lịch. Một chuyến đi thường kết nối nơi lưu trú, điểm tham quan, nhà hàng, bến xe, cơ sở y tế, sự kiện và dịch vụ hỗ trợ. Xã mới rộng hơn nên nhiều hành trình ngắn đã nằm gọn trong một xã, điều trước đây rất ít khi xảy ra. Nhưng phần lớn hành trình dài hơn một ngày vẫn vượt ranh giới xã và khi đó chỉ cấp tỉnh có đủ phạm vi để ghép các mảnh dữ liệu thành một hệ thống có nghĩa.

Điều cấp xã có mà không cấp nào thay thế được là quan sát. Xã là nơi phát hiện đầu tiên phần lớn những thay đổi có tác động tới khả năng hành động của máy: một tuyến đường bị cấm tạm thời, một lễ hội bị hoãn, một điểm tham quan ngừng đón khách, một cơ sở lưu trú đã dừng hoạt động nhưng chưa xóa tên, một công trình đang thi công. Những thay đổi này diễn ra nhanh hơn nhiều so với chu kỳ cập nhật của bất kỳ cơ sở dữ liệu cấp trên nào.

Đối với một tác nhân AI, dữ liệu cũ không chỉ là dữ liệu kém chất lượng. Nó có thể dẫn tới một quyết định sai. Một địa điểm được ghi là đang mở nhưng thực tế đã đóng cửa, một con đường được chỉ dẫn nhưng đang thi công, một lễ hội đã kết thúc nhưng vẫn xuất hiện trong lịch, một cơ sở lưu trú còn tên trên bản đồ nhưng không còn hoạt động. Mỗi sai lệch như vậy làm suy giảm niềm tin của hệ thống đối với toàn bộ địa phương.

Vì thế vai trò của xã không dừng ở việc phát hiện rồi báo cáo. Nếu xã chỉ báo cáo, độ trễ sẽ quay về mức của mô hình ba cấp trước đây, chỉ khác là ít hơn một tầng giấy tờ. Xã cần quyền ghi trực tiếp vào hệ thống đối với những trường dữ liệu biến động nhanh mà xã quan sát được, kèm trách nhiệm về chất lượng của chính những

trường đó. Phạm vi này không bao gồm giá, tồn kho và điều khoản thương mại của doanh nghiệp, cũng không bao gồm thông tin pháp lý chuyên ngành thuộc thẩm quyền của cơ quan khác.

Nhưng xã không thể làm phần còn lại. Phân giải thực thể, chuẩn hóa lược đồ, hợp nhất bản ghi trùng, công bố ra ngoài từ một nguồn duy nhất, điều phối dữ liệu giữa nhiều xã trong cùng một hành trình, tất cả đều đòi hỏi phạm vi và chuyên môn mà một xã không có. Đó là lý do mô hình hai cấp buộc phải sinh ra một lớp tích hợp ở cấp tỉnh, không phải như một phòng ban mới mà như một chức năng bắt buộc.

Nói cách khác, hai cấp không có nghĩa hai nơi làm cùng một việc. Xã quan sát và ghi nhận, tỉnh chuẩn hóa và công bố. AI có thể được xây dựng ở trung tâm, nhưng độ chính xác của nó được quyết định ở cơ sở. Trong cấu trúc quản trị mới, cấp xã là nơi dữ liệu thực địa được sinh ra và cấp tỉnh là nơi dữ liệu ấy trở thành thứ mà máy có thể dùng để suy luận.

**GIẢ THUYẾT RÒNG VỀ NĂNG LỰC SỐ CỦA MÔ HÌNH HAI CẤP**

*Việc kết thúc cấp huyện rút ngắn đường truyền dữ liệu từ thực địa lên tỉnh, đồng thời chuyển toàn bộ gánh nặng điều phối liên xã lên cấp tỉnh. Tác động ròng lên độ tươi của dữ liệu điểm đến phụ thuộc vào việc năng lực xã tăng nhanh hơn hay chậm hơn mức tải điều phối mới của tỉnh. Câu hỏi này chỉ có thể trả lời bằng dữ liệu quan sát trong giai đoạn 2026 đến 2028 và tác giả chưa thể có cơ sở để dự đoán chiều của kết quả.*

### 8.4.2. Người hiểu dữ liệu nhất lại ít được quyền cập nhật

Một nghịch lý phổ biến trong quản trị dữ liệu địa phương là người hiểu thông tin nhất thường không phải là người có quyền thay đổi thông tin đó trên hệ thống.

Người quản lý di tích biết chính xác giờ mở cửa đã thay đổi. Cán bộ văn hóa xã biết lễ hội năm nay tổ chức vào ngày nào. Chủ homestay biết cơ sở còn phòng hay đã ngừng hoạt động. Người phụ trách bến tàu biết chuyến nào bị hủy. Cán bộ địa chính biết tên đường, số nhà và địa giới đã được điều chỉnh. Nhưng trong nhiều trường hợp, những người này không thể trực tiếp cập nhật dữ liệu chính thức. Thông tin phải đi qua nhiều tầng báo cáo, biểu mẫu, phê duyệt và tổng hợp trước khi xuất hiện trên một cơ sở dữ liệu nào đó. Trong khoảng thời gian ấy, thực tế đã thay đổi nhưng dữ liệu vẫn đứng yên.

Cơ chế này có thể phù hợp với dữ liệu hành chính cần kiểm soát chặt chẽ, nhưng lại không phù hợp với dữ liệu điểm đến có tốc độ thay đổi cao. Giờ mở cửa, giá vé, lịch sự kiện, khả năng tiếp cận, tình trạng giao thông hay tồn kho dịch vụ không thể chờ một quy trình cập nhật kéo dài nhiều tuần. Đối với người dùng và tác nhân AI, giá trị của những dữ liệu này giảm rất nhanh theo thời gian.

Vấn đề không đơn thuần là thiếu nhân lực công nghệ. Cốt lõi nằm ở cách phân bổ quyền dữ liệu. Hệ thống thường được thiết kế theo tư duy tập trung: cấp trên sở hữu cơ sở dữ liệu, cấp dưới cung cấp số liệu. Nhưng dữ liệu điểm đến lại được tạo ra theo mô hình phân tán. Mỗi cơ sở, mỗi đơn vị, mỗi xã chỉ nắm một phần thông tin và chính họ là nơi phát hiện sai lệch sớm nhất.

Khi quyền cập nhật bị tách khỏi nơi phát sinh dữ liệu, ba hệ quả xuất hiện. Thứ nhất, dữ liệu luôn có độ trễ. Thứ hai, trách nhiệm trở nên mơ hồ vì người phát hiện sai không có quyền sửa, còn người có quyền sửa lại không đủ thông tin để xác minh. Thứ ba, các đơn vị cơ sở dần mất động lực báo cáo vì không nhìn thấy dữ liệu của mình được sử dụng như thế nào.

Trong khi đó, các nền tảng thương mại lại vận hành theo logic ngược lại. Chủ khách sạn có thể cập nhật số phòng, nhà hàng có thể sửa giờ mở cửa, ban tổ chức có thể đổi lịch sự kiện, người dùng có thể báo địa điểm đã đóng. Dữ liệu được cập nhật gần nguồn phát sinh hơn, vì vậy thường nhanh hơn dữ liệu chính thức, dù không phải lúc nào cũng chính xác hơn.

Bài toán của địa phương không phải lựa chọn giữa kiểm soát tập trung và cập nhật phân tán. Điều cần thiết là một kiến trúc quyền phù hợp với từng loại dữ liệu. Dữ liệu pháp lý như địa giới, tên chính thức hay giấy phép có thể cần phê duyệt chặt chẽ. Nhưng dữ liệu vận hành như giờ mở cửa, lịch dịch vụ, tình trạng hoạt động và khả năng tiếp cận phải được phép cập nhật ngay tại nguồn, kèm theo nhật ký thay đổi, xác thực danh tính và cơ chế kiểm tra sau. Nói cách khác, địa phương cần chuyển từ mô hình "một nơi được quyền sửa mọi dữ liệu" sang mô hình "mỗi chủ thể được quyền cập nhật phần dữ liệu mà mình chịu trách nhiệm". Chính quyền không nhất thiết phải nhập thay doanh nghiệp, điểm tham quan hay đơn vị tổ chức sự kiện. Vai trò của chính quyền là xác lập chuẩn dữ liệu, phân quyền, giám sát và xử lý tranh chấp.

Khi người hiểu dữ liệu nhất được quyền cập nhật dữ liệu, độ trễ giảm, trách nhiệm rõ hơn và thông tin trở nên sống động hơn. Ngược lại, nếu quyền cập nhật vẫn nằm xa nguồn dữ liệu,

địa phương sẽ tiếp tục sở hữu những cơ sở dữ liệu rất lớn nhưng luôn chậm hơn thực tế.

Trong nền kinh tế tác nhân, quyền cập nhật không chỉ là một quyền hành chính. Đó là quyền quyết định phiên bản nào của địa phương sẽ được máy nhìn thấy. Khoảng cách quyền cập nhật. Độ tươi của dữ liệu vận hành giảm khi khoảng cách giữa người biết sự thật và người có quyền ghi sự thật tăng. Khoảng cách này được đo bằng số bước phê duyệt, không bằng khoảng cách địa lý.

Ai biết động Phong Nha đóng cửa hôm nay. Vào mùa mưa, mực nước sông Son có thể dâng tới mức không cho phép thuyền vào động. Quyết định tạm ngừng đón khách được đưa ra tại chỗ, đôi khi chỉ trước vài giờ. Người biết trước nhất là đơn vị vận hành bến thuyền. Người có thẩm quyền công bố lại thường ở một cấp khác. Trong khoảng thời gian giữa hai thời điểm ấy, các hệ thống bên ngoài vẫn hiển thị điểm đến đang hoạt động bình thường và một tác nhân được hỏi lúc đó sẽ trả lời rằng có thể đi. Đây là dạng sai lệch mà không một mô hình ngôn ngữ nào sửa được, vì thông tin đúng chưa từng tồn tại dưới dạng dữ liệu. Nó cũng là dạng sai lệch gây tổn hại lớn nhất tới niềm tin, vì hậu quả rơi trực tiếp vào một chuyến đi cụ thể của một người cụ thể. Nguồn và mức xác thực: Cơ chế tạm ngừng đón khách theo mực nước là đặc điểm đã được ghi nhận của tuyến tham quan này. Diễn biến cụ thể trong hộp là tình huống dựng lại để minh họa quy luật khoảng cách quyền cập nhật, không phải một sự việc được ghi chép.

Cấp xã không đủ nhân lực và chuyên môn để làm dữ liệu. Nên giao toàn bộ việc này cho cấp tỉnh. Trả lời: Phần đúng: năng lực chuyên môn về dữ liệu ở cấp xã hiện thấp và tập trung về tỉnh giúp chuẩn hóa nhanh hơn, kiểm soát chất lượng tốt hơn, chi phí trên mỗi bản ghi thấp hơn. Nhưng có một thứ không thể tập

trung được, đó là quan sát. Tỉnh có thể chuẩn hóa một bản ghi, tỉnh không thể biết một bến thuyền vừa ngừng hoạt động sáng nay. Tập trung toàn bộ về tỉnh sẽ cải thiện tính nhất quán và làm giảm độ tươi, mà độ tươi lại là thuộc tính quyết định trong các quyết định vận hành. Cách tổ chức phù hợp hơn không phải chọn một trong hai cấp mà phân tách theo loại công việc: xã quan sát và ghi nhận, tỉnh chuẩn hóa và công bố. Mô hình này chỉ chạy được nếu xã có quyền ghi trực tiếp, điều mà nhiều nơi hiện chưa có.

### 8.4.3. Một cán bộ văn hóa có thể làm thay đổi khả đọc của cả địa phương

Cần nói ngay để tránh một cách hiểu sai: cán bộ văn hóa không phải người nhập toàn bộ dữ liệu của địa phương. Họ là ví dụ rõ nhất cho một nguyên tắc, rằng người gần tài sản nhất phải có vai trò trực tiếp trong việc xác nhận trạng thái của tài sản ấy. Trong tư duy quản lý truyền thống, một cán bộ văn hóa cấp xã được xem là người tổ chức lễ hội, quản lý di tích và triển khai hoạt động văn hóa cộng đồng. Vai trò đó nay có thêm một chiều mới: những gì họ ghi lại quyết định phần lớn những gì máy nhìn thấy về nhóm tài sản họ phụ trách. Điều này không xuất phát từ quyền lực hành chính, mà từ quyền tạo ra và duy trì dữ liệu.

Một lễ hội chỉ thực sự tồn tại với AI khi có thời gian bắt đầu và kết thúc rõ ràng, địa điểm chính xác, mô tả đầy đủ, thông tin vé, khả năng tiếp cận, lịch biểu và trạng thái cập nhật. Một di tích chỉ trở thành điểm đến có thể được đề xuất khi giờ mở cửa, giá vé, tọa độ, hình ảnh, đơn vị quản lý và các dịch vụ liên quan luôn phản ánh đúng thực tế. Tất cả những thông tin này đều bắt nguồn từ những người làm công tác văn hóa tại địa phương.

Nếu dữ liệu không được cập nhật, AI không biết lễ hội đã bắt đầu hay kết thúc. Nếu giờ mở cửa sai, tác nhân có thể loại điểm đến khỏi danh sách đề xuất. Nếu địa chỉ không chính xác, hệ thống điều hướng sẽ mất niềm tin. Nếu thông tin về khả năng tiếp cận của người khuyết tật, bãi đỗ xe hay hướng dẫn tham quan không tồn tại, AI sẽ ưu tiên những địa điểm khác có dữ liệu đầy đủ hơn.

Phần lớn các quyết định này không đòi hỏi trí tuệ nhân tạo phức tạp. Chúng chỉ đòi hỏi dữ liệu đúng. Trong nhiều trường hợp, một thay đổi nhỏ như cập nhật lịch lễ hội đúng ngày, bổ sung thông tin đặt vé trực tuyến hoặc xác nhận một điểm tham quan vẫn đang hoạt động có thể tạo ra tác động lớn hơn nhiều so với một chiến dịch quảng bá tốn kém. Bởi AI không đọc khẩu hiệu, AI đọc dữ liệu. Điều đó khiến vai trò của cán bộ văn hóa thay đổi căn bản. Họ không chỉ quản lý di sản, mà còn quản lý "phiên bản số" của di sản. Họ không chỉ tổ chức sự kiện, mà còn đảm bảo sự kiện có thể được các hệ thống số nhận biết, hiểu và đề xuất. Nói cách khác, họ đang vận hành một phần của song sinh số địa phương.

Mỗi cán bộ ở tuyến cơ sở đều có thể trở thành người duy trì niềm tin dữ liệu của địa phương. Một bản cập nhật chính xác hôm nay có thể được tái sử dụng trong rất nhiều lần truy vấn và đề xuất về sau. Vì vậy, công việc nhập liệu không còn là thao tác hành chính đơn thuần mà đã trở thành một mắt xích trong chuỗi tạo ra giá trị kinh tế.

Địa phương muốn AI hiểu mình không chỉ cần chuyên gia công nghệ ở cấp tỉnh hay trung ương. Địa phương cần những người quản lý tài sản ở cấp cơ sở nhận thức rằng mỗi thông tin họ cập nhật đều là một tín hiệu mà máy sẽ sử dụng để đưa ra quyết định. Người biết sự thật và người được ghi sự thật càng xa nhau, dữ liệu càng cũ.

### 8.4.4. Năng lực số không phải mua phần mềm

Sai lầm phổ biến của nhiều chương trình chuyển đổi số là đồng nhất năng lực số với việc sở hữu nhiều phần mềm hơn. Khi hệ thống vận hành chưa thay đổi, phần mềm mới chỉ số hóa quy trình cũ, chứ không tạo ra năng lực mới.

Câu hỏi quan trọng không phải là địa phương đang sử dụng phần mềm nào, mà là ai chịu trách nhiệm cập nhật dữ liệu, dữ liệu được cập nhật theo quy trình nào và mất bao lâu để thay đổi ngoài thực địa xuất hiện trên hệ thống.

Nếu một di tích đổi giờ mở cửa vào sáng nay, chiều nay AI đã biết chưa? Nếu một tuyến đường bị cấm tạm thời, bao lâu sau dữ liệu được đồng bộ? Nếu một lễ hội bị hoãn vì thời tiết, thông tin đó được cập nhật trong vài phút hay vài tuần? Đây mới là những chỉ số phản ánh năng lực số thực sự.

Vì vậy, năng lực số của địa phương được xây trên ba trụ cột. Thứ nhất là quy trình (process). Mỗi loại dữ liệu phải có một quy trình cập nhật rõ ràng, từ phát hiện thay đổi, xác minh, chuẩn hóa, công bố đến đồng bộ với các hệ thống khác. Quy trình càng ngắn, dữ liệu càng phản ánh đúng thực tế. Thứ hai là đầu mối chịu trách nhiệm (data stewardship). Mỗi trường dữ liệu phải có một đầu mối, không nhất thiết là chủ sở hữu theo nghĩa pháp lý. Ai chịu trách nhiệm về giờ mở cửa. Ai xác nhận giá vé. Ai cập nhật lịch sự kiện. Khi câu hỏi này có tên người trả lời, dữ liệu mới có người chịu trách nhiệm về chất lượng của nó. Thứ ba là độ tươi của dữ liệu (freshness). Trong nền kinh tế tác nhân, dữ liệu không chỉ cần đúng mà còn phải đúng vào thời điểm được hỏi. Một thông tin chính xác nhưng đã cũ nhiều tuần có thể gây hại hơn một dữ liệu chưa tồn tại, vì nó khiến AI đưa ra quyết định dựa trên một thực tế không còn tồn tại.

Độ tươi vì thế trở thành một năng lực vận hành chứ không phải một thuộc tính kỹ thuật. Nó phản ánh tốc độ mà địa phương có thể chuyển đổi biến động ngoài đời thành dữ liệu số. Địa phương nào làm được điều này sẽ tạo ra song sinh số gần với thực tế hơn, từ đó tăng khả năng được AI tin cậy.

Nhìn dưới góc độ này, đầu tư quan trọng nhất không phải là mua thêm một hệ thống phần mềm mới, mà là thiết kế lại cách dữ liệu được tạo ra, cập nhật và duy trì. Phần mềm chỉ là công cụ. Quy trình tạo ra dòng chảy dữ liệu. Trách nhiệm tạo ra chất lượng dữ liệu. Còn độ tươi tạo ra niềm tin của máy.

Ba loại quyền vì thế cần được phân biệt và không nên gộp lại. Quyền cập nhật đi theo nguồn phát sinh sự thật, nên nó thuộc về nơi quan sát được thay đổi, có thể là xã đối với trạng thái tài sản công hoặc doanh nghiệp đối với giá và tồn kho của chính mình. Quyền xác nhận đi theo thẩm quyền, nên nó thuộc về cơ quan có thẩm quyền xác nhận tình trạng pháp lý và danh tính chính thức. Quyền công bố đi theo kiến trúc quản trị, nên nó thuộc về cấp giữ nguồn duy nhất phát ra bên ngoài. Trao cả ba cho một cấp sẽ tạo ra nút thắt, còn để cả ba trôi nổi sẽ tạo ra ba phiên bản của cùng một sự thật.

Địa phương mạnh về số không phải là địa phương có nhiều ứng dụng nhất, mà là địa phương có khả năng biến mọi thay đổi trong thế giới thực thành dữ liệu đáng tin cậy nhanh nhất. Đây mới là năng lực cạnh tranh cốt lõi của quản trị địa phương trong kỷ nguyên AI.

### 8.4.5. Một bộ chỉ số tối thiểu cho cấp xã

Năng lực số chỉ có thể được cải thiện khi nó được đo lường. Tuy nhiên, nhiều địa phương vẫn đánh giá chuyển đổi số bằng số lượng phần mềm đã triển khai, số tài khoản được cấp, số hồ sơ được số hóa hoặc số lượt truy cập cổng thông tin. Những chỉ số này phản ánh mức độ sử dụng công cụ, nhưng chưa cho biết dữ liệu của địa phương có thực sự hữu ích đối với người dân, doanh nghiệp và các tác nhân AI hay không.

Đối với cấp xã, một bộ chỉ số hiệu quả không cần quá phức tạp. Điều quan trọng là đo được năm thuộc tính cơ bản của dữ liệu: mức độ mới, mức độ bao phủ, tính nhất quán, khả năng sử dụng và tốc độ phản ứng.

**Độ tươi: Dữ liệu mới đến đâu?**

Độ tươi đo khoảng cách thời gian giữa một thay đổi ngoài thực tế và thời điểm thay đổi đó xuất hiện trên hệ thống.

Một điểm tham quan đổi giờ mở cửa vào thứ Hai nhưng đến thứ Sáu dữ liệu mới được cập nhật thì độ trễ là bốn ngày. Một lễ hội bị hoãn nhưng thông tin trực tuyến vẫn giữ lịch cũ cho đến sau ngày tổ chức thì dữ liệu gần như không còn giá trị vận hành.

Mỗi loại dữ liệu cần có một ngưỡng độ tươi khác nhau. Địa giới hành chính có thể được rà soát theo quý hoặc theo năm. Thông tin cơ sở dịch vụ có thể cần xác nhận hằng tháng. Lịch sự kiện, tình trạng giao thông, giờ mở cửa và lượng vé còn lại phải được cập nhật theo ngày, theo giờ, thậm chí theo thời gian thực.

Cấp xã vì vậy không nên chỉ hỏi dữ liệu có tồn tại hay không, mà phải hỏi, dữ liệu này được xác nhận lần cuối khi nào?

**Độ phủ: Địa phương đã được mô tả đầy đủ chưa?**

Độ phủ đo tỷ lệ tài sản, dịch vụ và sự kiện đã được đưa vào cơ sở dữ liệu so với những gì thực sự tồn tại trên địa bàn.

Một xã có thể đã số hóa toàn bộ di tích được xếp hạng nhưng lại bỏ sót hàng trăm nhà hàng, làng nghề, điểm ngắm cảnh, cơ sở lưu trú nhỏ, không gian văn hóa cộng đồng và dịch vụ vận chuyển. Khi đó, cơ sở dữ liệu chính thức chỉ mô tả được phần hành chính của địa phương, chứ chưa phản ánh đầy đủ nền kinh tế trải nghiệm.

Độ phủ cũng phải được đo ở cấp độ thuộc tính. Một điểm đến có tên và địa chỉ chưa có nghĩa là đã được mô tả đầy đủ. Nó còn cần tọa độ, giờ mở cửa, giá vé, phương thức liên hệ, khả năng tiếp cận, hình ảnh, trạng thái hoạt động và liên kết đặt dịch vụ.

Một xã có độ phủ cao không chỉ có nhiều bản ghi, mà có đủ các thực thể và đủ thông tin để máy có thể hiểu, so sánh và kết nối chúng.

**Tính nhất quán: Các hệ thống có đang nói cùng một sự thật?**

Tính nhất quán đo mức độ thống nhất của dữ liệu giữa các phòng ban, cơ sở dữ liệu và nền tảng công khai.

Một địa điểm có thể mang tên khác nhau trên cổng thông tin xã, bản đồ số, hồ sơ ngành văn hóa và nền tảng du lịch. Giá vé trên trang web có thể khác với bảng giá tại điểm đến. Một lễ hội có thể được công bố vào ba ngày khác nhau trên ba hệ thống. Những sai lệch này khiến người dùng bối rối và làm giảm niềm tin của AI.

Tính nhất quán không đòi hỏi các hệ thống phải lưu trữ dữ liệu theo cùng một cách. Nhưng các thông tin cốt lõi như tên chuẩn, địa chỉ, tọa độ, trạng thái hoạt động, đơn vị quản lý và thời gian tổ chức phải cùng tham chiếu đến một thực thể thống nhất.

Cấp xã cần theo dõi tỷ lệ bản ghi có xung đột và thời gian cần thiết để giải quyết xung đột. Càng nhiều phiên bản khác nhau của cùng một sự thật, khả năng đọc của địa phương càng thấp.

**Khả năng dùng được: dữ liệu có thực sự sử dụng được không?**

Khả năng dùng được không chỉ có nghĩa là hệ thống đang hoạt động. Nó còn đo khả năng dữ liệu được tìm thấy, truy cập và sử dụng bởi các chủ thể bên ngoài.

Một cơ sở dữ liệu có thể rất đầy đủ nhưng bị khóa trong một phần mềm nội bộ, chỉ xuất được dưới dạng PDF hoặc không có cơ chế chia sẻ. Trong trường hợp đó, dữ liệu tồn tại về mặt hành chính nhưng gần như không tồn tại đối với hệ sinh thái số.

Dữ liệu có khả năng dùng được cao cần được công bố ở định dạng có cấu trúc, có địa chỉ truy cập ổn định, có cơ chế tìm kiếm và có thể tích hợp với các hệ thống khác. Những thông tin cần thiết cho người dùng và doanh nghiệp phải được mở ở mức phù hợp, trong khi dữ liệu cá nhân và dữ liệu nhạy cảm vẫn được bảo vệ.

Đối với nền kinh tế tác nhân, một thông tin không thể truy cập bằng máy gần như tương đương với một thông tin chưa được số hóa.

### Thời gian phản hồi: Mất bao lâu để sửa một dữ liệu sai?

Thời gian phản hồi đo tốc độ phản ứng của bộ máy khi phát hiện dữ liệu sai, thiếu hoặc lỗi thời.

Khi người dân báo một địa điểm đã đóng cửa, bao lâu thông tin được xác minh? Khi doanh nghiệp yêu cầu sửa địa chỉ, bao lâu thay đổi xuất hiện trên hệ thống? Khi một lỗi dữ liệu ảnh hưởng đến nhiều nền tảng, ai tiếp nhận và xử lý nó?

Chỉ số này đặc biệt quan trọng vì không có cơ sở dữ liệu nào hoàn toàn không có lỗi. Khác biệt giữa một hệ thống tốt và một hệ thống yếu không nằm ở việc có phát sinh sai sót hay không, mà ở khả năng phát hiện và sửa sai nhanh đến mức nào.

Cấp xã cần có một đầu mối tiếp nhận phản hồi, phân loại mức độ ưu tiên và quy định thời hạn xử lý cho từng nhóm dữ liệu. Sai giờ mở cửa có thể cần sửa trong ngày. Sai tọa độ của một điểm cấp cứu hoặc tuyến giao thông có thể phải xử lý ngay. Sai mô tả lịch sử có thể được xác minh trong thời gian dài hơn.

Năm chỉ số này tạo thành một bảng điều khiển tối thiểu:

- *Độ tươi cho biết dữ liệu có còn mới.*
- *Độ phủ cho biết địa phương đã được mô tả đầy đủ đến đâu.*
- *Tính nhất quán cho biết các hệ thống có thống nhất với nhau không.*
- *Khả năng dùng được cho biết dữ liệu có thể được sử dụng hay chỉ đang được lưu trữ.*
- *Thời gian phản hồi cho biết bộ máy có khả năng sửa sai nhanh đến mức nào.*

Bộ chỉ số này không nhằm tạo thêm một tầng báo cáo hành chính. Mục tiêu của nó là giúp lãnh đạo xã và lãnh đạo tỉnh nhìn thấy những điểm mà địa phương đang dần trở nên vô hình đối với máy.

Bộ chỉ số năm thành tố sẽ nhanh chóng trở thành một tầng báo cáo hành chính mới, tốn nguồn lực mà không cải thiện dữ liệu. Nguy cơ này là thật. Nó có tiền lệ. Nhiều bộ chỉ số chuyển đổi số đã đi đúng con đường đó: được thiết kế để đo năng lực, kết thúc bằng việc đo mức độ hoàn thành báo cáo. Khác biệt nằm ở chỗ số liệu được sinh ra như thế nào. Nếu năm chỉ số được kê khai định kỳ bằng biểu mẫu, chúng sẽ trở thành báo cáo. Nếu chúng được đo tự động từ chính hệ thống dữ liệu, tức đếm dấu thời gian của lần cập nhật cuối, đếm số bản ghi thiếu trường bắt buộc, đếm số xung đột giữa hai nguồn, thì chi phí đo gần bằng không và số liệu không thể kê khai theo ý muốn. Bộ chỉ số này vì thế chỉ có nghĩa khi không dừng ở việc ban hành mà chỉ đo những gì hệ thống tự sinh được. Không thu thập một trường dữ liệu nào mà chưa xác định đơn vị chịu trách nhiệm cập nhật trường đó và thời hạn cập nhật. Không đo năng lực số bằng số phần mềm đã triển khai, số tài khoản đã cấp hay số hồ sơ đã số hóa. Không coi một hạng mục khả đọc là hoàn thành khi chưa có cơ chế tiếp nhận báo lỗi và thời hạn xử lý cho từng nhóm dữ liệu.

Một xã không thể chỉ báo cáo rằng mình có cơ sở dữ liệu. Xã phải chứng minh rằng dữ liệu ấy đủ mới, đủ rộng, đủ nhất quán, luôn sẵn sàng và có thể được sửa chữa nhanh chóng. Đó mới là tiêu chuẩn tối thiểu của một địa phương có năng lực số.

Nếu bạn là chủ tịch Ủy ban nhân dân xã. Lập danh sách toàn bộ tài sản trải nghiệm trên địa bàn, gồm cả những cơ sở nhỏ

không thuộc diện quản lý chuyên ngành và xác định ai đang biết trạng thái thực của từng cơ sở. Đề nghị quyền ghi trực tiếp cho những trường dữ liệu biến động nhanh, giờ mở cửa, trạng thái hoạt động, lịch sự kiện, tình trạng đường tiếp cận. Đặt thời hạn cập nhật theo loại dữ liệu chứ không theo kỳ báo cáo và công khai thời hạn đó để người dân và doanh nghiệp có thể đối chiếu.

## 8.5. QUAN HỆ GIỮA CƠ QUAN QUẢN LÝ VÀ DOANH NGHIỆP

Khả đọc của địa phương không thể do chính quyền hoặc doanh nghiệp tạo ra một mình. Dữ liệu của điểm đến phân tán trong toàn bộ hệ sinh thái.

Trong nền kinh tế tác nhân, không một tổ chức nào sở hữu đầy đủ dữ liệu để giúp máy hiểu và giao dịch với một địa phương. Chính quyền nắm dữ liệu quy hoạch, hạ tầng và quản lý công. Doanh nghiệp lại nắm dữ liệu về sản phẩm, giá bán, tồn kho, lịch hoạt động và trải nghiệm khách hàng. Các nền tảng số lưu giữ đánh giá, hành vi tìm kiếm và tín hiệu tương tác. Khi những mảnh dữ liệu này tồn tại rời rạc, AI chỉ nhìn thấy một bức tranh không hoàn chỉnh, dẫn đến việc hiểu sai, đề xuất sai hoặc không thể thực hiện giao dịch.

Năng lực số của địa phương không còn là năng lực của riêng cơ quan quản lý hay từng doanh nghiệp, mà là năng lực của toàn bộ hệ sinh thái. Giá trị của một điểm đến được quyết định bởi khả năng kết nối các nguồn dữ liệu thành một mô hình thống nhất, nơi mọi thực thể, từ điểm tham quan, khách sạn, nhà hàng đến sự kiện, phương tiện và dịch vụ, đều được mô tả bằng cùng một ngôn ngữ mà máy có thể đọc và liên kết.

Điều này cũng làm thay đổi vai trò của chính quyền. Thay vì cố gắng tự xây dựng và cập nhật toàn bộ dữ liệu, cơ quan quản lý cần trở thành kiến trúc sư của hệ sinh thái dữ liệu: thiết lập tiêu chuẩn, xác thực thực thể, điều phối chia sẻ thông tin và tạo cơ chế để doanh nghiệp tham gia cập nhật liên tục. Ngược lại, doanh nghiệp không còn chỉ là đơn vị cung cấp dịch vụ mà trở thành nhà cung cấp dữ liệu vận hành theo thời gian thực cho toàn bộ điểm đến.

Một địa phương chỉ thật sự khả đọc với máy khi chính quyền, doanh nghiệp và các nền tảng cùng tạo ra một "nguồn sự thật chung" (shared source of truth). Khi đó, AI không chỉ biết địa phương có gì, mà còn biết điều gì đang mở cửa, còn chỗ, có thể đặt, có thể thanh toán và có thể hoàn thành giao dịch ngay lập tức. Khả đọc vì thế không phải là thành quả của một cơ quan riêng lẻ, mà là kết quả của năng lực hợp tác giữa toàn bộ hệ sinh thái địa phương.

Thiếu người tích hợp. Khi không chủ thể nào chịu trách nhiệm hợp nhất, mỗi hệ thống dữ liệu mới được xây thêm mà không dùng chung định danh và chuẩn liên thông có thể làm tăng số phiên bản của cùng một sự thật. Trong điều kiện đó, đầu tư dữ liệu có thể làm giảm khả đọc thay vì làm tăng.

### 8.5.1. Chính quyền giữ danh tính

Tài sản quan trọng nhất mà chính quyền nắm giữ không phải là toàn bộ dữ liệu của điểm đến, mà là danh tính số (digital identity) của các thực thể. Chính quyền giữ thẩm quyền xác nhận danh tính pháp lý, địa giới, tên chính thức và tình trạng công nhận của các thực thể thuộc phạm vi quản lý. Đây là lớp dữ liệu nền tảng mà không doanh nghiệp riêng lẻ nào thay thế được. Cần nói thêm một vế để tránh hiểu quá rộng: cộng đồng và thị trường vẫn tạo ra những thực thể trải nghiệm mà chính quyền không trực tiếp khai sinh, từ khu phố ẩm thực hình thành tự phát tới những địa điểm chỉ có tên gọi dân gian và các thực thể ấy vẫn cần được ghi nhận trong lớp dữ liệu điểm đến.

Nếu danh tính không thống nhất, toàn bộ hệ sinh thái sẽ bị phân mảnh. Một điểm tham quan có thể xuất hiện dưới nhiều

tên khác nhau trên Google Maps, OTA, mạng xã hội và hệ thống AI. Một khách sạn có thể thay đổi địa chỉ sau khi sáp nhập đơn vị hành chính nhưng các nền tảng cập nhật không đồng thời. Một bảo tàng đổi tên nhưng các liên kết cũ vẫn tồn tại. Đối với con người, đây chỉ là những sai khác nhỏ. Đối với máy, đó có thể là nhiều thực thể hoàn toàn khác nhau.

Vì vậy, nhiệm vụ của chính quyền không phải là quản lý giá phòng khách sạn hay số lượng vé còn bán, mà là duy trì một hệ thống định danh thống nhất cho toàn bộ địa phương. Mỗi điểm đến cần có một thực thể chuẩn, một bộ định danh ổn định, các tên thay thế, lịch sử thay đổi, vị trí chính xác và quan hệ với những thực thể liên quan. Khi lớp danh tính này ổn định, mọi doanh nghiệp và nền tảng đều có thể tham chiếu tới cùng một đối tượng, giúp AI hợp nhất dữ liệu thay vì chia nhỏ niềm tin.

Nói cách khác, chính quyền không cần sở hữu mọi dữ liệu, nhưng cần chịu trách nhiệm quản trị và bảo đảm tính có thẩm quyền của chuẩn danh tính địa phương. Đây là hạ tầng niềm tin để cả hệ sinh thái số vận hành trên cùng một bản đồ thực thể, bất kể dữ liệu được cập nhật từ bao nhiêu nguồn.

### 8.5.2. Doanh nghiệp giữ tồn kho

Nếu chính quyền chịu trách nhiệm về danh tính của điểm đến, thì doanh nghiệp nắm giữ phần lớn dữ liệu vận hành và giao dịch của các dịch vụ thương mại tại địa phương. Những gì một tác nhân cần để hoàn thành giao dịch, gồm số phòng còn trống, số ghế còn bán, giờ phục vụ, giá hiện tại, lịch biểu diễn, suất ăn và vé tham quan, đều nằm trong hệ thống của đơn vị vận hành chứ không nằm trong cơ sở dữ liệu quản lý nhà nước. Cần nói

thêm cho đủ: dữ liệu vận hành của tài sản công như bảo tàng, giao thông công cộng, đường sá hay lễ hội do cơ quan nhà nước tổ chức vẫn thuộc các đơn vị công.

Đây là loại dữ liệu có tốc độ thay đổi rất cao. Một khách sạn có thể hết phòng chỉ sau vài phút. Một nhà hàng có thể kín chỗ vào cuối tuần nhưng còn nhiều bàn vào chiều thứ Hai. Một khu vui chơi có thể tạm đóng vì thời tiết hoặc bảo trì. Những thay đổi này diễn ra liên tục và chỉ doanh nghiệp mới có khả năng cập nhật theo thời gian thực.

Điều đó có nghĩa là khả đọc của địa phương không chỉ phụ thuộc vào việc có dữ liệu hay không, mà còn phụ thuộc vào việc tồn kho có được chia sẻ và đồng bộ hay không. Nếu doanh nghiệp chỉ cập nhật thông tin trên trang web riêng, trong khi OTA, Google Maps và các hệ thống AI vẫn sử dụng dữ liệu cũ, máy sẽ không thể biết đâu là trạng thái hiện tại của dịch vụ. Kết quả là AI có thể đề xuất một nơi đã hết chỗ, báo giá không còn hiệu lực hoặc từ chối giới thiệu vì không xác minh được khả năng giao dịch.

Vai trò mới của doanh nghiệp vì thế không chỉ là bán dịch vụ mà còn là cung cấp dữ liệu vận hành liên tục cho hệ sinh thái. Mỗi lần cập nhật giá, thay đổi lịch mở cửa, bổ sung sự kiện hay điều chỉnh tồn kho đều là một tín hiệu giúp AI hiểu đúng trạng thái của điểm đến. Trong nền kinh tế tác nhân, dữ liệu vận hành trở thành một phần của chính dịch vụ.

Nếu chính quyền cung cấp danh tính đáng tin cậy, thì doanh nghiệp cung cấp khả năng giao dịch đáng tin cậy. Chỉ khi hai lớp dữ liệu này kết hợp với nhau, một địa phương mới thực sự trở thành một thực thể mà AI có thể đọc, suy luận và hành động.

### 8.5.3. Nền tảng giữ hành vi

Nếu chính quyền nắm danh tính và doanh nghiệp nắm dữ liệu vận hành, thì các nền tảng số nắm giữ dữ liệu hành vi, một trong những lớp dữ liệu quan trọng ảnh hưởng đến cách hệ thống xếp hạng và khuyến nghị. Mỗi lượt tìm kiếm, mỗi lần nhấp chuột, mỗi đánh giá, mỗi lượt lưu địa điểm và mỗi giao dịch đều để lại dấu vết về cách con người thực sự sử dụng một điểm đến.

Đây là loại dữ liệu mà cả chính quyền lẫn doanh nghiệp riêng lẻ đều rất khó quan sát đầy đủ. Một khách sạn biết khách đã đặt phòng của mình, nhưng không biết trước đó khách đã so sánh với bao nhiêu khách sạn khác. Một bảo tàng biết lượng khách tham quan, nhưng không biết vì sao hàng nghìn người khác quyết định không đến. Ngược lại, một số nền tảng lớn có thể quan sát nhiều đoạn của hành trình hơn từng doanh nghiệp hoặc cơ quan quản lý riêng lẻ. Không nền tảng nào mặc nhiên thấy toàn bộ hành trình, vì hành trình thật thường đi qua nhiều nền tảng.

Chính vì vậy, các nền tảng không chỉ là nơi phân phối khách mà còn là nơi tích lũy tri thức về nhu cầu của thị trường. Qua thời gian, chúng hiểu mùa nào khách thích đi đâu, nhóm khách nào quan tâm đến loại trải nghiệm nào, mức giá nào tạo ra tỷ lệ đặt chỗ cao nhất và những yếu tố nào làm thay đổi quyết định lựa chọn. Những mô hình hành vi này trở thành nhiên liệu để các thuật toán ngày càng tối ưu hơn.

Điều này tạo ra một sự dịch chuyển quyền lực quan trọng. Người sở hữu dữ liệu hành vi không nhất thiết sở hữu điểm đến, nhưng lại sở hữu khả năng dự đoán và điều phối dòng khách đến điểm đến đó. Càng nhiều giao dịch diễn ra trên một nền tảng,

nền tảng càng hiểu người dùng, càng hiểu người dùng, nền tảng càng có khả năng ảnh hưởng đến những giao dịch tiếp theo. Đây là vòng lặp học hỏi khiến dữ liệu hành vi trở thành một tài sản chiến lược của nền kinh tế tác nhân.

### 8.5.4. Khi không ai là người tích hợp

Vấn đề lớn nhất của nhiều địa phương không phải là thiếu dữ liệu mà là không có ai chịu trách nhiệm tích hợp dữ liệu. Chính quyền quản lý cơ sở dữ liệu hành chính. Doanh nghiệp quản lý hệ thống bán hàng và vận hành. Các nền tảng lưu giữ dữ liệu hành vi. Mỗi bên đều sở hữu một phần của bức tranh, nhưng không bên nào nhìn thấy toàn bộ bức tranh.

Kết quả là cùng một điểm đến xuất hiện dưới nhiều phiên bản khác nhau. Tên gọi không thống nhất, tọa độ khác nhau, giờ mở cửa không đồng bộ, thông tin liên hệ mâu thuẫn, giá và tình trạng phục vụ cập nhật không cùng thời điểm. Hệ thống phải ghép nối những mảnh rời rạc và thường không đủ bằng chứng để xác định đâu là thông tin chính xác nhất. Khi mức không chắc chắn vượt ngưỡng chấp nhận, hệ thống có thể giảm mức ưu tiên, yêu cầu xác minh thêm hoặc loại thực thể khỏi một số bước có rủi ro cao.

Đây là khoảng trống mà nhiều địa phương chưa nhận ra. Họ đầu tư xây dựng thêm trang web, thêm ứng dụng, thêm cơ sở dữ liệu mới, nhưng không đầu tư cho lớp tích hợp, nơi hợp nhất danh tính, dữ liệu vận hành và dữ liệu hành vi thành một mô hình thống nhất. Thiếu lớp này, mỗi hệ thống tiếp tục phát triển theo cách riêng, làm gia tăng phân mảnh thay vì giảm phân mảnh.

Đến đây có thể nảy ra một cách hiểu quá đơn giản, rằng chỉ cần dựng một nền tảng chung là xong. Bốn chủ thể không chỉ giữ những mảnh dữ liệu khác nhau. Họ có động cơ khác nhau và động cơ mới là phần khó.

Năng lực cạnh tranh của địa phương không chỉ được đo bằng chất lượng từng nguồn dữ liệu mà còn bằng khả năng hợp nhất các nguồn dữ liệu thành một hệ thống duy nhất. Vai trò mới của chính quyền vì thế không phải là sở hữu toàn bộ dữ liệu, mà là tổ chức kiến trúc tích hợp, thiết lập tiêu chuẩn chia sẻ và tạo cơ chế phối hợp giữa doanh nghiệp, nền tảng và các đơn vị quản lý. Chỉ khi có một lớp tích hợp chung, địa phương mới hình thành được song sinh số và đồ thị tri thức đủ đầy để AI có thể đọc, suy luận và hành động một cách tin cậy. Khi không ai chịu trách nhiệm hợp nhất, mỗi hệ thống mới lại thêm một phiên bản của cùng một sự thật.

### 8.5.5. Cơ chế liên thông tín hiệu địa phương

Một hướng tiếp cận phù hợp là chuyển từ tư duy quản lý dữ liệu tập trung sang tư duy liên minh dữ liệu (data alliance). Đây là đề xuất của tác giả, không phải mô tả một xu thế đã được đo lường. Mục tiêu của liên minh không phải là gom toàn bộ dữ liệu về một nơi, mà là xây dựng các tiêu chuẩn chung để mọi thành viên có thể chia sẻ và đồng bộ những phần dữ liệu cần thiết, trong khi vẫn duy trì quyền kiểm soát và trách nhiệm đối với phần dữ liệu của mình. Chính quyền không cần quản lý tồn kho khách sạn, doanh nghiệp không cần lưu cơ sở dữ liệu hành chính, nền tảng không cần sở hữu hồ sơ di sản. Điều cần thiết là tất cả đều tham chiếu đến cùng một hệ thống thực thể và cùng một ngôn ngữ dữ liệu.

Trong mô hình đó, vai trò của từng thành phần trở nên rõ ràng.

- ***Chính quyền xác lập danh tính số của địa phương,*** ban hành tiêu chuẩn dữ liệu, xác thực thực thể và điều phối toàn bộ hệ sinh thái.

- ***Doanh nghiệp cập nhật liên tục dữ liệu vận hành*** như giá, tồn kho, lịch hoạt động, khả năng phục vụ và các dịch vụ có thể giao dịch.

- ***Nền tảng phản hồi tín hiệu từ thị trường*** thông qua dữ liệu tìm kiếm, hành vi, đánh giá, xu hướng và hiệu quả phân phối.

- ***Cộng đồng bổ sung lớp tri thức sống*** của địa phương bằng nội dung do người dùng tạo ra, phản hồi, hình ảnh và những thay đổi hằng ngày mà không cơ quan nào cập nhật kịp, thông qua những kênh đóng góp có lưu vết, đối chiếu và xác minh.

Doanh nghiệp không có động cơ chia sẻ dữ liệu tồn kho cho chính quyền, nên liên minh dữ liệu là ý tưởng khó khả thi. Phản bác này mô tả đúng thực tế. Tồn kho và giá là tài sản cạnh tranh. Không doanh nghiệp nào chia sẻ chúng vì mục tiêu quản lý nhà nước. Nó chỉ sai ở giả định về phạm vi dữ liệu cần chia sẻ. Để một tác nhân biết một khách sạn còn khả năng nhận khách, hệ thống không cần biết còn bao nhiêu phòng và giá bao nhiêu. Nó cần biết cơ sở này đang hoạt động, đang nhận khách và có kênh đặt chỗ hoạt động. Đó là tín hiệu trạng thái, không phải dữ liệu thương mại. Ranh giới giữa hai loại này là điều cần được thiết kế ngay từ đầu, vì một liên minh yêu cầu quá nhiều sẽ không có ai tham gia, còn một liên minh yêu cầu quá ít sẽ không tạo được khả hành động.

Khi bốn nguồn dữ liệu này vận hành độc lập, địa phương chỉ có những mảnh thông tin rời rạc. Khi chúng được kết nối, địa phương hình thành một song sinh số luôn được cập nhật và một đồ thị tri thức phản ánh đúng thực trạng vận hành. AI không còn phải suy đoán từ các nguồn dữ liệu mâu thuẫn mà có thể dựa trên một hệ sinh thái thông tin thống nhất, đáng tin cậy và liên tục được làm mới.

Vì vậy, lợi thế cạnh tranh của địa phương trong nền kinh tế tác nhân sẽ không chỉ thuộc về nơi có nhiều dữ liệu, mà ngày càng thuộc về nơi hợp tác và tích hợp dữ liệu tốt nhất. Cuộc cạnh tranh không còn là giữa các cơ quan hay doanh nghiệp riêng lẻ, mà là giữa các hệ sinh thái dữ liệu. Địa phương nào xây dựng được một cơ chế liên thông tín hiệu hiệu quả sẽ có khả năng xuất hiện nhiều hơn trong quá trình suy luận của AI, được lựa chọn nhiều hơn trong các giao dịch tự động và giữ được vị thế trong chuỗi giá trị số của tương lai.

Nếu bạn là doanh nghiệp dịch vụ tại địa phương. Tách rõ hai loại dữ liệu trong hệ thống của mình, tín hiệu trạng thái có thể chia sẻ và dữ liệu thương mại không chia sẻ. Việc tách này là điều kiện để tham gia liên minh dữ liệu mà không mất lợi thế cạnh tranh. Tham chiếu định danh thực thể do tỉnh công bố trong hồ sơ của mình, thay vì tự đặt tên và địa chỉ theo cách riêng. Báo lỗi dữ liệu về đầu mối của xã hoặc tỉnh khi phát hiện sai lệch, thay vì chỉ sửa trên trang của mình.

## 8.6. GIÀU DI SẢN, NGHÈO DỮ LIỆU

Có lẽ nghịch lý lớn nhất của Việt Nam trong kỷ nguyên AI không phải là thiếu tài nguyên, mà là khoảng cách rất lớn giữa tài sản trải nghiệm và mức độ hiện diện trước máy. Chúng ta sở hữu hàng nghìn di tích lịch sử, hàng chục di sản được thế giới công nhận, hàng nghìn lễ hội, làng nghề, món ăn, nghệ thuật dân gian và cảnh quan tự nhiên đặc sắc. Tuy nhiên, phần lớn những tài sản đó vẫn tồn tại dưới dạng ký ức văn hóa, tư liệu giấy hoặc thông tin rời rạc, thay vì trở thành dữ liệu có cấu trúc mà AI có thể đọc, hiểu và sử dụng.

Nói cách khác, Việt Nam là một quốc gia giàu nội dung nhưng lại hạn chế về khả đọc. Một ngôi làng cổ có thể nổi tiếng với người dân địa phương nhưng không có thực thể số rõ ràng. Một lễ hội có lịch tổ chức hằng năm nhưng thiếu dữ liệu chuẩn về thời gian, địa điểm và khả năng tham gia. Một món ăn truyền thống có lịch sử hàng trăm năm nhưng lại xuất hiện khá mờ nhạt trong các đồ thị tri thức toàn cầu với đầy đủ mối quan hệ, biến thể và bối cảnh văn hóa. Đối với con người, những giá trị này vẫn hiện hữu. Đối với AI, nhiều khi chúng gần như vô hình.

Đó là thách thức trung tâm của giai đoạn tới. Bài toán không còn là tạo thêm tài nguyên mà là biểu diễn tài nguyên đã có theo cách máy đọc được và đây là việc đòi hỏi tiêu chuẩn cùng quy trình nhiều hơn là nguồn kinh phí.

Trong tập lựa chọn của máy, một tài sản chưa được biểu diễn thành thực thể, thuộc tính và quan hệ có vị thế tương đương một tài sản không tồn tại. Giá trị thực của tài sản không bù được sự thiếu vắng của lớp biểu diễn.

Khối tài sản trải nghiệm đã được định danh một phần. Đến giữa năm 2026, cả nước có hơn 21.000 sản phẩm thuộc chương trình mỗi xã một sản phẩm được xếp hạng ba sao trở lên.[1] Đây là một trong ít khối tài sản trải nghiệm của Việt Nam đã có danh mục, có xếp hạng và có chủ thể chịu trách nhiệm, tức đã đi được bước đầu tiên của khả đọc. Phần còn thiếu là lớp vận hành và lớp giao dịch: phần lớn các sản phẩm này chưa gắn với một thực thể có giờ đón khách, có tọa độ chính xác và có kênh đặt trực tiếp mà máy đọc được.

### 8.6.1. Di sản nhiều không đồng nghĩa được AI biết đến

Một quốc gia có thể sở hữu hàng nghìn năm lịch sử nhưng điều đó không có nghĩa AI sẽ biết đến từng giá trị của quốc gia ấy. AI không học từ giá trị nội tại của di sản mà học từ dấu vết dữ liệu mà di sản để lại trong thế giới số. Một ngôi đình có lịch sử năm trăm năm nhưng chỉ có vài dòng mô tả trên internet sẽ "nhỏ" hơn trong mắt AI so với một quán cà phê mới mở nhưng có dữ liệu đầy đủ, hàng nghìn đánh giá và thông tin được cập nhật mỗi ngày.

Lợi thế của Việt Nam vì thế không tự động đến từ việc sở hữu nhiều di sản. Nó đến từ việc chuyển những di sản ấy thành thực thể có định danh, có thuộc tính và có quan hệ, tức thành thứ mà một hệ thống có thể so sánh và đưa vào một hành trình cụ thể.

1. Bộ Nông nghiệp và Môi trường, số liệu chương trình mỗi xã một sản phẩm đến giữa năm 2026. Số liệu quản lý nhà nước, tổng hợp từ các quyết định phân hạng của địa phương.

### 8.6.2. Từ nội dung truy hồi được đến thực thể giao dịch được

Con người có thể quyết định đi đến một nơi vì nghe bạn bè kể, vì đọc một cuốn sách hay vì tình cờ nhìn thấy trên đường. Một tác nhân chọn trong phạm vi những gì nó truy cập, xác minh và so sánh được. Nhưng cần phân biệt ba mức tồn tại khác nhau, vì gộp chúng lại sẽ dẫn tới một phát biểu quá mạnh và dễ bị phản bác.

Mức thứ nhất là tồn tại dưới dạng nội dung có thể truy hồi. Một bài báo, một đoạn video, một cuốn sách đã được số hóa đều thuộc mức này và máy hoàn toàn có thể biết đến một làng nghề từ những nguồn ấy. Mức thứ hai là tồn tại dưới dạng thực thể có cấu trúc, tức có định danh, tọa độ, phạm vi và quan hệ. Mức thứ ba là tồn tại dưới dạng khả năng giao dịch, tức có giờ hoạt động, giá và kênh đặt chỗ mà máy gọi được.

Khoảng cách giữa ba mức là nơi phần lớn tài sản của Việt Nam đang mắc lại. Một làng nghề được nhắc trong hàng trăm bài báo có thể đạt mức thứ nhất mà không đạt mức thứ hai, nên khó được đưa vào một lịch trình có thể xác minh và thực hiện. Một lễ hội có tên trong danh mục văn hóa có thể đạt mức thứ hai mà không đạt mức thứ ba, nên tác nhân chưa thể hoàn tất việc đăng ký hoặc đặt chỗ. Phát biểu chính xác vì thế không phải là những gì không được số hóa thì không tồn tại, mà là: những gì không được biểu diễn thành dữ liệu đủ xác định, đủ mới và đủ khả hành động rất khó có mặt trong tập lựa chọn thực tế của một tác nhân.

Khoảng cách giữa tồn tại trong thực tế và tồn tại trong tập lựa chọn vì thế chính là khoảng cách giữa di sản và dữ liệu. Mỗi tài sản được lựa chọn để công khai và biểu diễn đúng cách sẽ thu hẹp khoảng cách ấy một phần.

### 8.6.3. Địa phương nhỏ có cơ hội mới

Trong thời đại truyền thông đại chúng, địa phương có ngân sách quảng bá lớn thường chiếm ưu thế. Trong nhiều nhiệm vụ có tính lập kế hoạch và giao dịch, dữ liệu đầy đủ, đáng tin cậy và khả hành động có thể giúp một địa phương nhỏ thu hẹp một phần bất lợi về ngân sách. Cần nói rõ mức độ của phát biểu này: các hệ thống xếp hạng vẫn chịu ảnh hưởng của độ phổ biến, dữ liệu hành vi, sức mạnh thương hiệu, quan hệ thương mại và cơ chế xếp hạng riêng của từng nền tảng. Dữ liệu tốt không xóa được những yếu tố đó, nó chỉ giành lại phần cơ hội mà dữ liệu kém đang làm mất.

Một xã miền núi có thể không đủ ngân sách để tổ chức các chiến dịch truyền thông quy mô quốc gia, nhưng hoàn toàn có thể xây dựng một cơ sở dữ liệu chuẩn cho toàn bộ điểm đến của mình. Nếu mỗi homestay đều có dữ liệu tồn kho theo thời gian thực, mỗi điểm tham quan có thực thể số đầy đủ, mỗi lễ hội có lịch tổ chức chuẩn hóa và toàn bộ dịch vụ đều có thể được AI đọc và đặt trực tiếp, địa phương đó có thể trở thành lựa chọn ưu tiên trong nhiều tình huống, dù ngân sách quảng bá chỉ bằng một phần rất nhỏ so với các trung tâm du lịch lớn.

Đây là chuyển dịch từ ưu tiên truyền thông sang ưu tiên khả đọc. Trong nhiều trường hợp, chi phí chuẩn hóa lớp dữ liệu tối thiểu có thể thấp hơn chi phí duy trì truyền thông dài hạn và hiệu quả bền hơn vì dữ liệu tiếp tục được dùng lại trong rất nhiều lần suy luận sau đó.

Dữ liệu tốt không xóa được bất lợi về thương hiệu, độ phổ biến hay quyền lực nền tảng. Nó chỉ giúp địa phương nhỏ giành lại phần cơ hội đang bị mất vì dữ liệu kém. Đây là một giả thuyết

chứ không phải kết luận đã kiểm chứng và cả hai chiều lập luận được ghi như đã nêu ngay dưới đây.

Lợi thế quy mô nhỏ. Một địa phương ít thực thể có thể đạt độ phủ và độ tươi cao hơn một địa phương nhiều thực thể với cùng một mức nỗ lực, vì khối lượng phải chuẩn hóa nhỏ hơn. Nếu đúng, khả đọc sẽ là một trong ít lĩnh vực mà địa phương nhỏ có thể vượt địa phương lớn. Cần lưu ý chiều ngược lại cũng có lý: nơi ít thực thể thường cũng ít nhân lực chuyên trách và ít doanh nghiệp có hệ thống dữ liệu, nên lợi thế quy mô nhỏ có thể bị triệt tiêu bởi thiếu hụt năng lực. Giả thuyết này chưa được kiểm chứng.

### 8.6.4. Việt Nam không thiếu tài nguyên, thiếu lớp biểu diễn

Điểm yếu lớn nhất của Việt Nam không nằm ở tài nguyên trải nghiệm mà nằm ở lớp biểu diễn số (representation layer). Chúng ta có cảnh quan, văn hóa, ẩm thực, di sản, con người và những câu chuyện đặc sắc. Điều còn thiếu là một cách biểu diễn thống nhất để toàn bộ những giá trị đó có thể được máy nhận diện và xử lý.

Đối với AI, một điểm đến không tồn tại dưới dạng phong cảnh hay truyền thuyết, mà tồn tại dưới dạng thực thể, thuộc tính và quan hệ. Một lễ hội không chỉ là một sự kiện văn hóa mà còn phải có thời gian, địa điểm, chu kỳ tổ chức, quy mô, đối tượng tham gia và liên kết với các dịch vụ xung quanh. Một món ăn không chỉ cần hình ảnh đẹp mà còn cần được kết nối với vùng nguyên liệu, nhà hàng, làng nghề, lịch sử hình thành và các trải nghiệm liên quan. Một làng nghề không chỉ là một cái tên mà phải được mô tả bằng dữ liệu mà AI có thể liên kết với tuyến du lịch, nghệ nhân, sản phẩm và hoạt động trải nghiệm.

Thiếu lớp biểu diễn, những giá trị ấy chỉ tồn tại như các đảo dữ liệu rời rạc: có trang web nhưng không có thực thể, có hình ảnh nhưng không có ngữ nghĩa, có thực thể nhưng không có quan hệ, có quan hệ nhưng không có khả năng giao dịch. Vì vậy, nhiệm vụ của giai đoạn tới không chỉ là số hóa hồ sơ hay xây thêm cổng thông tin, mà là xây lớp biểu diễn số của địa phương, nơi mọi tài sản, dịch vụ, sự kiện và cộng đồng đều được mô hình hóa thành dữ liệu đọc được, liên kết được và hành động được.

Đây là thông điệp lớn của cả chương. Chúng ta không bước vào kỷ nguyên AI với điểm xuất phát nghèo nàn. Điều còn thiếu không phải thêm di sản, thêm lễ hội hay thêm điểm đến, mà là một lớp biểu diễn đủ tốt để những tài sản ấy được nhìn thấy. Chúng ta không nghèo di sản. Chúng ta đang thiếu nghèo lớp biểu diễn những di sản ấy.

Đầu tư cho khả đọc là làm việc không lương cho các nền tảng nước ngoài. Địa phương bỏ chi phí chuẩn hóa, nền tảng thu phần lớn giá trị. Trả lời: Đây là phản bác mạnh nhất trong chương này. Nó không có câu trả lời trọn vẹn. Chương 6 đã chỉ ra rằng phần giá trị lớn nhất trong nền kinh tế tác nhân chảy về lớp kiểm soát xếp hạng và giao dịch, không về lớp tạo ra trải nghiệm. Chi phí chuẩn hóa dữ liệu do địa phương chịu, còn lợi ích từ dữ liệu đã chuẩn hóa thì được chia lại theo tương quan quyền lực hiện có và tương quan đó không có lợi cho địa phương. Ba điểm có thể nói được. Thứ nhất, phần lớn hạng mục chuẩn hóa nêu trong chương này cũng là điều kiện để chính địa phương vận hành được hệ thống của mình, nên nó không phải chi phí thuần cho bên ngoài. Thứ hai, dữ liệu đã chuẩn hóa là tài sản khả chuyển, nó có thể được cấp cho nền tảng khác khi tương quan thay đổi, khác với dữ liệu bị khóa trong một hệ thống. Thứ ba, không chuẩn hóa cũng không giúp giữ giá trị, nó chỉ chuyển địa phương từ vị thế thương lượng yếu sang vị thế không có mặt. Nhưng cả ba điểm đó đều không phủ nhận nội dung của phản bác.

## 8.7. KẾT LUẬN CHƯƠNG

Việt Nam không cần bắt đầu cuộc cạnh tranh trong nền kinh tế tác nhân bằng một mô hình AI riêng hay một hạ tầng điện toán khổng lồ. Điểm khởi đầu thực tế hơn nằm ở khả năng tổ chức lại những gì đất nước đã có: hợp nhất danh tính, duy trì lịch sử dữ liệu, bổ sung lớp vận hành, trao quyền cập nhật cho nơi gần sự thật nhất và xây dựng cơ chế tích hợp giữa chính quyền, doanh nghiệp, nền tảng và cộng đồng. Thứ tự ấy quan trọng. Danh tính phải đi trước dữ liệu vận hành, dữ liệu vận hành phải đi trước truyền thông. Làm ngược lại thì hình ảnh mới sẽ tiếp tục được dựng trên một nền dữ liệu cũ và phân mảnh.

Nhưng khả đọc mới chỉ là điều kiện để có mặt, chưa phải điều kiện để giữ quyền lực hay giữ giá trị. Bên bỏ chi phí chuẩn hóa chưa chắc là bên hưởng lợi nhiều nhất từ dữ liệu đã chuẩn hóa. Một địa phương có thể được tác nhân nhìn thấy, hiểu đúng và giao dịch được nhưng vẫn phụ thuộc vào nền tảng đang kiểm soát việc xếp hạng, phân phối và quan hệ với người dùng. Làm tốt những việc của chương này không tự động đưa địa phương vào vị thế mạnh, nó chỉ giúp địa phương bước vào bàn thương lượng thay vì tiếp tục đứng ngoài tập lựa chọn.

Và khi một địa phương đã có khả năng tự biểu diễn, một lớp câu hỏi mới lập tức xuất hiện. Ai có quyền định nghĩa hình ảnh số của nơi đó? Ai được sửa nó? Ai được sử dụng nó? Ai chịu trách nhiệm khi máy mô tả sai hoặc hành động sai dựa trên một mô tả lệch? Chương này trả lời cách để địa phương được máy đọc đúng. Chương sau bắt đầu từ quyền của địa phương đối với chính hình ảnh số ấy: quyền định nghĩa, quyền giám sát và quyền sửa sai. Vô hình trong dữ liệu, sớm hay muộn, sẽ thành vô hình trong kinh tế.

## PHỤ LỤC 8A. DANH MỤC VIỆC CẦN LÀM TRONG MƯỜI HAI THÁNG

Danh mục này không phải một kế hoạch chuẩn để áp dụng cho mọi địa phương. Mỗi hạng mục đều kèm một đầu ra kiểm tra được, để nó có thể kiểm tra được thay vì chỉ được báo cáo là đã làm. Nó là cách sắp xếp thứ tự ưu tiên theo logic của chương: hợp nhất danh tính trước, dữ liệu vận hành sau, truyền thông sau cùng. Thứ tự này quan trọng hơn thời hạn cụ thể.

- ***Chín mươi ngày đầu:***

• *Kiểm kê toàn bộ tài sản trải nghiệm trên địa bàn, ghi rõ ai biết trạng thái thực của từng cơ sở. Chủ trì: Xã, tỉnh tổng hợp. Đầu ra kiểm tra được: Danh mục thực thể có định danh ổn định, tọa độ và tên đầu mối chịu trách nhiệm từng cơ sở.*

• *Lập bảng ánh xạ từ mọi tên cũ sang tên mới, kèm ngày hiệu lực. Chủ trì: Tỉnh. Đầu ra kiểm tra được: Tệp công bố ở định dạng máy đọc được, có địa chỉ truy cập ổn định.*

- ***Tới hết tháng thứ sáu:***

• *Trao quyền ghi trực tiếp cho cấp xã đối với các trường biến động nhanh. Chủ trì: Tỉnh quyết định, xã thực hiện. Đầu ra kiểm tra được: Văn bản phân quyền theo trường dữ liệu, kèm thời hạn cập nhật cho từng loại.*

• *Dựng lớp tích hợp, chuyển hướng địa chỉ web cũ, hợp nhất bản ghi trùng trên các nền tảng ngoài. Chủ trì: Tỉnh. Đầu ra kiểm tra được: Không còn địa chỉ cũ trả về lỗi, mỗi điểm đến chỉ còn một bản ghi chính trên từng nền tảng.*

- ***Tới hết tháng thứ chín:***

• *Hoàn thiện thuộc tính vận hành cho nhóm tài sản chính. Chủ trì: Xã, doanh nghiệp. Đầu ra kiểm tra được: Tỷ lệ điểm*

*đến có đủ giờ mở cửa, giá, trạng thái hoạt động và điều kiện tiếp cận.*

- *Chốt ranh giới chia sẻ dữ liệu với doanh nghiệp. Chủ trì: Tỉnh. Đầu ra kiểm tra được: Thỏa thuận liên minh nêu rõ trường nào chia sẻ và trường nào không.*

▪ ***Ba tháng cuối:***

- *Vận hành đầu mối nhận báo lỗi với thời hạn xử lý công khai. Chủ trì: Xã, tỉnh. Đầu ra kiểm tra được: Nhật ký xử lý báo lỗi, thời gian phản hồi tính theo phân vị.*
- *Kiểm toán khả đọc và chỉ khi đó mới triển khai nhận diện thương hiệu mới. Chủ trì: Tỉnh. Đầu ra kiểm tra được: Báo cáo năm chỉ số đo tự động: độ tươi, độ phủ, tính nhất quán, khả năng dùng được, thời gian phản hồi.*

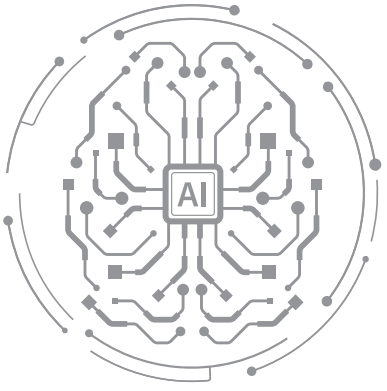
Danh mục dưới đây cần đọc theo hai tầng, nếu không nó sẽ biến một tỉnh thành một công ty công nghệ. Tầng bắt buộc gồm những việc mọi địa phương đều làm được mà không cần năng lực kỹ thuật đặc biệt: một mã định danh chuẩn, một bảng tên thay thế, một nguồn công bố ổn định, một đầu mối báo lỗi, dấu thời gian trên mỗi trường dữ liệu và phân quyền cập nhật rõ ràng. Tầng có điều kiện chỉ nên làm khi đã có ngân sách duy trì và người chịu trách nhiệm: giao diện vận hành, tồn kho thời gian thực, khả năng giao dịch bằng máy, đàm phán dữ liệu với nền tảng và kiểm toán đa tác nhân nâng cao. Một địa phương bỏ tầng thứ nhất để nhảy sang tầng thứ hai sẽ tiêu tiền vào một hạ tầng không có móng.

***Ghi chú:*** *Thứ tự quan trọng hơn thời hạn. Hợp nhất danh tính đi trước dữ liệu vận hành, dữ liệu vận hành đi trước truyền thông. Đảo thứ tự này là sai lầm phổ biến nhất sau mỗi lần thay đổi địa giới.*

## CHƯƠNG 9

# CHỦ QUYỀN VÀ TRÁCH NHIỆM

*Ai sở hữu hình ảnh của một địa phương, chủ quyền dữ liệu khác chủ quyền thuật toán ở đâu, quyền mờ như một lựa chọn quản trị và chuỗi trách nhiệm khi một tác nhân đặt sai.*


AI

Trước khi đi tiếp, một giới hạn cần đặt ra về chính chữ chủ quyền. Chủ quyền trong nghĩa pháp lý gắn với thẩm quyền tối cao, lãnh thổ và khả năng cưỡng chế. Một tỉnh không có và không thể có chủ quyền theo nghĩa ấy đối với cách một mô hình tư nhân suy luận về mình. Chữ chủ quyền trong chương này có nghĩa hẹp hơn: năng lực quản trị và tự quyết tương đối đối với hình ảnh số của mình, gồm quyền yêu cầu sửa dữ kiện sai kiểm chứng được, quyền mang dữ liệu của mình đi nơi khác và quyền được nhận diện nhất quán giữa các hệ thống.

Sự phân biệt này có hệ quả thực hành. Quyền yêu cầu sửa một địa chỉ sai, một mức giá lỗi thời hay một thực thể bị nhầm là quyền có cơ sở và nên được thiết chế hóa. Quyền can thiệp vào một đánh giá, một suy luận xác suất hay một quan điểm phê bình thì không và một địa phương đòi quyền đó sẽ mất luôn tính chính đáng của những yêu cầu hợp lý còn lại. Đầu mối hiệu chỉnh được đề nghị ở tiết 9.7 chỉ nên xử lý nhóm thứ nhất.

Chương cuối trở lại chi tiết nhỏ nhất và cũng khó nhất của mô hình Chương 2: cái van. Bốn bộ lọc nói về việc được đọc. Cái van nói về quyền không được đọc hết. Toàn bộ phần bàn về quyền mờ trong chương này là phần bàn về cái van ấy, lần này không còn ở tầng khái niệm mà ở tầng thẩm quyền: ai được quyết định phần nào của một vùng đất đi vào lược đồ và phần nào ở lại bên ngoài. Nếu bốn bộ lọc là câu hỏi về năng lực, cái van là câu hỏi về chủ quyền.

Chủ quyền không phải quyền được đọc đúng. Chủ quyền là quyền quyết định mình được đọc tới đâu. Một quyền không thể sửa sai thì chỉ là một tuyên bố. Trong nền kinh tế tác nhân, chủ quyền của một địa phương được đo bằng khoảng cách từ lúc máy mô tả sai đến lúc mô tả ấy được sửa.

Chương 1 xác lập rằng một phần quyền sàng lọc và dựng tập lựa chọn đang chuyển từ con người sang máy. Chương 2 giải thích điều kiện để một địa phương trở nên khả đọc và Chương 3 mô tả kiến trúc mà máy sử dụng để hiểu thế giới. Các Chương 4, 5 và 6 phân tích cách niềm tin, dòng khách và giá trị kinh tế được phân phối lại. Chương 7 chỉ ra cách một địa phương phải thiết kế lại chính mình để máy có thể nhận diện, truy vấn và thực hiện giao dịch với các dịch vụ của mình, còn Chương 8 đặt yêu cầu ấy vào bối cảnh Việt Nam với dấu thanh, địa danh và năng lực cấp xã. Chương cuối cùng này đặt ra câu hỏi lớn nhất của toàn bộ cuốn sách: ai có quyền quyết định hình ảnh số của địa phương và ai chịu trách nhiệm khi hình ảnh ấy sai?

Đó không còn là câu hỏi của công nghệ, cũng không chỉ là câu hỏi của quản trị dữ liệu. Đó là câu hỏi về chủ quyền trong một thời đại mà ngày càng nhiều quyết định kinh tế bắt đầu bằng một quá trình suy luận của máy.

Trong nhiều thế kỷ, chủ quyền của một địa phương được xác định bằng những yếu tố hữu hình: lãnh thổ, dân cư, pháp luật, bộ máy hành chính và quyền thực thi quyền lực công. Những yếu tố ấy vẫn giữ nguyên giá trị. Tuy nhiên, cùng với sự xuất hiện của các hệ thống AI và tác nhân tự động, địa phương đồng thời tồn tại trong một không gian khác: không gian của dữ liệu và biểu diễn số. Ở đó, một thành phố không xuất hiện dưới dạng quảng trường, bảo tàng hay dòng sông, mà dưới dạng thực thể, quan hệ, thuộc tính, mô tả, bản đồ tri thức và các khả năng giao dịch mà máy có thể đọc, so sánh và hành động.

Đối với con người, hình ảnh của một địa phương được hình thành qua trải nghiệm, ký ức và cảm xúc. Đối với máy, hình ảnh ấy được hình thành từ dữ liệu, từ các quan hệ giữa thực thể, từ những nguồn được coi là có thẩm quyền và từ vô số vòng lặp suy luận diễn ra liên tục. Khi ngày càng nhiều quyết định về du lịch, đầu tư, tiêu dùng hay dịch chuyển được ủy nhiệm cho tác nhân AI, quyền kiểm soát lớp biểu diễn số này dần trở thành một dạng chủ quyền mới. Không phải chủ quyền đối với lãnh thổ, mà là chủ quyền đối với cách địa phương được nhận diện, được diễn giải và được hành động trong thế giới số.

Tuy nhiên, chương này cũng đưa ra một lập luận quan trọng hơn, chủ quyền chỉ thực sự tồn tại khi đi kèm trách nhiệm. Một địa phương có thể tuyên bố mình sở hữu dữ liệu, sở hữu thương hiệu hay sở hữu hình ảnh chính thức, nhưng nếu khi một hệ thống AI mô tả sai địa phương, dẫn khách đến nhầm nơi, gán nhầm di sản, hiểu sai quy định hay tạo ra một chuỗi suy luận gây tổn hại đến danh tiếng mà không tồn tại bất kỳ cơ chế nào để phát hiện, phản hồi và sửa chữa, thì quyền đó mới chỉ là một tuyên bố mang tính hình thức.

Chính vì vậy, chương này không xem chủ quyền và trách nhiệm là hai chủ đề song song. Trách nhiệm không phải phần bổ sung của chủ quyền mà là phép thử của chủ quyền. Một quyền không thể thực thi, không thể bảo vệ và không thể khắc phục khi bị xâm phạm thì chưa phải là một quyền hoàn chỉnh. Nói cách khác, khả năng sửa sai cũng quan trọng không kém khả năng định nghĩa ban đầu.

Từ lập luận đó, phần đầu của chương sẽ làm rõ khái niệm chủ quyền thuật toán và phân biệt nó với chủ quyền dữ liệu truyền thống. Cuốn sách sẽ chỉ ra rằng điều quan trọng không chỉ là ai lưu giữ dữ liệu, mà còn là ai có quyền định nghĩa thực thể,

ai quyết định cách dữ liệu được biểu diễn, dữ liệu có thể được mang sang hệ thống khác hay không và các hệ thống có thể liên thông với nhau như thế nào. Đồng thời, chương cũng lập luận rằng không phải mọi dữ liệu đều nên được tối đa hóa khả đọc, một phần giá trị của địa phương nằm ở quyền được giữ lại những khoảng mờ, những yếu tố văn hóa, thiêng liêng và bản sắc không nên bị nén hoàn toàn thành cấu trúc dữ liệu.

Phần thứ hai chuyển trọng tâm từ quyền sang trách nhiệm. Khi một tác nhân AI đặt sai dịch vụ, mô tả sai điểm đến hoặc tạo ra thiệt hại kinh tế và thương hiệu, chuỗi trách nhiệm sẽ thuộc về ai? Nhà cung cấp dữ liệu, đơn vị vận hành dịch vụ, nền tảng trung gian, nhà phát triển mô hình hay chính quyền địa phương? Những câu hỏi ấy hiện vẫn chưa có lời giải rõ ràng ở phần lớn quốc gia, nhưng chúng sẽ nhanh chóng trở thành những vấn đề quản trị thường nhật khi tác nhân AI bắt đầu thay mặt con người thực hiện hàng triệu giao dịch mỗi ngày. Cuốn sách vì thế đề xuất rằng bên cạnh hạ tầng dữ liệu, các địa phương sẽ phải xây dựng cả hạ tầng quy trách nhiệm: quy trình kiểm toán tri thức, cơ chế hiệu chỉnh mô tả, kênh khiếu nại đối với hệ thống AI và bộ máy đủ thẩm quyền để duy trì sự chính xác của hình ảnh số.

Đây là chương có tính chính trị nhất của cuốn sách và cũng là chương khép lại toàn bộ lập luận của nó. Nếu các chương trước trả lời câu hỏi làm thế nào để một địa phương tồn tại trong nền kinh tế tác nhân, thì chương này trả lời câu hỏi cuối cùng: một địa phương chỉ thật sự có chủ quyền trong nền kinh tế tác nhân khi không chỉ được quyền định nghĩa cách mình được máy nhìn thấy, mà còn có năng lực bảo vệ, sửa chữa và chịu trách nhiệm đối với hình ảnh ấy. Đó cũng là thời điểm khái niệm chủ quyền được mở rộng khỏi biên giới địa lý để bước vào một không gian mới: không gian của dữ liệu, thuật toán và niềm tin số.

PHẦN A
# CHỦ QUYỀN

## 9.1. QUYỀN ĐỊNH NGHĨA HÌNH ẢNH ĐỊA PHƯƠNG

Trong phần lớn lịch sử, câu hỏi "ai sở hữu hình ảnh của một địa phương" dường như có câu trả lời khá rõ ràng. Chính quyền quản lý biểu tượng, tên gọi và thông tin chính thức, doanh nghiệp xây dựng thương hiệu dịch vụ, báo chí và các nhà làm truyền thông kể những câu chuyện, còn du khách tạo ra ký ức và truyền miệng. Hình ảnh của địa phương vì thế luôn là kết quả của nhiều chủ thể cùng tham gia, nhưng quyền định hình cuối cùng vẫn chủ yếu thuộc về con người.

Nền kinh tế tác nhân làm thay đổi cấu trúc ấy. Hình ảnh của một địa phương không còn chỉ tồn tại trong các bài viết, video quảng bá hay trang web du lịch. Nó còn tồn tại trong hàng nghìn cơ sở dữ liệu, kho tri thức, bản đồ số, đồ thị tri thức và mô hình ngôn ngữ mà các hệ thống AI sử dụng để suy luận. Khi một tác nhân AI được giao nhiệm vụ lập kế hoạch du lịch, tìm địa điểm đầu tư hay lựa chọn nơi tổ chức sự kiện, nó không "xem" địa phương theo cách con người xem. Nó đọc một thực thể số đã được chuẩn hóa, liên kết và xếp hạng từ trước. Điều đó khiến câu hỏi trung tâm không còn là ai quảng bá địa phương, mà là ai định nghĩa thực thể mà máy đang đọc.

Đây là một sự dịch chuyển rất lớn về bản chất của quyền lực. Trong nền kinh tế trải nghiệm, địa phương cạnh tranh để thu hút sự chú ý của con người. Trong nền kinh tế tác nhân, trước khi cạnh tranh để được lựa chọn, địa phương phải bảo đảm mình được các hệ thống AI nhận diện đúng. Một thành phố có thể đầu

tư rất nhiều vào truyền thông, nhưng nếu thực thể số của nó bị phân mảnh, mô tả sai hoặc bị một nguồn dữ liệu kém chất lượng chi phối, thì tác nhân AI vẫn có thể đưa ra những quyết định bất lợi. Nói cách khác, quyền lực không còn nằm hoàn toàn ở việc kể câu chuyện hấp dẫn nhất, mà ở việc xác lập được phiên bản có thẩm quyền của thực thể mà máy sử dụng.

Điều đáng chú ý là thực thể này thường không nằm dưới sự kiểm soát của một tổ chức duy nhất. Một địa phương ngày nay đồng thời tồn tại trên bản đồ số, trong cơ sở dữ liệu của các nền tảng đặt phòng, trong kho dữ liệu của các công cụ tìm kiếm, trong đồ thị tri thức của các mô hình AI và trong nhiều nguồn dữ liệu mở khác. Chỉ cần các phiên bản ấy bắt đầu khác nhau về tên gọi, ranh giới, quan hệ hay khả năng giao dịch, máy sẽ không còn đọc một địa phương duy nhất mà đọc nhiều thực thể song song và hình ảnh số bị chia nhỏ ngay từ tầng dữ liệu, trước khi bất kỳ hoạt động quảng bá nào bắt đầu.

Vì vậy, quyền sở hữu hình ảnh trong nền kinh tế tác nhân không nên được hiểu theo nghĩa độc quyền. Không một địa phương nào có thể sở hữu toàn bộ dữ liệu về chính mình, cũng không thể yêu cầu mọi hệ thống AI chỉ sử dụng một nguồn duy nhất. Điều địa phương có thể và cần sở hữu là quyền định nghĩa phiên bản chuẩn của chính mình: tên gọi chính thức, quan hệ giữa các thực thể, mô tả có thẩm quyền, dữ liệu nguồn và các quy trình để những hệ thống khác có thể đối chiếu và đồng bộ. Đây chính là nền tảng của chủ quyền thuật toán.

Chủ quyền của địa phương vì thế bắt đầu từ một quyền rất cơ bản nhưng trước đây ít được chú ý: quyền được định nghĩa chính mình trong thế giới số. Chỉ khi tồn tại một thực thể số thống nhất, có thẩm quyền và được các hệ thống khác thừa nhận, địa

phương mới có thể nói đến những tầng chủ quyền cao hơn như quyền kiểm soát dữ liệu, quyền bảo vệ thương hiệu hay quyền tham gia bình đẳng vào nền kinh tế tác nhân. Nếu ngay cả thực thể mà máy đang đọc còn không phải là thực thể do địa phương xác lập, thì phần lớn các cuộc cạnh tranh về thương hiệu sau đó thực chất đã bắt đầu từ một vị thế bất lợi.

Quyền định nghĩa. Trong nền kinh tế tác nhân, chủ thể định hình hình ảnh của một địa phương không phải chủ thể kể câu chuyện hay nhất mà chủ thể xác lập được thực thể mà máy coi là chuẩn. Truyền thông tác động lên lớp trên, quyền định nghĩa tác động lên lớp dưới và lớp dưới quyết định lớp trên có điểm tựa hay không.

Ba thiết chế đã từng giải quyết bài toán danh tính. Bài toán ai có thẩm quyền xác nhận một danh tính khi các bên tranh chấp không mới. Ít nhất ba thiết chế đã giải quyết nó trong ba lĩnh vực khác nhau. Với tên miền, ngành công nghiệp internet lập ra một quy trình giải quyết tranh chấp thống nhất từ năm 1999, gắn với hệ thống đăng ký tập trung.[1] Với báo chí, pháp luật nhiều nước quy định quyền yêu cầu cải chính kèm thời hạn và hình thức đăng cải chính, ở Việt Nam là Điều 34 Luật Báo chí số 126/2025/QH15, có hiệu lực từ ngày 1 tháng 7 năm 2026. Với dữ liệu cá nhân, quyền yêu cầu chỉnh sửa thông tin không chính xác được ghi trong Quy định bảo vệ dữ liệu chung của Liên minh châu Âu và trong Điều 4 Luật Bảo vệ dữ liệu cá nhân số 91/2025/QH15 của Việt Nam, có hiệu lực từ ngày 01 tháng 01 năm 2026. Cả ba thiết chế có chung ba thành phần: một sổ đăng ký để đối chiếu, một quy trình xác minh và một điểm kết thúc vụ việc. Tri thức mà tác nhân AI đang dùng để mô tả một địa phương hiện thiếu cả ba và đó là nội dung của phần B.

1. Quy trình giải quyết tranh chấp tên miền thống nhất do ICANN thông qua ngày 24 tháng 10 năm 1999.

### 9.1.1. Hình ảnh địa phương không còn nằm trong brochure

Trong phần lớn thế kỷ XX, hình ảnh của một địa phương chủ yếu được xây dựng thông qua những phương tiện truyền thông mà con người tiếp nhận trực tiếp: brochure du lịch, trang web giới thiệu, phim quảng bá, bài báo, sách hướng dẫn hay các chiến dịch marketing. Thành công của thương hiệu địa phương phụ thuộc vào việc có thể kể một câu chuyện hấp dẫn, tạo được cảm xúc và khiến du khách muốn đến trải nghiệm. Có thể nói, cuộc cạnh tranh diễn ra chủ yếu trên bề mặt của truyền thông.

Ngày nay, lớp truyền thông ấy vẫn tồn tại nhưng không còn là nơi đầu tiên mà quyết định được hình thành. Khi một người giao việc cho tác nhân AI: "hãy tìm giúp tôi một điểm đến phù hợp", hệ thống gần như không đọc brochure, không xem video quảng bá và cũng không duyệt từng trang web như con người. Thay vào đó, nó truy cập vào các cơ sở tri thức đã được chuẩn hóa, đối chiếu nhiều nguồn dữ liệu khác nhau, xác định các thực thể liên quan rồi suy luận từ mạng lưới quan hệ giữa chúng. Điều mà máy "nhìn thấy" trước tiên không phải là một cuốn cẩm nang du lịch mà là một đồ thị tri thức.

Sự thay đổi này làm dịch chuyển trung tâm của thương hiệu địa phương. Nếu brochure trả lời câu hỏi "chúng ta muốn kể câu chuyện gì?", thì đồ thị tri thức trả lời câu hỏi "máy đang tin điều gì về địa phương này?". Hai câu hỏi có thể dẫn đến hai kết quả rất khác nhau. Một thành phố có thể sở hữu bộ nhận diện thương hiệu đẹp, trang web hiện đại và những bộ phim quảng bá chất lượng cao, nhưng nếu trong đồ thị tri thức của máy, địa phương ấy thiếu quan hệ với các điểm tham quan, không có dữ liệu về giờ mở cửa, khả năng giao thông, dịch vụ hay các thực thể liên quan, thì tác nhân AI vẫn khó có thể đề xuất hoặc giao dịch với địa phương đó.

### 9.1.2. Một địa phương, nhiều bản sao trong thế giới của máy

Một địa phương không chỉ tồn tại một lần trong thế giới số. Mỗi hệ thống AI, nền tảng tìm kiếm, tác nhân AI hay mô hình ngôn ngữ có thể tạo hoặc truy xuất một cách biểu diễn riêng về cùng một địa phương. Có thể hình dung rằng thay vì chỉ có một "bản đồ chính thức", hiện nay đang tồn tại hàng nghìn bản sao của cùng một thành phố, một khu di tích hay một điểm đến trong các hệ thống khác nhau. Những bản sao ấy không hoàn toàn giống nhau, bởi chúng được hình thành từ những nguồn dữ liệu khác nhau, được cập nhật ở những thời điểm khác nhau và được tổ chức theo những kiến trúc tri thức khác nhau.

Ở cấp độ kỹ thuật, một địa phương có thể được biểu diễn dưới nhiều dạng: trong trọng số của mô hình, trong các vector ngữ nghĩa, trong chỉ mục tìm kiếm, trong đồ thị tri thức, hoặc trong những kho dữ liệu được truy xuất ngay tại thời điểm tác nhân thực hiện nhiệm vụ. Ở dạng vector ngữ nghĩa, thường gọi là biểu diễn vector (embedding), thông tin về địa phương được nén thành một vị trí trong không gian ngữ nghĩa: nơi có đặc điểm tương tự nằm gần nhau, nơi khác biệt nằm xa nhau. Cách biểu diễn ấy cho phép máy tìm kiếm, so sánh và suy luận mà không cần đọc lại toàn bộ dữ liệu gốc.

Bên cạnh đó, nhiều hệ thống còn kết nối với hoặc tự duy trì các kho tri thức chứa các thực thể, thuộc tính và quan hệ mà mô hình có thể tham chiếu khi trả lời hoặc thực hiện nhiệm vụ. Trong kho tri thức ấy, địa phương được liên kết với hàng loạt thực thể khác: di sản, khách sạn, nhà hàng, tuyến giao thông, lễ hội, nhân vật lịch sử, doanh nghiệp, cơ quan quản lý và nhiều lớp thông tin khác. Chính mạng lưới quan hệ này tạo nên khả năng suy luận của máy.

Quan trọng hơn, mỗi lần tác nhân AI thực hiện một nhiệm vụ, nó còn xây dựng một ngữ cảnh xử lý riêng (context) cho phiên làm việc đó. Nếu người dùng hỏi về du lịch, địa phương sẽ được đặt trong ngữ cảnh của trải nghiệm, điểm tham quan và đặt dịch vụ. Nếu câu hỏi liên quan đến đầu tư, cùng địa phương ấy lại được nhìn qua hạ tầng, nhân lực, quy hoạch và chính sách. Nếu nhiệm vụ là ứng phó thiên tai, các thuộc tính ưu tiên sẽ hoàn toàn khác. Điều này có nghĩa là hình ảnh của địa phương không phải một đối tượng cố định mà liên tục được tái cấu trúc theo mục tiêu suy luận của từng tác nhân.

Hệ quả là không tồn tại một hệ thống AI duy nhất đang đọc địa phương. Mỗi hệ thống đang tạo hoặc truy xuất một cách biểu diễn khác nhau về cùng một nơi. Khi những cách biểu diễn ấy ngày càng lệch nhau về dữ liệu, quan hệ hoặc mức độ cập nhật, hình ảnh số của địa phương sẽ bị phân mảnh. Vì vậy, mục tiêu không phải là tối ưu cho một nền tảng cụ thể mà là tạo ra một phiên bản chuẩn, đủ thẩm quyền và đủ khả năng liên thông để các hệ thống có thể đối chiếu và đồng bộ theo thời gian.

Trong nền kinh tế tác nhân, vì vậy, câu hỏi không còn là địa phương có xuất hiện trên internet hay không. Câu hỏi là bao nhiêu bản sao của địa phương đang tồn tại trong các hệ thống AI, chúng giống nhau đến mức nào và bản sao nào đang được máy tin tưởng nhất khi đưa ra quyết định.

### 9.1.3. Quyền định nghĩa dịch chuyển khỏi chính quyền

Trong phần lớn lịch sử hiện đại, chính quyền được mặc nhiên xem là chủ thể có quyền định nghĩa địa phương. Chính quyền quyết định tên gọi chính thức, địa giới hành chính, hệ thống

phân loại, quy hoạch, các biểu tượng và những thông tin được công bố ra công chúng. Đó là nền tảng của chủ quyền hành chính và cũng là cơ sở để xây dựng thương hiệu địa phương.

Nhưng trong nền kinh tế tác nhân, quyền định nghĩa bắt đầu dịch chuyển khỏi không gian hành chính truyền thống. Điều này không có nghĩa là chính quyền mất quyền pháp lý đối với địa phương, mà có nghĩa là quyền tạo ra phiên bản mà máy thực sự sử dụng để ra quyết định ngày càng nằm trong tay các hạ tầng số.

Một nền tảng đặt dịch vụ trực tuyến có thể quyết định địa phương nào xuất hiện đầu tiên khi người dùng tìm kiếm chỗ nghỉ. Một công cụ tìm kiếm có thể tổng hợp hàng trăm nguồn dữ liệu để tạo ra đoạn mô tả mà hàng triệu người đọc mỗi ngày. Một mô hình ngôn ngữ lớn có thể kết hợp dữ liệu từ nhiều kho tri thức khác nhau rồi sinh ra một câu trả lời hoàn toàn mới. Một nền tảng có thể thiết kế giao diện, tiêu chí xếp hạng và cơ chế gợi ý khiến một số địa phương luôn được ưu tiên còn những nơi khác gần như biến mất khỏi tầm nhìn của người dùng.

Điểm quan trọng là phần lớn các quyết định này không cần thay đổi bất kỳ văn bản pháp luật nào. Một địa phương vẫn giữ nguyên tên gọi, ranh giới và tư cách pháp lý, nhưng hình ảnh mà hàng triệu tác nhân AI sử dụng lại được hình thành từ những hệ thống nằm ngoài phạm vi quản lý trực tiếp của địa phương. Quyền lực vì thế dịch chuyển từ quyền công bố thông tin sang quyền tổ chức và diễn giải thông tin.

Đây là điểm đáng dừng lại lâu hơn. Quyền pháp lý vẫn nằm ở chính quyền, nhưng phiên bản mà máy đang đọc có thể do những nguồn khác tạo ra và theo thời gian chính phiên bản ấy trở thành sự thật vận hành: cái mà hàng triệu tác nhân dùng làm

căn cứ, bất kể văn bản chính thức viết gì. Khoảng cách giữa sự thật pháp lý và sự thật vận hành là khoảng cách mà chương này gọi là mất chủ quyền biểu diễn.

Nói cách khác, quyền định nghĩa không biến mất mà dịch chuyển từ người sở hữu thông tin sang người kiểm soát kiến trúc biểu diễn và phân phối thông tin. Đây có thể là một trong những thay đổi quyền lực đáng chú ý nhất mà các địa phương phải đối diện trong thập niên tới.

### 9.1.4. Từ thương hiệu sang quyền biểu diễn

Trong nhiều thập niên, chiến lược của các địa phương tập trung vào việc xây dựng thương hiệu. Họ thiết kế logo, khẩu hiệu, chiến dịch truyền thông, bộ nhận diện và những câu chuyện nhằm tạo ra một hình ảnh khác biệt trong tâm trí con người. Thành công của thương hiệu được đo bằng mức độ nhận biết, thiện cảm và khả năng thu hút du khách, nhà đầu tư hay nhân tài.

Nền kinh tế tác nhân không làm thương hiệu trở nên kém quan trọng, nhưng nó đặt ra một tầng vấn đề mới nằm bên dưới thương hiệu: quyền biểu diễn (representation). Trước khi một địa phương có thể tạo ra cảm xúc, nó phải được máy biểu diễn đúng. Nếu tầng biểu diễn sai, mọi nỗ lực xây dựng thương hiệu ở tầng trên đều bị suy yếu.

Khái niệm biểu diễn ở đây không chỉ là cách địa phương được mô tả bằng ngôn từ. Nó bao gồm toàn bộ cách địa phương được mã hóa trong thế giới số: thực thể nào được tạo ra, các thuộc tính nào được gắn vào thực thể đó, những quan hệ nào được thiết lập và nguồn dữ liệu nào được coi là có thẩm quyền. Đó là lớp hạ tầng vô hình quyết định máy sẽ hiểu địa phương theo cách nào.

Điều này cũng làm thay đổi vai trò của tự sự. Trong nền kinh tế trải nghiệm, câu chuyện là công cụ trung tâm của thương hiệu. Câu chuyện tạo ra cảm xúc, ý nghĩa và bản sắc. Nhưng với AI, một câu chuyện có khả năng tham gia sâu hơn vào quá trình suy luận khi các nhân vật, địa điểm, sự kiện và quan hệ trong đó có thể được nhận diện thành thực thể và thuộc tính. Văn bản phi cấu trúc vẫn có giá trị với mô hình ngôn ngữ, chỉ là độ ổn định và khả năng đối chiếu thấp hơn. Một truyền thuyết, một lễ hội hay một bản sắc văn hóa nếu chỉ tồn tại dưới dạng văn bản sẽ rất khó trở thành một phần của quá trình ra quyết định của tác nhân AI.

Có thể nói gọn sự phân biệt này như sau: thương hiệu là điều địa phương muốn được cảm nhận, còn biểu diễn là cấu trúc mà máy dùng để nhận diện. Vì vậy, một trong những thay đổi lớn nhất mà cuốn sách đề xuất là chuyển trọng tâm từ tư duy xây dựng thương hiệu sang tư duy quản trị quyền biểu diễn. Thương hiệu vẫn là mục tiêu cuối cùng, nhưng trước khi địa phương có thể tác động đến cách mình được cảm nhận, họ phải làm chủ cách mình được máy nhận diện. Trong thời đại tác nhân AI, quyền biểu diễn trở thành lớp nền của thương hiệu, giống như bản đồ là điều kiện để một hành trình có thể bắt đầu.

### 9.1.5. Chủ quyền bắt đầu từ quyền sửa tên mình

Mọi hình thức chủ quyền đều bắt đầu từ quyền được xác định danh tính của chính mình. Một quốc gia có quyền quyết định tên gọi của mình trên trường quốc tế. Một doanh nghiệp có quyền xác lập thương hiệu của mình trên thị trường. Một cá nhân có quyền được nhận diện bằng đúng danh tính của mình. Đối với địa phương trong nền kinh tế tác nhân cũng vậy, biểu hiện đầu tiên của chủ quyền không phải là khả năng kiểm soát mọi dữ liệu về mình, mà là quyền được máy nhận diện đúng thực thể mà mình là.

Nghe có vẻ hiển nhiên, nhưng đây lại là điều chưa hề được bảo đảm trong thế giới số hiện nay. Một địa phương có thể đồng thời tồn tại dưới nhiều tên gọi khác nhau, nhiều định danh khác nhau và nhiều mô tả khác nhau trong các hệ thống AI. Có nơi vẫn được lưu bằng tên cũ trước khi sáp nhập. Có nơi bị tách thành nhiều thực thể do khác biệt về dấu tiếng Việt, cách phiên âm hay cách viết tắt. Có nơi bị gộp nhầm với một địa danh khác có tên gần giống. Khi điều đó xảy ra, địa phương không chỉ mất tính nhất quán về dữ liệu mà còn mất quyền kiểm soát đối với chính danh tính số của mình.

Đó là lý do khái niệm thực thể chuẩn trở nên đặc biệt quan trọng. Trong khoa học dữ liệu, một thực thể chuẩn (canonical entity) là bản ghi tham chiếu có thẩm quyền, được dùng để hợp nhất và đối chiếu mọi biến thể của cùng một thực thể. Mọi tên gọi khác, mọi cách viết khác hay mọi định danh lịch sử đều phải quy chiếu về thực thể chuẩn này. Đối với nền kinh tế tác nhân, thực thể chuẩn không chỉ là một giải pháp kỹ thuật mà là nền móng của chủ quyền. Chỉ khi tồn tại một thực thể chuẩn được cộng đồng và các hệ thống công nhận, địa phương mới có thể duy trì được sự nhất quán của hình ảnh số qua thời gian.

Từ thực thể chuẩn dẫn đến khái niệm danh tính số. Danh tính của một địa phương không chỉ là tên gọi trên bản đồ mà còn là toàn bộ tập hợp các thuộc tính và quan hệ tạo nên bản sắc số của nó: ranh giới hành chính, di sản, điểm tham quan, doanh nghiệp, lễ hội, dịch vụ, lịch sử và các mối liên kết với những thực thể khác. Nếu danh tính này bị thay đổi hoặc bị diễn giải sai, máy sẽ bắt đầu suy luận từ một nền tảng sai lệch. Khi đó, những quyết định tiếp theo cũng có nguy cơ sai lệch theo.

Tuy nhiên, một danh tính chỉ thực sự có giá trị khi có nguồn có thẩm quyền, một nguồn có đủ thẩm quyền để xác nhận và

hiệu chỉnh danh tính ấy. Trong thế giới vật lý, giấy khai sinh cần cơ quan hộ tịch xác nhận, bản đồ cần cơ quan đo đạc ban hành và văn bản pháp luật cần cơ quan nhà nước có thẩm quyền công bố. Trong thế giới số cũng vậy, phải tồn tại những nguồn có thẩm quyền mà các nền tảng và hệ thống AI có thể tin cậy khi cần xác minh hoặc sửa sai. Nếu không có một cơ chế xác lập thẩm quyền rõ ràng, mỗi hệ thống sẽ tiếp tục duy trì phiên bản riêng của địa phương và sự phân mảnh sẽ ngày càng lớn.

Đó là lý do chương này coi quyền sửa tên mình là điểm khởi đầu của toàn bộ khái niệm chủ quyền thuật toán. Khi một địa phương có thể xác lập thực thể chuẩn, duy trì danh tính nhất quán và xây dựng được nguồn dữ liệu có thẩm quyền để các hệ thống khác tham chiếu, họ mới thực sự bước vào vị thế chủ động trong nền kinh tế tác nhân. Ngược lại, nếu ngay cả cách máy nhận diện mình cũng do người khác quyết định, thì mọi chiến lược thương hiệu, truyền thông hay phát triển điểm đến đều phải xây dựng trên một nền móng không thuộc quyền kiểm soát của chính địa phương.

Đây chỉ là tối ưu hóa tìm kiếm được gọi bằng một cái tên khác. Tối ưu hóa tìm kiếm là kỹ thuật cải thiện thứ hạng của một trang trong một hệ thống. Vấn đề ở đây khác cả về đối tượng lẫn về câu hỏi. Đối tượng không phải một trang mà là một thực thể được rất nhiều hệ thống cùng lưu giữ, mỗi hệ thống một phiên bản. Câu hỏi không phải làm sao xếp cao hơn mà ai có thẩm quyền xác nhận thực thể ấy khi các phiên bản mâu thuẫn. Một kỹ thuật xếp hạng không trả lời được câu hỏi thẩm quyền và đó là lý do mục này dùng chữ quyền chứ không dùng chữ tối ưu. Ai xác lập được thực thể thì người ấy nắm điểm bắt đầu của mọi câu trả lời.

## 9.2. CHỦ QUYỀN DỮ LIỆU VÀ CHỦ QUYỀN THUẬT TOÁN LÀ HAI KHÁI NIỆM KHÁC NHAU

Trong nhiều năm, các địa phương xem dữ liệu là tài sản số quan trọng nhất của mình. Họ đầu tư xây dựng cơ sở dữ liệu về dân cư, di sản, điểm tham quan, lễ hội, doanh nghiệp, lưu trú và hạ tầng với kỳ vọng rằng ai sở hữu dữ liệu sẽ nắm quyền kiểm soát quá trình chuyển đổi số. Điều đó đúng trong giai đoạn dữ liệu chủ yếu được con người truy cập. Nhưng trong nền kinh tế tác nhân, khi phần lớn quyết định được tạo ra bởi AI, quyền lực không còn dừng ở việc ai sở hữu dữ liệu, mà chuyển sang câu hỏi quan trọng hơn: ai kiểm soát cách dữ liệu được đọc, diễn giải và sử dụng để ra quyết định.

Đó là sự khác biệt giữa chủ quyền dữ liệu (data sovereignty) và chủ quyền thuật toán (algorithmic sovereignty). Chủ quyền dữ liệu nói về quyền tạo lập, quản lý, lưu trữ và cập nhật dữ liệu. Chủ quyền thuật toán không phải quyền trực tiếp quyết định cách mô hình suy luận, vì quyền ấy nằm ở phía nền tảng. Nó là năng lực duy trì ảnh hưởng đối với những tín hiệu, nguồn tham chiếu và cơ chế hiệu chỉnh mà quá trình suy luận dựa vào. Một địa phương có thể hoàn toàn sở hữu dữ liệu của mình nhưng vẫn không kiểm soát được cách các nền tảng AI biểu diễn địa phương đó trước hàng triệu người dùng.

Điều này tạo ra một nghịch lý mới. Một bảo tàng có thể công bố đầy đủ thông tin trên trang web chính thức, một khu du lịch có thể cập nhật chính xác giờ mở cửa và giá vé trong hệ thống của mình, nhưng nếu các mô hình AI học từ những nguồn khác, hoặc xếp các nguồn đó có mức độ tin cậy cao hơn, thì phiên bản mà người dùng nhìn thấy vẫn có thể thiếu, cũ hoặc sai. Quyền sở hữu dữ liệu không tự động chuyển thành quyền kiểm soát kết quả mà thuật toán tạo ra.

Vì vậy, trong nền kinh tế tác nhân, chủ quyền không còn được đo bằng số lượng dữ liệu mà một địa phương nắm giữ, mà bằng khả năng khiến dữ liệu của mình trở thành nguồn tham chiếu có thẩm quyền trong hệ sinh thái số. Chủ quyền thuật toán tăng lên khi dữ liệu của địa phương ngày càng được các công cụ tìm kiếm, nền tảng, đồ thị tri thức và mô hình AI nhận diện như một nguồn tham chiếu đáng tin cậy.

Đó cũng là lý do chương này chuyển trọng tâm từ câu hỏi "Ai sở hữu dữ liệu?" sang một câu hỏi sâu hơn: "Ai kiểm soát quá trình hình thành tri thức của máy?" Trong nền kinh tế tác nhân, hai câu hỏi ấy không còn có cùng một đáp án.

Ảnh hưởng của một chủ thể trong hệ sinh thái số tăng theo mức độ gần gũi của chủ thể ấy với lớp diễn giải, không theo khối lượng dữ liệu mà chủ thể ấy nắm giữ. Một bên có thể không tạo ra bất kỳ dữ liệu gốc nào mà vẫn quyết định phần lớn kết quả. Sở hữu dữ liệu là điều kiện cần của ảnh hưởng, chưa phải điều kiện đủ.

Bốn nghĩa của chữ chủ quyền trong chương này. Chương này dùng chữ chủ quyền theo bốn nghĩa không thay thế được cho nhau và mỗi lần dùng đều có định ngữ đi kèm. Chủ quyền nhà nước là quyền lực công đối với lãnh thổ, dân cư và pháp luật, phần này không bàn tới. Chủ quyền dữ liệu là quyền tạo lập, lưu giữ, cập nhật và quyết định điều kiện truy cập dữ liệu. Chủ quyền thuật toán là khả năng tác động lên những đầu vào và tín hiệu thẩm quyền mà hệ thống suy luận sử dụng, không phải quyền sửa mô hình. Chủ quyền biểu diễn là quyền xác lập và duy trì phiên bản chuẩn của thực thể để các hệ thống khác tham chiếu. Ba nghĩa sau xếp thành một thứ tự: không có chủ quyền biểu diễn thì chủ quyền dữ liệu chỉ là quyền lưu giữ và không có chủ quyền dữ liệu thì chủ quyền thuật toán không có chỗ để tác động.

> ▪ ***Ba lớp từ dữ liệu đến hành động và điểm tập trung quyền lực:***
> • ***Lớp 3 – Hành động:*** *đặt chỗ, mua vé, thanh toán, phản hồi kết quả. Câu hỏi của lớp này: ai được phép thực thi thay người dùng.*
> • ***Lớp 2 – Diễn giải:*** *xếp hạng, suy luận, đề xuất, chọn nguồn để tin. Câu hỏi của lớp này: ai quyết định điều gì là quan trọng.*
> • ***Lớp 1 – Dữ liệu:*** *thực thể, thuộc tính, quan hệ, nguồn, dấu thời gian.*

Câu hỏi của lớp này: ai giữ và ai xác nhận. Địa phương thường chỉ hiện diện ở lớp 1. Quyền lực tập trung ở lớp 2. Giá trị kinh tế được ghi nhận ở lớp 3. Bảng 9.1 so sánh chủ quyền dữ liệu và chủ quyền thuật toán theo bảy tiêu chí.

**Bảng 9.1.** Chủ quyền dữ liệu và chủ quyền thuật toán so theo bảy tiêu chí

| Tiêu chí | Chủ quyền dữ liệu | Chủ quyền thuật toán |
|---|---|---|
| **Câu hỏi trung tâm** | Ai giữ dữ liệu và theo điều kiện nào | Ai quyết định dữ liệu nào được tin và được dùng |
| **Đối tượng** | Tập dữ liệu, nơi lưu trữ, quyền truy cập | Tín hiệu thẩm quyền, lược đồ, thứ tự ưu tiên |
| **Công cụ pháp lý** | Luật dữ liệu, hợp đồng, quy định truy cập | Tiêu chuẩn, nghĩa vụ minh bạch, quy trình khiếu nại |
| **Chủ thể nắm giữ chủ yếu** | Cơ quan quản lý và doanh nghiệp | Nền tảng và nhà phát triển mô hình |
| **Chỉ báo đạt được** | Dữ liệu đầy đủ, an toàn, cập nhật | Nguồn của địa phương được máy dẫn lại và ưu tiên khi mâu thuẫn |
| **Cách mất đi** | Dữ liệu bị rò rỉ, bị khóa vào nền tảng, không được duy trì | Nguồn khác trở thành nguồn được tin hơn, dù kém chính xác hơn |
| **Việc phải làm** | Chuẩn hóa, phân lớp, cập nhật, bảo vệ | Phát tín hiệu thẩm quyền, kiểm toán tri thức, hiệu chỉnh liên tục |

### 9.2.1. Có dữ liệu nhưng không có quyền suy luận

Một địa phương có thể lưu giữ đầy đủ dữ liệu về di tích, khách sạn, tuyến tham quan, lễ hội và doanh nghiệp. Tuy nhiên, địa phương không quyết định AI sẽ kết hợp những dữ liệu đó như thế nào, ưu tiên tín hiệu nào, loại bỏ tín hiệu nào hay cuối cùng sẽ đề xuất điểm đến nào cho người dùng. Những quyết định ấy nằm trong mô hình suy luận của thuật toán.

Đây là sự dịch chuyển quan trọng của quyền lực số. Nếu dữ liệu trả lời câu hỏi "có gì", thì suy luận trả lời câu hỏi "điều gì quan trọng". Hai hệ thống có thể cùng đọc một tập dữ liệu nhưng đưa ra hai kết luận hoàn toàn khác nhau vì sử dụng hai mô hình suy luận khác nhau. Chính quá trình lựa chọn, kết nối và xếp hạng ấy mới quyết định giá trị thực tế của dữ liệu.

Có thể hình dung dữ liệu như các quân cờ trên bàn cờ, còn suy luận là người chơi. Người sở hữu quân cờ chưa chắc là người quyết định ván đấu. Trong nền kinh tế tác nhân cũng vậy, nhiều địa phương nắm dữ liệu nhưng không kiểm soát được cách AI sử dụng dữ liệu để xây dựng tri thức, xếp hạng lựa chọn hay thực hiện giao dịch. Vì thế, quyền lực đang dịch chuyển từ dữ liệu thô sang suy luận, từ quyền lưu trữ sang quyền diễn giải.

### 9.2.2. Chủ quyền dữ liệu trả lời "ai giữ"

Khái niệm chủ quyền dữ liệu ra đời để trả lời một nhóm câu hỏi rất cụ thể: dữ liệu được lưu ở đâu, thuộc quyền sở hữu của ai, ai được quyền truy cập, ai có trách nhiệm cập nhật và hệ thống pháp luật nào điều chỉnh việc quản lý dữ liệu đó. Đây là những vấn đề về lưu trữ và sở hữu.

Ở cấp địa phương, chủ quyền dữ liệu thể hiện qua việc chính quyền quản lý cơ sở dữ liệu hành chính, doanh nghiệp quản lý dữ liệu khách hàng, còn các đơn vị vận hành quản lý dữ liệu dịch vụ của mình. Việc dữ liệu được lưu trữ an toàn, cập nhật đầy đủ và không bị sử dụng trái phép là điều kiện nền tảng của chuyển đổi số. Tuy nhiên, chủ quyền dữ liệu chỉ bảo đảm rằng địa phương giữ được dữ liệu, chứ không bảo đảm rằng địa phương giữ được ảnh hưởng đối với cách dữ liệu ấy được sử dụng. Khi dữ liệu được công khai hoặc được các nền tảng thu thập, các hệ thống AI có thể kết hợp chúng với hàng nghìn nguồn khác để tạo ra những kết luận vượt xa ý định ban đầu của người tạo dữ liệu.

Khoảng cách giữa câu hỏi ai giữ dữ liệu và câu hỏi ai quyết định ý nghĩa của dữ liệu chính là nơi xuất hiện khái niệm chủ quyền thuật toán, lớp quyền lực mới được phân tích ở các mục tiếp theo.

### 9.2.3. Chủ quyền thuật toán trả lời "ai quyết định"

Nếu chủ quyền dữ liệu quan tâm đến việc dữ liệu thuộc về ai, thì chủ quyền thuật toán quan tâm đến việc ai quyết định dữ liệu được sử dụng như thế nào. Đây là lớp quyền lực mới hình thành khi AI trở thành trung gian giữa thế giới số và quyết định của con người.

Mỗi ngày, các mô hình AI phải xử lý hàng tỷ tín hiệu khác nhau. Chúng không thể sử dụng tất cả. Vì vậy, chúng liên tục thực hiện ba hành động quan trọng: xếp hạng (ranking), suy luận (reasoning) và đề xuất (recommendation). Ba hành động này quyết định điều gì sẽ được nhìn thấy, điều gì bị bỏ qua và cuối cùng người dùng sẽ lựa chọn điều gì.

Khi một du khách hỏi AI "Nên ở đâu tại Đà Nẵng?", câu trả lời không chỉ phụ thuộc vào việc Đà Nẵng có bao nhiêu khách sạn hay bao nhiêu dữ liệu được công bố. Điều quan trọng hơn là thuật toán đánh giá nguồn nào đáng tin cậy hơn, hiểu sở thích của người dùng ra sao, cân nhắc khoảng cách, giá cả, đánh giá, khả năng đặt chỗ và hàng chục tín hiệu khác để tạo ra danh sách đề xuất. Chính thuật toán đang thực hiện vai trò của một "người biên tập" đối với toàn bộ thế giới số.

Vì vậy chủ quyền thuật toán cần được định nghĩa hẹp và dè dặt, nếu không, nó sẽ hứa nhiều hơn mức một địa phương làm được. Nó không phải quyền sửa mô hình, cũng không phải quyền tham gia vào suy luận của một hệ thống tư nhân. Nó là năng lực gồm năm khả năng thực tế: xác lập nguồn chuẩn, phát tín hiệu có thẩm quyền, quan sát và kiểm toán kết quả biểu diễn, yêu cầu hiệu chỉnh và giảm mức phụ thuộc vào một nền tảng duy nhất.

Đó là lý do trong nền kinh tế tác nhân, quyền lực không chỉ nằm ở việc có dữ liệu, mà nằm ở việc tham gia vào quá trình quyết định của thuật toán. Chủ thể thiết kế hoặc vận hành lớp xếp hạng, suy luận và đề xuất đang nắm một phần đáng kể quyền phân phối sự chú ý, dòng khách và giá trị kinh tế.

### 9.2.4. Quyền lực nằm ở lớp diễn giải

Dữ liệu không tự quyết định giá trị mà nó sẽ tạo ra. Giá trị xuất hiện ở lớp nằm giữa dữ liệu và quyết định, lớp diễn giải hay lớp suy luận (inference layer). Đây là nơi các mô hình AI biến hàng triệu mẩu thông tin rời rạc thành tri thức có thể hành động.

Chính lớp trung gian này tạo ra quyền lực mới. Hai nền tảng có thể sử dụng gần như cùng một tập dữ liệu nhưng đưa ra các

khuyến nghị hoàn toàn khác nhau vì cách chúng diễn giải và suy luận khác nhau. Nói cách khác, sự khác biệt cạnh tranh ngày càng ít nằm ở việc ai có nhiều dữ liệu hơn, mà nằm ở việc ai diễn giải dữ liệu tốt hơn.

Điều này giải thích vì sao các công ty AI, công cụ tìm kiếm và nền tảng số trở thành những tác nhân có ảnh hưởng lớn đến kinh tế địa phương dù họ không trực tiếp sở hữu các tài sản du lịch. Họ sở hữu lớp diễn giải, nơi quyết định thông tin nào được biến thành tri thức, tri thức nào được biến thành khuyến nghị và khuyến nghị nào được chuyển thành giao dịch.

Với địa phương, thay đổi này chạm tới cách phân bổ nguồn lực. Mục tiêu không còn chỉ là xây dựng cơ sở dữ liệu đầy đủ mà là thiết kế dữ liệu theo cách giúp máy diễn giải đúng ngữ nghĩa và nhận diện đúng quan hệ giữa các thực thể. Đây là một trong những dịch chuyển sâu nhất của quyền lực số trong thập niên AI.

### 9.2.5. Không thể bảo vệ thương hiệu chỉ bằng dữ liệu mở

Trong nhiều năm, dữ liệu mở được xem là một trụ cột của chuyển đổi số, vì nó tăng minh bạch, thúc đẩy đổi mới và giúp các nền tảng hiểu địa phương đầy đủ hơn. Tuy nhiên trong nền kinh tế tác nhân, dữ liệu mở chỉ là điều kiện cần, không còn là điều kiện đủ để bảo vệ thương hiệu số của một địa phương.

Lý do nằm ở chỗ dữ liệu mở chỉ giải quyết vấn đề khả năng tiếp cận, chứ không giải quyết vấn đề quản trị việc sử dụng dữ liệu. Một khi dữ liệu được đưa vào hệ sinh thái AI, nó sẽ được kết hợp với vô số nguồn khác, được tóm tắt, diễn giải, xếp hạng và tái sử dụng trong những ngữ cảnh hoàn toàn mới. Quá trình này

có thể tạo ra những giá trị rất lớn, nhưng cũng có thể khiến hình ảnh của địa phương dần tách khỏi ý nghĩa ban đầu mà chính địa phương mong muốn truyền tải.

Một địa phương có thể công khai toàn bộ dữ liệu về di sản văn hóa, nhưng nếu các nền tảng chỉ học và khuếch đại những đánh giá tiêu cực, những thông tin lỗi thời hoặc những thuộc tính dễ tối ưu như giá rẻ và khoảng cách gần, thì bản sắc văn hóa phong phú vẫn có thể bị nén thành vài dòng mô tả đơn giản. Vấn đề không nằm ở việc dữ liệu có được mở hay không, mà ở việc ai đang quyết định cách dữ liệu được kết hợp và trình bày.

Chính vì vậy, dữ liệu mở phải đi cùng quản trị. Quản trị trong nền kinh tế tác nhân không còn chỉ là quản trị cơ sở dữ liệu, mà là quản trị toàn bộ vòng đời của tri thức số: chuẩn hóa thực thể, xác định nguồn dữ liệu có thẩm quyền, duy trì tính cập nhật, theo dõi cách AI biểu diễn địa phương, phát hiện sai lệch và xây dựng cơ chế hiệu chỉnh liên tục. Nói cách khác, địa phương không chỉ cần "mở dữ liệu", mà còn phải "quản trị khả năng được máy hiểu đúng".

Không một địa phương nào can thiệp được vào thuật toán của nền tảng, nên chủ quyền thuật toán chỉ là một cách nói cho sang. Phản bác này đúng ở tiền đề và sai ở suy luận. Không địa phương nào sửa được trọng số của một mô hình. Nhưng chủ quyền thuật toán ở đây không được định nghĩa là quyền sửa mô hình mà là quyền tác động lên những đầu vào mà mô hình coi là đáng tin. Toàn bộ biện pháp thực tế nằm trong khoảng cách giữa hai điều đó: nguồn có thẩm quyền, dấu thời gian, khả năng đối chiếu, kênh hiệu chỉnh. Nếu định nghĩa chủ quyền là kiểm soát mô hình thì kết luận bi quan là tất yếu, nhưng định nghĩa ấy cũng không dùng được cho việc gì. Dữ liệu là nguyên liệu. Diễn giải là quyền lực.

## 9.3. BA MẶT CỦA CHỦ QUYỀN

Các mục trước đã chỉ ra ba lớp mà quyền lực số vận hành qua: dữ liệu, diễn giải và hành động. Một địa phương có thể sở hữu rất nhiều dữ liệu mà không quyết định được cách dữ liệu ấy được diễn giải. Ngược lại, một nền tảng có thể không tạo ra dữ liệu gốc nào mà vẫn chi phối quá trình suy luận, xếp hạng và phân phối. Nhưng ba lớp ấy chưa phải ba mặt của chủ quyền. Chúng mô tả nơi quyền lực nằm, còn chủ quyền phải được đo bằng những năng lực giúp địa phương không bị lệ thuộc trong các lớp ấy.

Ba năng lực ấy có thể kiểm chứng bằng ba câu hỏi. Địa phương có xác lập được chính xác mình là ai trong thế giới của máy hay không, đây là chủ quyền biểu diễn. Có mang được danh tính và dữ liệu ấy sang hệ thống khác mà không phải làm lại từ đầu hay không, đây là chủ quyền khả chuyển. Có kết nối được với các hệ thống khác mà không bị khóa vào một nhà cung cấp duy nhất hay không, đây là chủ quyền liên thông. Nói gọn lại, ba lớp dữ liệu, diễn giải và hành động mô tả nơi quyền lực vận hành, còn ba mặt biểu diễn, khả chuyển và khả liên thông mô tả năng lực mà địa phương cần có để không bị lệ thuộc trong những lớp ấy.

Ba mặt này không cộng vào nhau mà nhân với nhau. Một danh tính số rõ ràng mà không mang đi được thì chỉ có giá trị trong một hệ sinh thái. Dữ liệu mang đi được nhưng mỗi nơi hiểu một cách thì tính nhất quán biến mất. Biểu diễn và khả chuyển đều tốt mà các hệ thống không nói cùng ngôn ngữ số thì chuỗi hành động vẫn đứt. Vì vậy đầu tư lệch vào một mặt thường không tạo ra chuyển biến nào đo được và đây là khung tư duy cho các mục tiếp theo.

Chủ quyền số của một địa phương gần với một phép nhân hơn một phép cộng: nó là kết quả kết hợp của chủ quyền biểu

diễn, chủ quyền khả chuyển và chủ quyền liên thông. Khi một mặt tiến về không, hai mặt còn lại gần như không cải thiện được năng lực thực tế, nên đầu tư lệch vào một mặt thường không tạo ra chuyển biến.

### 9.3.1. Chủ quyền biểu diễn

Một địa phương tồn tại trong thực tế và đồng thời tồn tại qua nhiều cách biểu diễn số. Nếu lớp biểu diễn này sai lệch, không đầy đủ hoặc bị phân mảnh, thì mọi hoạt động suy luận của tác nhân cũng sẽ dựa trên một hình ảnh méo mó. Vì vậy, trước cả chủ quyền dữ liệu hay chủ quyền thuật toán, địa phương cần có chủ quyền biểu diễn (representation sovereignty): quyền xác lập và duy trì bản sắc số chính thức của mình.

Chủ quyền biểu diễn trả lời câu hỏi rất cơ bản: máy đang nghĩ địa phương này là ai. Một điểm đến có tên chính thức là gì, ranh giới ở đâu, thuộc đơn vị hành chính nào, có những điểm tham quan nào, sự kiện nào là chính thức, doanh nghiệp nào đang hoạt động và các thực thể ấy liên hệ với nhau ra sao. Đây không chỉ là thông tin mô tả mà là nền tảng để các hệ thống AI xây dựng mô hình nhận thức về địa phương.

Vì vậy, chủ quyền biểu diễn không phải là quyền kiểm soát mọi thông tin về mình mà là quyền xác lập phiên bản chuẩn để những hệ thống có nhu cầu có thể tham chiếu. Đó là nền móng của niềm tin số. Khi địa phương làm chủ được cách mình được biểu diễn, mọi lớp dữ liệu, suy luận và giao dịch phía trên mới có điểm tựa thống nhất.

### 9.3.2. Chủ quyền khả chuyển

Nếu chủ quyền biểu diễn quyết định máy hiểu địa phương như thế nào, thì chủ quyền khả chuyển (portability sovereignty) quyết định địa phương có thể đi cùng máy đến đâu. Nói đơn giản, đó là việc đem được danh tính và dữ liệu của mình sang một hệ thống khác mà không phải làm lại từ đầu. Một thực thể số chỉ thực sự có chủ quyền khi danh tính, dữ liệu và năng lực của nó có thể được sử dụng nhất quán trên nhiều nền tảng, nhiều tác nhân và nhiều hệ thống khác nhau.

Trong nhiều năm, chiến lược số của không ít địa phương gắn chặt với một trang web, một ứng dụng hoặc một nền tảng duy nhất. Điều đó tạo ra sự phụ thuộc: khi người dùng chuyển sang một hệ sinh thái AI khác, phần lớn tài sản số của địa phương phải được xây dựng lại từ đầu. Trong nền kinh tế tác nhân, đây là một rủi ro lớn bởi tác nhân không trung thành với bất kỳ giao diện nào, chúng sẽ liên tục di chuyển giữa các mô hình AI, công cụ tìm kiếm, nền tảng đặt dịch vụ và hệ thống giao dịch khác nhau.

Vì vậy, khả chuyển trước máy trở thành một năng lực cốt lõi. Dữ liệu cần tuân theo các tiêu chuẩn mở, thực thể cần có mã định danh ổn định, quan hệ cần được mô tả theo những mô hình mà nhiều hệ thống đều có thể đọc, còn các dịch vụ cần được cung cấp thông qua các giao diện có khả năng tích hợp rộng rãi. Mục tiêu không phải là gắn chặt địa phương vào một nền tảng, mà là để địa phương có thể xuất hiện nhất quán ở mọi nơi mà tác nhân hoạt động.

Chủ quyền khả chuyển vì thế không đồng nghĩa với việc sao chép dữ liệu sang nhiều nơi. Nó là khả năng giữ nguyên danh tính, ngữ nghĩa và khả năng hành động khi di chuyển giữa các hệ sinh thái số khác nhau. Một địa phương có chủ quyền khả

chuyển sẽ không bị "mắc kẹt" trong bất kỳ nền tảng nào, cũng không phải xây dựng lại từ đầu mỗi khi công nghệ thay đổi. Chính khả năng này giúp tài sản số tích lũy được giá trị lâu dài thay vì liên tục bị phân mảnh theo các chu kỳ công nghệ mới.

### 9.3.3. Chủ quyền liên thông

Nếu chủ quyền biểu diễn giúp địa phương có một danh tính thống nhất, còn chủ quyền khả chuyển giúp danh tính ấy có thể đi qua nhiều hệ thống, thì chủ quyền liên thông quyết định các hệ thống ấy có thể thực sự hợp tác với nhau hay không. Nói đơn giản, đó là việc các hệ thống khác nhau hiểu được nhau và phối hợp được với nhau. Đây là năng lực để dữ liệu, dịch vụ và tác nhân của nhiều tổ chức cùng vận hành trên một nền tảng ngữ nghĩa chung mà không cần phụ thuộc vào một nhà cung cấp duy nhất.

Trong nền kinh tế tác nhân, không một địa phương nào có thể tự xây dựng toàn bộ hệ sinh thái số. Một hành trình của du khách có thể đồng thời đi qua cơ quan quản lý, doanh nghiệp lưu trú, đơn vị vận tải, cổng thanh toán, bảo hiểm, bản đồ số và nhiều tác nhân AI khác nhau. Nếu mỗi bên sử dụng một chuẩn dữ liệu riêng, một mô hình thực thể riêng hoặc một giao diện tích hợp riêng, toàn bộ chuỗi hành động sẽ bị chia cắt. Máy phải liên tục chuyển đổi định dạng, suy đoán ngữ nghĩa hoặc nhập lại thông tin, làm giảm tốc độ, độ chính xác và mức độ tin cậy.

Vì vậy, chủ quyền liên thông không có nghĩa là mọi hệ thống phải giống nhau. Điều quan trọng là chúng nói cùng một ngôn ngữ số. Khả liên thông (interoperability) cho phép các hệ thống khác nhau trao đổi dữ liệu và phối hợp hành động mà vẫn giữ được tính độc lập. Điều này chỉ có thể đạt được khi địa phương ưu tiên tiêu chuẩn mở thay vì các định dạng đóng hoặc các giải pháp khóa người dùng vào một nền tảng duy nhất.

Do đó, chủ quyền liên thông không đối lập với hợp tác. Trái lại, đó là điều kiện để hợp tác mà vẫn giữ được quyền tự chủ. Địa phương không cần sở hữu mọi hệ thống, nhưng cần bảo đảm rằng bất kỳ hệ thống nào muốn kết nối cũng phải có thể đọc, hiểu và giao tiếp với hạ tầng dữ liệu của mình thông qua các tiêu chuẩn chung.

### 9.3.4. Thiếu một mặt, hai mặt còn lại suy yếu

Ba mặt của chủ quyền không tồn tại độc lập mà tạo thành một cấu trúc phụ thuộc lẫn nhau. Chỉ cần một mắt xích bị đứt, toàn bộ năng lực số của địa phương sẽ suy giảm.

Một địa phương có thể xây dựng được danh tính số rất rõ ràng, nhưng nếu dữ liệu không thể di chuyển giữa các nền tảng thì giá trị của danh tính ấy chỉ bị giới hạn trong một hệ sinh thái. Ngược lại, dữ liệu có thể di chuyển dễ dàng nhưng nếu mỗi nền tảng lại hiểu thực thể theo một cách khác nhau thì sự nhất quán sẽ biến mất. Tương tự, dù biểu diễn và khả chuyển đều tốt, nếu các hệ thống không liên thông thì chuỗi hành động vẫn bị ngắt quãng, tác nhân có thể tìm thấy địa phương nhưng không thể hoàn thành giao dịch.

Đó chính là hiện tượng hệ sinh thái đứt gãy. Sự đứt gãy không nhất thiết xuất phát từ việc thiếu dữ liệu hay thiếu công nghệ, mà từ việc các thành phần không thể phối hợp với nhau. Người dùng nhìn thấy một trải nghiệm liền mạch, nhưng phía sau là hàng chục hệ thống phải trao đổi thông tin chính xác và liên tục. Chỉ cần một điểm không tương thích, toàn bộ chuỗi giá trị có thể bị gián đoạn.

Chủ quyền trong nền kinh tế tác nhân không thể được xây dựng theo từng dự án riêng lẻ. Đầu tư vào dữ liệu mà bỏ qua tiêu chuẩn liên thông, hay xây dựng API mà không quản lý danh tính thực thể, đều chỉ tạo ra những năng lực cục bộ. Giá trị chỉ xuất hiện khi ba mặt của chủ quyền được phát triển đồng thời và hỗ trợ lẫn nhau.

Nói cách khác, chủ quyền biểu diễn giúp máy biết địa phương là ai, chủ quyền khả chuyển giúp địa phương hiện diện ở mọi nơi máy hoạt động, còn chủ quyền liên thông giúp toàn bộ hệ sinh thái có thể cùng hành động. Thiếu bất kỳ mặt nào, hai mặt còn lại cũng không thể phát huy đầy đủ giá trị. Đây là lý do chủ quyền trong nền kinh tế tác nhân phải được nhìn như một kiến trúc hoàn chỉnh, chứ không phải một tập hợp các sáng kiến số rời rạc. Bảng 9.2 tổng hợp ba mặt ấy cùng câu hỏi kiểm tra và chỉ báo suy giảm.

**Bảng 9.2.** Ba mặt của chủ quyền: câu hỏi kiểm tra, chỉ báo suy giảm và biện pháp

| Mặt | Câu hỏi kiểm tra | Chỉ báo suy giảm | Biện pháp then chốt |
|---|---|---|---|
| **Biểu diễn** | Các hệ thống có đang đọc cùng một thực thể hay không | Số bản trùng lặp tăng, tên gọi khác nhau giữa các nền tảng, quan hệ bị gắn nhầm | Thực thể chuẩn có mã định danh ổn định và nguồn xác thực công khai |
| **Khả chuyển** | Địa phương có lấy lại được tài sản số của mình hay không | Dữ liệu chỉ tồn tại trong một nền tảng, xuất ra không đọc được bằng máy | Điều khoản khả chuyển trong hợp đồng và định dạng mở |
| **Khả liên thông** | Các hệ thống có nói cùng một ngôn ngữ số hay không | Mỗi bên một lược đồ, phải nhập lại thông tin, chuỗi hành động bị ngắt | Tiêu chuẩn chung do cơ quan có thẩm quyền ban hành và đưa thành yêu cầu trong mua sắm công |

### 9.3.5. Chủ quyền là khả năng tồn tại nhất quán ở mọi hệ thống

Cuối cùng, có thể quy toàn bộ khái niệm chủ quyền trong nền kinh tế tác nhân về một tiêu chuẩn rất đơn giản: một địa phương có còn là chính mình khi đi qua mọi hệ thống số hay không. Nếu tên gọi, danh tính, quan hệ, dữ liệu và khả năng giao dịch đều được giữ nguyên dù xuất hiện trên công cụ tìm kiếm, bản đồ số, OTA, mạng xã hội, mô hình AI hay các tác nhân tự động khác nhau, thì địa phương đang có chủ quyền. Nếu mỗi hệ thống lại tạo ra một phiên bản khác nhau của cùng một thực thể, chủ quyền ấy đã bị suy giảm.

Điều này cho thấy chủ quyền không nằm ở việc sở hữu một trang web, một cơ sở dữ liệu hay một ứng dụng. Chủ quyền nằm ở tính nhất quán, khả năng duy trì sự nhất quán của thực thể trên toàn bộ hệ sinh thái số. Một địa phương phải có cùng một danh tính, cùng một ngữ nghĩa và cùng một mức độ tin cậy ở bất kỳ nơi nào máy tiếp cận. Chỉ khi đó, mọi suy luận của AI mới tích lũy trên cùng một nền tảng tri thức thay vì bị chia cắt thành nhiều phiên bản rời rạc.

Tuy nhiên, nhất quán thôi chưa đủ. Chủ quyền còn đòi hỏi tính bền vững, khả năng tồn tại bền vững trước sự thay đổi của công nghệ. Trong vài thập kỷ qua, internet đã liên tục chuyển từ trang web sang mạng xã hội, từ ứng dụng di động sang nền tảng đám mây và hiện nay là các tác nhân AI. Trong tương lai, giao diện và mô hình tương tác sẽ còn tiếp tục thay đổi. Nếu mỗi chu kỳ công nghệ đều buộc địa phương phải xây dựng lại toàn bộ danh tính số của mình, thì đó không phải là chủ quyền mà chỉ là sự thích nghi tạm thời.

Đó cũng là ý nghĩa sâu xa của chủ quyền trong nền kinh tế tác nhân. Nó không được đo bằng quyền kiểm soát một lãnh thổ số mà bằng khả năng để thực thể của mình vẫn được nhận diện đúng qua mọi thế hệ công nghệ, với cùng một danh tính, cùng một ngữ nghĩa và cùng một năng lực giao dịch.

Ba mặt của chủ quyền chỉ là một cách chia lại cùng một việc quản trị dữ liệu. Ba mặt đáng được tách ra vì chúng hỏng theo ba cách khác nhau và do ba chủ thể khác nhau sửa. Biểu diễn hỏng khi thực thể bị phân mảnh và người sửa là đơn vị giữ nguồn chuẩn. Khả chuyển hỏng khi tài sản số bị khóa vào một nền tảng và người sửa là đơn vị mua sắm công nghệ. Khả liên thông hỏng khi các bên dùng lược đồ riêng và người sửa là đơn vị ban hành tiêu chuẩn. Gộp ba mặt lại thì mất khả năng chẩn đoán, đó là lý do Bảng 9.2 tách chỉ báo suy giảm cho từng mặt. Chủ quyền số gần với một phép nhân hơn một phép cộng.

## 9.4. QUYỀN MỞ VÀ DỮ LIỆU NHƯ HẠ TẦNG CÔNG

Một trong những ngộ nhận phổ biến của kỷ nguyên AI là càng nhiều dữ liệu được công khai thì địa phương càng có lợi. Thực tế không đơn giản như vậy. Nếu các mục trước bàn về chủ quyền như quyền duy trì cách máy biểu diễn, suy luận và giao dịch, thì mục này đặt ra một giới hạn quan trọng của chủ quyền: không phải mọi dữ liệu đều nên được tối đa hóa. Một nền kinh tế tác nhân lành mạnh không chỉ cần dữ liệu mở, mà còn cần biết dữ liệu nào phải được mở, dữ liệu nào chỉ nên chia sẻ có điều kiện và dữ liệu nào cần được giữ lại như một tài sản chiến lược.

Điểm đến không chỉ gồm các tọa độ, giờ mở cửa hay giá vé. Giá trị của địa phương còn nằm trong những tri thức ngầm, những tập quán văn hóa, những không gian linh thiêng, những nguồn tài nguyên mong manh và cả những trải nghiệm mà chính sự không hoàn toàn minh bạch mới tạo nên sức hấp dẫn. Khi mọi thứ đều bị số hóa và tối ưu hóa cho máy, địa phương có thể vô tình đánh mất khả năng kiểm soát nhịp độ khai thác, làm suy giảm bản sắc hoặc khiến những tài sản dễ tổn thương trở thành đối tượng bị khai thác quá mức.

Điều đó không có nghĩa là quay lưng với dữ liệu mở. Ngược lại, dữ liệu công là một dạng hạ tầng thiết yếu của nền kinh tế số, giống như đường sá, điện hay viễn thông. Những dữ liệu nền tảng về thực thể, vị trí, dịch vụ công, giao thông hay thông tin du lịch cần được duy trì như một hạ tầng công cộng để mọi doanh nghiệp và tác nhân AI đều có thể xây dựng dịch vụ mới trên đó. Nhưng cũng giống như hạ tầng vật chất, hạ tầng dữ liệu phải đi kèm với các nguyên tắc quản trị: phân tầng quyền truy cập, bảo vệ lợi ích cộng đồng và cân bằng giữa đổi mới với bảo tồn.

Ngưỡng khả đọc. Giá trị của khả đọc không tăng đơn điệu theo lượng dữ liệu được mở. Với mỗi loại tài sản của địa phương tồn tại một ngưỡng mà vượt qua đó, thông tin bổ sung chuyển từ lợi thế cạnh tranh thành rủi ro khai thác. Quản trị dữ liệu là công việc xác định ngưỡng ấy cho từng lớp, không phải công việc mở tối đa mọi lớp.

Chủ quyền dữ liệu bản địa và bộ nguyên tắc CARE. Quyền mờ không phải một khái niệm được nghĩ ra cho nền kinh tế tác nhân. Trong hai thập niên qua, các cộng đồng bản địa ở Australia, New Zealand, Canada và Hoa Kỳ đã xây dựng một hệ thống lập luận về chủ quyền dữ liệu bản địa, với lập luận trung tâm là dữ liệu về một cộng đồng thuộc về cộng đồng ấy, kể cả khi do bên ngoài thu thập. Bộ nguyên tắc CARE, công bố trên Data Science Journal tháng 11 năm 2020, nêu bốn yêu cầu: lợi ích tập thể, quyền kiểm soát của cộng đồng, trách nhiệm của bên sử dụng dữ liệu và ràng buộc đạo đức. Nhóm tác giả nêu rõ CARE được đề xuất để đi cùng chứ không thay thế nhóm nguyên tắc FAIR về dữ liệu khoa học, vốn nhấn vào khả năng tìm thấy, truy cập, liên thông và tái sử dụng, với khẩu hiệu vừa FAIR vừa CARE. Một bài tiếp theo năm 2021 trên Scientific Data bàn cách vận hành đồng thời hai bộ nguyên tắc.[1] Cặp CARE và FAIR là một cách diễn đạt khác của chính căng thẳng mà mục này bàn: khả đọc là một giá trị, nhưng không phải giá trị duy nhất.

1. Stephanie Russo Carroll và cộng sự, "The CARE Principles for Indigenous Data Governance", Data Science Journal, tập 19, bài 43, ngày 4 tháng 11 năm 2020. Bài tiếp theo là Stephanie Russo Carroll, Edit Herczog, Maui Hudson, Keith Russell và Shelley Stall, "Operationalizing the CARE and FAIR Principles for Indigenous data futures", Scientific Data, tập 8, bài 108, năm 2021. Bộ nguyên tắc FAIR theo Mark D. Wilkinson và cộng sự, "The FAIR Guiding Principles for scientific data management and stewardship", Scientific Data, 2016.

### 9.4.1. Không phải mọi thứ đều cần khả đọc

Một trong những cám dỗ lớn của kỷ nguyên AI là tin rằng mọi thứ đều nên được số hóa, cấu trúc hóa và mở cho máy đọc. Tuy nhiên, một địa phương không chỉ là tập hợp của dữ liệu. Nó còn là tập hợp của những giới hạn. Có những giá trị càng được tối ưu hóa cho máy thì càng mất đi ý nghĩa đối với con người.

Những không gian linh thiêng là ví dụ rõ nhất. Một ngôi đền, một khu rừng cấm, một nghĩa trang cổ hay một nghi lễ tín ngưỡng không chỉ có tọa độ địa lý. Chúng còn có quy tắc tiếp cận, ngữ cảnh văn hóa và những giới hạn được cộng đồng gìn giữ qua nhiều thế hệ. Nếu AI chỉ nhìn thấy đây là một "điểm tham quan hấp dẫn" để tối đa hóa lượng khách, chính thuật toán có thể trở thành tác nhân làm xói mòn giá trị mà địa phương muốn bảo vệ.

Tương tự, không phải mọi dữ liệu đều nên được công khai. Thông tin về các loài quý hiếm, vị trí khảo cổ chưa khai quật, nguồn dược liệu tự nhiên, hệ sinh thái mong manh hay các tuyến đường truyền thống của cộng đồng bản địa đều có thể trở thành mục tiêu khai thác nếu bị phổ biến không kiểm soát. Trong trường hợp này, khả đọc quá cao lại tạo ra rủi ro lớn hơn lợi ích.

Ngay cả dữ liệu thương mại cũng cần có ranh giới. Tồn kho, doanh thu, chiến lược giá hay hành vi khách hàng là những tài sản cạnh tranh của doanh nghiệp địa phương. Nếu mọi dữ liệu đều trở thành dữ liệu công, địa phương có thể thúc đẩy đổi mới trong ngắn hạn nhưng lại làm suy yếu động lực đầu tư của chính hệ sinh thái kinh doanh.

Vì vậy, mục tiêu không phải là làm cho mọi thứ trở nên khả đọc, mà là làm cho đúng những gì cần thiết trở nên khả đọc. Một

hạ tầng dữ liệu trưởng thành luôn biết phân loại dữ liệu thành dữ liệu công, dữ liệu chia sẻ có điều kiện và dữ liệu cần được bảo vệ. Chủ quyền dữ liệu không chỉ là quyền mở dữ liệu mà còn là quyền quyết định giới hạn của sự minh bạch.

### 9.4.2. Quyền được giữ mờ

Trong nhiều thập niên, các chính sách dữ liệu tập trung vào quyền được biết (right to know), quyền tiếp cận thông tin và dữ liệu mở. Nhưng trong nền kinh tế tác nhân, cần bổ sung một quyền mới: quyền được giữ mờ. Đó không phải là quyền che giấu thông tin vì thiếu minh bạch, mà là quyền duy trì những khoảng không mà máy không nên tự động khai thác hoặc diễn giải.

"Mờ" ở đây không đồng nghĩa với "bí mật". Một lễ hội vẫn có thể được quảng bá rộng rãi nhưng không nhất thiết phải công bố toàn bộ cấu trúc vận hành nội bộ. Một làng nghề có thể giới thiệu sản phẩm mà không cần công khai toàn bộ bí quyết truyền thống. Một khu bảo tồn có thể cho phép tham quan nhưng không nhất thiết phải cung cấp vị trí chính xác của mọi loài quý hiếm cho mọi hệ thống AI.

Cần một hàng rào rõ ràng để khái niệm này không bị dùng sai. Quyền mờ không áp dụng đối với những thông tin mà pháp luật yêu cầu phải công khai, những dữ liệu liên quan đến an toàn công cộng, hay những thông tin cần thiết để bảo vệ quyền lợi người tiêu dùng. Nó chỉ áp dụng đối với phần tri thức mà việc phổ biến vô điều kiện có thể làm tổn hại đến cộng đồng, hệ sinh thái, không gian thiêng hoặc một lợi thế cạnh tranh chính đáng.

Quyền được giữ mờ còn là một cơ chế bảo vệ sự đa dạng văn hóa. Khi thuật toán liên tục chuẩn hóa dữ liệu theo cùng một

khuôn mẫu, những yếu tố khó mô tả, khó phân loại hoặc mang tính biểu tượng dễ bị xem là "nhiễu" và dần biến mất khỏi không gian số. Nhưng chính những phần không hoàn toàn chuẩn hóa ấy lại thường là linh hồn của một địa phương.

Ví dụ rõ nhất về cái van không nằm ở một thành phố di sản mà ở các cộng đồng bản địa. Nhiều nhóm dân tộc thiểu số ở vùng cao có những tri thức mà việc công bố sẽ làm mất chính giá trị của nó: vị trí một khu rừng thiêng, lịch một nghi lễ chỉ dành cho người trong cộng đồng, công thức một bài thuốc gia truyền. Đưa những thứ ấy vào lược đồ không phải là minh bạch mà là chuyển giao. Bộ nguyên tắc CARE về quản trị dữ liệu bản địa đặt vấn đề đúng ở chỗ này: chủ thể của dữ liệu phải là người quyết định dữ liệu được dùng vào việc gì, chứ không chỉ được thông báo rằng dữ liệu đã được dùng.

Ở góc độ quản trị, quyền được giữ mờ không phủ nhận dữ liệu mở mà đặt ra nguyên tắc mở có chọn lọc. Địa phương cần xác định rõ lớp dữ liệu nào phục vụ lợi ích công cộng và đổi mới sáng tạo, lớp nào chỉ nên chia sẻ với các đối tác đủ điều kiện và lớp nào phải được bảo vệ lâu dài vì liên quan đến an ninh, văn hóa, môi trường hoặc lợi thế cạnh tranh.

Một nền kinh tế tác nhân trưởng thành vì thế không được đo bằng lượng dữ liệu công khai tối đa, mà bằng khả năng thiết kế ranh giới thông minh giữa khả đọc và khả bảo tồn. Chủ quyền không chỉ là quyền xuất hiện trong thế giới của máy, mà còn là quyền quyết định phần nào của mình sẽ mãi chỉ thuộc về con người.

### 9.4.3. Dữ liệu công khác dữ liệu thương mại

Một trong những điều kiện để xây dựng chủ quyền dữ liệu là phân biệt rõ dữ liệu công và dữ liệu thương mại. Hai loại dữ liệu này đều có giá trị kinh tế nhưng phục vụ những mục tiêu hoàn toàn khác nhau. Nếu đánh đồng chúng, địa phương hoặc sẽ quá đóng khiến hệ sinh thái không thể phát triển, hoặc quá mở đến mức làm mất động lực đầu tư của doanh nghiệp.

Dữ liệu công là lớp dữ liệu nền mà toàn bộ xã hội cùng hưởng lợi khi được chia sẻ. Đó là bản đồ hành chính, thực thể địa danh, địa chỉ chuẩn hóa, giao thông, thông tin di tích, giờ mở cửa của cơ quan công quyền, hạ tầng công cộng, lịch sự kiện, quy hoạch, cảnh báo thiên tai hay các tiêu chuẩn dữ liệu dùng chung. Đây là những dữ liệu càng được chuẩn hóa và mở thì chi phí giao dịch của toàn bộ nền kinh tế càng giảm. Chúng tạo ra một "ngôn ngữ chung" để mọi doanh nghiệp, nhà phát triển và tác nhân AI có thể kết nối.

Ngược lại, dữ liệu thương mại phản ánh lợi thế cạnh tranh của từng tổ chức. Tồn kho phòng khách sạn theo thời gian thực, chiến lược định giá động, dữ liệu khách hàng, lịch sử giao dịch, mô hình dự báo nhu cầu hay thuật toán đề xuất của doanh nghiệp là những tài sản được tạo ra bằng vốn đầu tư và năng lực vận hành. Chúng không nên mặc nhiên được coi là dữ liệu công, bởi nếu mọi lợi thế đều bị buộc phải chia sẻ, động lực đổi mới sẽ suy giảm.

Điều quan trọng là hai lớp dữ liệu này phải bổ sung cho nhau. Hạ tầng dữ liệu công giúp doanh nghiệp kết nối với thị trường hiệu quả hơn, còn dữ liệu thương mại tạo ra sự khác biệt, cạnh tranh và đổi mới. Một nền kinh tế tác nhân phát triển không phải là nơi mọi dữ liệu đều mở, mà là nơi dữ liệu nền được mở để khuyến khích sáng tạo, trong khi dữ liệu tạo lợi thế cạnh tranh vẫn được bảo vệ bằng những cơ chế quản trị rõ ràng.

### 9.4.4. Khi mở dữ liệu nhưng không mất bản sắc

Không ít địa phương lo ngại rằng càng chuẩn hóa dữ liệu cho AI thì bản sắc địa phương càng bị làm phẳng. Lo ngại đó chỉ đúng khi mở dữ liệu đồng nghĩa với việc giản lược mọi giá trị thành một vài thuộc tính giống nhau. Thực tế, dữ liệu mở và bản sắc không phải hai mục tiêu đối lập.

Điều cần mở là cấu trúc dữ liệu, chứ không phải làm nghèo nội dung. Một di tích có thể được mô tả bằng các chuẩn quốc tế về vị trí, thời gian mở cửa, khả năng tiếp cận hay thông tin vé, đồng thời vẫn lưu giữ đầy đủ câu chuyện lịch sử, truyền thuyết, nghi lễ, nhân vật và bối cảnh văn hóa riêng. Chuẩn hóa giúp máy hiểu cách đọc dữ liệu, còn bản sắc nằm ở nội dung mà dữ liệu truyền tải.

Thực tế, dữ liệu càng có ngữ cảnh thì AI càng có khả năng biểu diễn đúng bản sắc của địa phương. Một lễ hội không chỉ là một sự kiện diễn ra vào một ngày cố định, mà còn là kết quả của lịch sử, tín ngưỡng và ký ức cộng đồng. Nếu những lớp thông tin ấy được tổ chức thành dữ liệu có cấu trúc thay vì chỉ tồn tại trong các bài viết rời rạc, AI sẽ có nhiều khả năng diễn giải chính xác hơn thay vì rút gọn thành một vài nhãn phân loại đơn giản.

Mục tiêu của dữ liệu mở không phải tạo ra những địa phương giống nhau, mà là giúp mỗi địa phương có thể được máy nhận diện đúng với bản sắc riêng của mình. Chuẩn hóa là chuẩn hóa ngôn ngữ trao đổi dữ liệu, không phải chuẩn hóa văn hóa. Chính sự khác biệt được biểu diễn bằng một cấu trúc thống nhất mới là nền tảng của một hệ sinh thái AI đa dạng và công bằng.

### 9.4.5. Hạ tầng dữ liệu công giống đường và điện

Trong nhiều năm, đầu tư số thường được đánh giá thông qua những sản phẩm có thể nhìn thấy: Trang web mới, ứng dụng mới hay các chiến dịch truyền thông số. Nhưng hầu hết những yếu tố quyết định khả năng cạnh tranh của nền kinh tế tác nhân lại nằm ở những lớp hạ tầng mà người dùng gần như không bao giờ nhìn thấy.

Dữ liệu công cũng giống như đường sá, điện, nước hay viễn thông. Bản thân chúng không phải là trải nghiệm du lịch, cũng không trực tiếp tạo ra doanh thu. Không ai đến một địa phương chỉ vì hệ thống định danh thực thể được thiết kế tốt hay vì API của chính quyền hoạt động ổn định. Nhưng nếu thiếu những hạ tầng đó, phần lớn trải nghiệm hiện đại sẽ không thể vận hành một cách hiệu quả.

Một bản đồ thực thể chính xác giúp AI không nhầm lẫn giữa các địa danh. Một hệ thống dữ liệu sự kiện được cập nhật liên tục giúp tác nhân lập kế hoạch đáng tin cậy. Một chuẩn dữ liệu thống nhất về điểm tham quan, khách sạn, giao thông và dịch vụ công giúp hàng trăm doanh nghiệp có thể kết nối mà không phải xây dựng lại từ đầu. Cũng như một con đường tốt không chỉ phục vụ một chiếc xe, hạ tầng dữ liệu tốt phục vụ đồng thời hàng nghìn dịch vụ và hàng triệu giao dịch.

Giá trị lớn nhất của hạ tầng dữ liệu công nằm ở hiệu ứng lan tỏa. Mỗi bộ dữ liệu nền được chuẩn hóa và duy trì tốt sẽ làm giảm chi phí cho mọi doanh nghiệp, mọi nền tảng và mọi tác nhân AI tham gia vào hệ sinh thái. Đổi mới không còn phụ thuộc vào việc ai sở hữu nhiều dữ liệu nhất, mà vào việc toàn xã hội có một nền tảng dữ liệu đủ tin cậy để mọi chủ thể cùng xây dựng giá trị mới.

Đó cũng là lý do đầu tư vào hạ tầng dữ liệu công không nên được xem là một dự án công nghệ, mà là một khoản đầu tư phát triển tương đương với đường cao tốc, lưới điện hay mạng viễn thông quốc gia. Trong nền kinh tế tác nhân, đường dẫn cho dữ liệu cũng quan trọng như đường dẫn cho hàng hóa, bởi chính trên lớp hạ tầng vô hình ấy mà các trải nghiệm, giao dịch và chuỗi giá trị mới được hình thành.

Quyền được giữ mờ nghe hợp lý về văn hóa nhưng trên thực tế sẽ trở thành căn cứ để trì hoãn công bố những dữ liệu mà xã hội có quyền biết. Nguy cơ đó có thật và không nên xem nhẹ. Vì vậy quyền mờ chỉ có nghĩa khi đi kèm ba điều kiện: danh mục các lớp được giữ mờ phải công khai, lý do và thời hạn của mỗi lớp phải được ghi nhận và quyết định giữ mờ phải được rà soát định kỳ với sự tham gia của cộng đồng liên quan. Xã hội có thể không biết nội dung nhưng phải biết cái gì đang không được mở và vì sao. Thiếu ba điều kiện ấy thì đúng là quyền mờ chỉ là một cách nói khác của việc không công bố và nguyên tắc quản trị ở cuối mục này được viết ra để chặn chính khả năng đó.

> ▪ ***Ranh giới khả đọc:***
> • ***Phân lớp trước khi mở.*** *Mọi tập dữ liệu về địa phương được xếp vào một trong ba chế độ công, chia sẻ có điều kiện hoặc bảo vệ dài hạn. Không tập dữ liệu nào được để ở trạng thái chưa phân loại.*
> • ***Dữ liệu nền mở theo mặc định.*** *Thực thể, địa danh, địa chỉ, giao thông, lịch sự kiện và cảnh báo được coi là hạ tầng công, mở nếu không có lý do ngược lại được ghi nhận bằng văn bản.*
> • ***Mờ phải có lý do và thời hạn.*** *Mỗi quyết định giữ mờ ghi rõ căn cứ, phạm vi, thời hạn hiệu lực và đơn vị chịu trách nhiệm rà soát lại.*

> • ***Cộng đồng liên quan tham gia quyết định.*** *Với tri thức văn hóa, nghi lễ và không gian tín ngưỡng, chủ thể quyết định phần được đưa vào lược đồ phải bao gồm cộng đồng đang giữ tri thức ấy.*
> • ***Danh mục lớp mờ được công bố.*** *Xã hội biết cái gì đang không được mở và vì sao, dù không biết nội dung.*

Chủ quyền không chỉ là quyền được máy nhìn thấy mà còn là quyền quyết định máy được nhìn thấy đến đâu.

PHẦN B
# TRÁCH NHIỆM

## 9.5. TÁC NHÂN ĐẶT SAI THÌ AI CHỊU

Trong nền kinh tế truyền thống, khi một du khách tự mình lựa chọn khách sạn, nhà hàng hay điểm tham quan, việc xác định trách nhiệm tương đối rõ ràng. Nếu thông tin sai, đơn vị cung cấp phải chịu trách nhiệm. Nếu quảng cáo gây hiểu nhầm, nền tảng hoặc doanh nghiệp có thể bị khiếu nại. Nếu người dùng tự quyết định dù đã có đầy đủ thông tin, một phần trách nhiệm thuộc về chính họ.

Nhưng khi quyết định được ủy nhiệm cho tác nhân AI, cấu trúc trách nhiệm trở nên phức tạp hơn nhiều. Người dùng chỉ đưa ra mục tiêu, còn quá trình tìm kiếm, so sánh, suy luận và lựa chọn được thực hiện bởi một hệ thống trung gian mà họ không trực tiếp kiểm soát. Khi kết quả không như mong đợi, câu hỏi đầu tiên không còn là "ai đã chọn", mà là "ai đã khiến máy chọn như vậy".

Một tác nhân có thể đặt nhầm khách sạn vì dữ liệu cũ, chọn một điểm đến đã đóng cửa vì hệ thống chưa cập nhật, hoặc ưu tiên một nhà cung cấp chỉ vì thuật toán đánh giá đó là phương án tối ưu. Trong mỗi trường hợp, sai sót có thể bắt nguồn từ nhiều lớp khác nhau: dữ liệu, mô hình suy luận, nền tảng phân phối, nhà cung cấp dịch vụ hoặc chính yêu cầu ban đầu của người dùng. Càng nhiều lớp trung gian tham gia, ranh giới trách nhiệm càng trở nên mờ nhạt.

Điều này tạo ra một nghịch lý mới của nền kinh tế tác nhân. Quyền quyết định được tự động hóa nhưng trách nhiệm pháp lý

và đạo đức không thể tự động hóa theo. Xã hội không thể chấp nhận một hệ thống mà mọi sai sót đều được giải thích bằng câu trả lời đơn giản: "AI đã chọn". Ngược lại, mỗi quyết định của tác nhân phải luôn có khả năng truy vết về nguồn dữ liệu, quy trình suy luận và chủ thể chịu trách nhiệm đối với từng mắt xích trong chuỗi quyết định.

Vì vậy, vấn đề cốt lõi của nền kinh tế tác nhân không chỉ là làm cho AI thông minh hơn, mà còn là xây dựng một kiến trúc trách nhiệm đủ minh bạch để con người vẫn có thể xác định ai chịu trách nhiệm khi máy hành động thay mình. Trong bối cảnh đó, trách nhiệm không biến mất, nó chỉ dịch chuyển từ người thực hiện sang những người thiết kế, vận hành và cung cấp hạ tầng cho quá trình ra quyết định của máy. Đây cũng là lý do vì sao, sau khi bàn về chủ quyền dữ liệu và chủ quyền thuật toán, chương này chuyển sang một câu hỏi quan trọng không kém: làm thế nào để phân bổ trách nhiệm trong một thế giới mà tác nhân ngày càng thay con người đưa ra quyết định.

Trách nhiệm xếp chồng. Tự động hóa một quyết định không chuyển trách nhiệm sang máy mà phân bổ lại trách nhiệm giữa những chủ thể đã cung cấp dữ liệu, thiết kế mô hình, phân phối lựa chọn, vận hành hệ thống và đặt luật chơi. Trách nhiệm xếp chồng theo quyền lực mà mỗi lớp kiểm soát, không dịch chuyển dọc theo chuỗi kỹ thuật.

Ba nghĩa của chữ trách nhiệm. Tiếng Việt dùng một chữ trách nhiệm cho ba khái niệm mà tiếng Anh tách riêng và sự nhập nhằng ấy làm các cuộc thảo luận về AI thường đi vòng. Trách nhiệm pháp lý là nghĩa vụ bồi thường hoặc chịu chế tài khi có thiệt hại, do luật xác định. Trách nhiệm giải trình là nghĩa vụ chứng minh mình đã làm gì và vì sao, kể cả khi không có thiệt

hại. Trách nhiệm nghề nghiệp là chuẩn mực hành xử của một chủ thể trong lĩnh vực của mình, thường cao hơn mức luật đòi hỏi. Mục này bàn chủ yếu về trách nhiệm giải trình, vì đó là phần đang thiếu hạ tầng, còn phần trách nhiệm pháp lý phụ thuộc vào từng hệ thống pháp luật và chưa định hình ở phần lớn quốc gia.

### 9.5.1. Sai vì dữ liệu

Tác nhân không quan sát thế giới trực tiếp. Nó hành động trên phiên bản số của thế giới được tạo ra từ cơ sở dữ liệu, trang web, bản đồ, hệ thống đặt chỗ, đánh giá người dùng và các nguồn thông tin mà nó có thể truy cập. Vì vậy, một quyết định nhìn có vẻ sai về mặt thuật toán đôi khi chỉ là kết quả hợp lý được tạo ra từ dữ liệu sai.

Một khách sạn đã đóng cửa nhưng vẫn xuất hiện trên bản đồ. Một bảo tàng đã đổi giờ mở cửa nhưng trang web chưa cập nhật. Một tuyến đường đang sửa chữa nhưng hệ thống vẫn coi là có thể lưu thông. Một nhà hàng được gắn nhãn phù hợp với người dị ứng, trong khi thực đơn thực tế không còn bảo đảm điều đó. Trong những trường hợp này, tác nhân không nhất thiết suy luận sai. Nó có thể đã suy luận hoàn toàn nhất quán từ một bức tranh lỗi thời hoặc không đầy đủ.

> ▪ ***Sai vì dữ liệu thường xuất hiện dưới bốn hình thức:***
> • ***Thứ nhất là dữ liệu không chính xác.*** *Thông tin được nhập sai ngay từ đầu: vị trí sai, giá sai, mô tả sai hoặc thuộc tính dịch vụ không phản ánh thực tế.*
> • ***Thứ hai là dữ liệu đã lỗi thời.*** *Thông tin từng đúng nhưng không còn đúng tại thời điểm tác nhân ra quyết định. Đây là một rủi ro đặc biệt lớn đối với giá, tồn kho, lịch hoạt động,*

*điều kiện giao thông và khả năng tiếp cận.*

- ***Thứ ba là dữ liệu thiếu.*** *Hệ thống không biết rằng một điểm đến không phù hợp với người cao tuổi, một khách sạn không có thang máy hoặc một hoạt động chỉ diễn ra trong điều kiện thời tiết nhất định. Khi dữ liệu thiếu, máy thường phải suy đoán từ những thuộc tính gần nhất.*
- ***Thứ tư là dữ liệu mâu thuẫn.*** *Trang web chính thức, nền tảng đặt phòng, bản đồ số và đánh giá người dùng cùng mô tả một thực thể theo những cách khác nhau. Tác nhân phải tự quyết định nguồn nào đáng tin hơn, nhưng không phải lúc nào nó cũng có đủ cơ sở để làm việc đó.*

Vấn đề trách nhiệm bắt đầu từ đây. Nếu doanh nghiệp không cập nhật giờ mở cửa, doanh nghiệp phải chịu một phần trách nhiệm. Nếu nền tảng tiếp tục hiển thị thông tin cũ dù đã nhận được dữ liệu mới, trách nhiệm dịch chuyển sang nền tảng. Nếu dữ liệu được tổng hợp từ nhiều nguồn nhưng không ghi rõ nguồn gốc, đơn vị xây dựng hệ thống cũng phải chịu trách nhiệm về khả năng truy xuất và kiểm chứng.

Do đó, quản trị dữ liệu trong nền kinh tế tác nhân không thể chỉ dừng ở việc công bố thông tin. Mỗi thuộc tính quan trọng cần có nguồn, thời điểm cập nhật, chủ thể xác nhận và mức độ tin cậy. Một hệ thống không biết dữ liệu đến từ đâu cũng là một hệ thống không biết phải quy trách nhiệm cho ai khi dữ liệu sai.

Sai vì dữ liệu cho thấy một nguyên tắc căn bản, trước khi yêu cầu tác nhân chịu trách nhiệm về quyết định, cần bảo đảm rằng thế giới mà tác nhân đang đọc không phải là một bản sao méo mó của thực tế.

### 9.5.2. Sai vì thuật toán

Dữ liệu có thể đúng nhưng tác nhân vẫn lựa chọn sai. Khi đó, vấn đề nằm ở cách hệ thống diễn giải, xếp hạng và tối ưu dữ liệu.

Thuật toán không chỉ tìm kiếm thông tin. Nó quyết định thuộc tính nào quan trọng hơn, tiêu chí nào được ưu tiên và phương án nào được coi là "tốt nhất". Hai tác nhân cùng đọc một tập dữ liệu có thể đưa ra hai lựa chọn hoàn toàn khác nhau chỉ vì chúng được thiết kế với các hàm mục tiêu khác nhau.

Một hệ thống tối ưu giá thấp nhất có thể chọn khách sạn xa trung tâm. Một hệ thống tối ưu thời gian có thể bỏ qua trải nghiệm văn hóa. Một hệ thống ưu tiên xếp hạng phổ biến có thể liên tục đề xuất những điểm đã đông khách và loại bỏ những lựa chọn nhỏ nhưng phù hợp hơn. Một tác nhân tối ưu xác suất hoàn tất giao dịch có thể ưu tiên nhà cung cấp dễ đặt thay vì nhà cung cấp tốt nhất cho người dùng.

Sai vì thuật toán không nhất thiết là lỗi lập trình. Nhiều khi hệ thống hoạt động đúng như được thiết kế, nhưng mục tiêu thiết kế lại không trùng với lợi ích thực sự của người sử dụng.

> ▪ ***Có ít nhất ba dạng sai lệch thuật toán:***
> • ***Dạng thứ nhất là sai trong mô hình suy luận.*** *Tác nhân hiểu sai quan hệ giữa các biến, đánh giá quá cao một tín hiệu hoặc bỏ qua một điều kiện có ý nghĩa quyết định.*
> • ***Dạng thứ hai là sai trong mục tiêu tối ưu.*** *Hệ thống tối ưu doanh thu, tỷ lệ đặt thành công hoặc thời gian phản hồi trong khi người dùng quan tâm đến an toàn, bản sắc, sự riêng tư hoặc chất lượng trải nghiệm.*

> • ***Dạng thứ ba là sai trong cơ chế xếp hạng.*** *Một phương án được đưa lên đầu không phải vì phù hợp nhất mà vì được hệ thống ưu tiên, tài trợ hoặc có nhiều dữ liệu hơn.*

Trách nhiệm trong trường hợp này thuộc trước hết về chủ thể thiết kế và triển khai thuật toán. Họ quyết định biến nào được đưa vào mô hình, biến nào bị loại bỏ, trọng số được đặt như thế nào và hệ thống được phép tự chủ đến đâu. Không thể viện dẫn tính phức tạp của mô hình để phủ nhận trách nhiệm đối với mục tiêu mà mô hình được giao.

Một kiến trúc trách nhiệm đúng phải yêu cầu tác nhân giải thích ít nhất ba điều: nó đã tối ưu điều gì, dựa trên những tiêu chí nào và vì sao loại bỏ những phương án còn lại. Người dùng không cần nhìn thấy toàn bộ mã nguồn, nhưng phải hiểu được logic đủ để nhận ra quyết định đó phục vụ lợi ích của ai.

Sai vì thuật toán nhắc rằng tính chính xác của dữ liệu chưa đủ. Một hệ thống có thể đọc đúng thế giới nhưng vẫn hành động sai nếu nó được dạy tối ưu sai mục tiêu.

### 9.5.3. Sai vì nền tảng

Giữa tác nhân và nhà cung cấp dịch vụ thường tồn tại một lớp trung gian: nền tảng tìm kiếm, OTA, bản đồ số, hệ thống thanh toán, chợ ứng dụng hoặc mạng lưới phân phối dữ liệu. Lớp này không chỉ truyền tải thông tin. Nó định hình toàn bộ không gian lựa chọn mà tác nhân có thể nhìn thấy.

Một tác nhân không thể chọn một dịch vụ nếu dịch vụ đó không được nền tảng lập chỉ mục. Nó khó đề xuất một doanh

nghiệp nếu doanh nghiệp không có khả năng đặt trực tuyến. Nó có thể bỏ qua một điểm đến nếu nền tảng không hỗ trợ định dạng dữ liệu, phương thức thanh toán hoặc cơ chế xác nhận của nơi đó. Quyết định sai đôi khi không bắt nguồn từ dữ liệu hay thuật toán của tác nhân, mà từ cấu trúc của nền tảng mà tác nhân đang phụ thuộc.

Nền tảng có thể làm sai theo nhiều cách. Nó có thể loại bỏ một số lựa chọn khỏi tập kết quả. Nó có thể ưu tiên đối tác trả phí. Nó có thể trì hoãn cập nhật dữ liệu. Nó có thể gom nhiều thực thể khác nhau thành một. Nó cũng có thể áp đặt những điều kiện giao dịch khiến tác nhân chỉ nhìn thấy các nhà cung cấp đã tích hợp sâu với hệ sinh thái của nền tảng. Trong trường hợp đó, tác nhân có thể chọn đúng trong tập lựa chọn mà nó được cung cấp, nhưng tập lựa chọn ban đầu đã bị bóp méo. Đây là một dạng quyền lực đặc biệt vì nền tảng thường tự mô tả mình như một trung gian trung lập. Trên thực tế, mọi quyết định về tiêu chuẩn dữ liệu, thứ tự hiển thị, quyền truy cập API, mức phí và điều kiện tham gia đều có tác động trực tiếp đến việc ai được nhìn thấy và ai bị loại khỏi thị trường.

Trách nhiệm của nền tảng vì vậy không thể chỉ giới hạn ở việc "kết nối các bên". Nền tảng phải chịu trách nhiệm về tính toàn vẹn của thị trường mà nó tạo ra. Nếu kết quả tìm kiếm bị tài trợ, điều đó phải được công khai. Nếu một dịch vụ bị loại vì không tương thích kỹ thuật, hệ thống cần phân biệt rõ giữa "không tồn tại" và "không thể truy cập qua nền tảng". Nếu dữ liệu được ưu tiên theo quan hệ thương mại, tác nhân và người dùng cần được biết.

Ở cấp độ địa phương, rủi ro lớn nhất là để một vài nền tảng bên ngoài trở thành cửa ngõ duy nhất dẫn vào hệ sinh thái dịch

vụ. Khi đó, nền tảng không chỉ phân phối nhu cầu mà còn quyết định phiên bản nào của địa phương được tác nhân nhìn thấy.

Sai vì nền tảng cho thấy trách nhiệm không chỉ nằm ở người đưa ra lựa chọn cuối cùng. Chủ thể kiểm soát tập lựa chọn ban đầu cũng phải chịu trách nhiệm đối với những gì bị đưa vào, đẩy lên hoặc làm cho biến mất.

### 9.5.4. Sai vì người vận hành

Không phải mọi lỗi của hệ thống tác nhân đều bắt nguồn từ máy. Con người vẫn tham gia ở nhiều điểm: nhập dữ liệu, thiết lập quy tắc, lựa chọn nhà cung cấp, phê duyệt giao dịch, xử lý ngoại lệ và quyết định mức độ tự chủ của tác nhân. Vì vậy, một hệ thống có thể thất bại không phải vì AI thiếu năng lực, mà vì người vận hành sử dụng AI sai cách.

Người vận hành có thể giao cho tác nhân một mục tiêu mơ hồ. Họ có thể bỏ qua cảnh báo, thiết lập ngưỡng rủi ro quá cao hoặc cho phép hệ thống tự động thực hiện những giao dịch đáng lẽ phải có bước xác nhận của con người. Họ cũng có thể sử dụng tác nhân trong một lĩnh vực mà dữ liệu, quy trình hoặc năng lực giám sát chưa đủ trưởng thành.

Ví dụ, một đơn vị có thể yêu cầu hệ thống "đặt phương án rẻ nhất" nhưng không quy định tiêu chuẩn an toàn tối thiểu. Một doanh nghiệp có thể tự động hóa việc hủy đặt chỗ mà không xây dựng cơ chế xử lý tình huống đặc biệt. Một cơ quan quản lý có thể triển khai hệ thống đề xuất nhưng không phân công ai chịu trách nhiệm kiểm tra kết quả. Trong những trường hợp đó, lỗi không nằm ở một quyết định riêng lẻ mà ở thiết kế vận hành.

> ▪ ***Sai vì người vận hành thường xuất hiện ở bốn điểm:***
> • ***Thứ nhất là xác lập nhiệm vụ sai.*** *Mục tiêu không rõ ràng, mâu thuẫn hoặc thiếu các giới hạn cần thiết.*
> • ***Thứ hai là phân quyền sai.*** *Tác nhân được phép hành động vượt quá mức độ rủi ro mà tổ chức có thể kiểm soát.*
> • ***Thứ ba là giám sát không đầy đủ.*** *Không có người theo dõi, không có cảnh báo hoặc không có quy trình can thiệp khi hệ thống đi lệch.*
> • ***Thứ tư là không xử lý ngoại lệ.*** *Tổ chức giả định rằng mọi trường hợp đều có thể được tự động hóa, trong khi thực tế luôn tồn tại những tình huống cần phán đoán của con người.*

Trách nhiệm của người vận hành vì vậy không thể được chuyển hoàn toàn sang nhà phát triển công nghệ. Một công cụ có thể có giới hạn, nhưng tổ chức sử dụng công cụ phải hiểu giới hạn đó và thiết kế quy trình phù hợp. Việc lựa chọn sử dụng AI ở đâu, với quyền hạn nào và dưới sự giám sát của ai vẫn là một quyết định của con người.

Nguyên tắc quan trọng nhất là mức độ giám sát phải tăng theo mức độ hậu quả. Một tác nhân gợi ý nhà hàng có thể được trao quyền tự chủ cao hơn một tác nhân đặt dịch vụ y tế, xử lý dữ liệu nhạy cảm hoặc thực hiện giao dịch tài chính lớn. Không phải mọi quyết định đều cần con người phê duyệt, nhưng mọi quyết định có hậu quả đáng kể phải có một chủ thể con người chịu trách nhiệm cuối cùng.

Sai vì người vận hành là lời nhắc rằng tự động hóa không loại bỏ vai trò quản trị. Nó làm cho vai trò ấy quan trọng hơn. Khi máy được trao nhiều quyền hành động hơn, con người càng phải xác định rõ giới hạn, cơ chế giám sát và điểm dừng của hệ thống.

### 9.5.5. Chuỗi trách nhiệm mới

Trong nền kinh tế tác nhân, không còn tồn tại một chủ thể duy nhất có thể chịu trách nhiệm cho toàn bộ một quyết định. Mỗi giao dịch được hình thành từ một chuỗi các lớp hạ tầng, trong đó mỗi lớp đều đóng góp vào kết quả cuối cùng. Nếu một mắt xích sai, toàn bộ chuỗi có thể tạo ra một quyết định sai, ngay cả khi các mắt xích còn lại đều hoạt động đúng.

Khác với nền kinh tế truyền thống, nơi trách nhiệm thường tập trung vào doanh nghiệp trực tiếp phục vụ khách hàng, nền kinh tế tác nhân đòi hỏi một kiến trúc trách nhiệm phân tầng. Mỗi lớp phải chịu trách nhiệm đối với phần việc mà mình kiểm soát, thay vì đẩy toàn bộ trách nhiệm sang lớp kế tiếp hoặc viện dẫn rằng "AI đã quyết định".

Chuỗi trách nhiệm mới có thể được hình dung qua năm tầng:

- ***Tầng thứ nhất là nhà cung cấp dữ liệu (data provider).*** Đây là nơi tạo ra phiên bản số của thế giới thực. Doanh nghiệp, điểm đến, cơ quan quản lý và các tổ chức sở hữu dữ liệu phải chịu trách nhiệm về tính chính xác, tính đầy đủ và tính cập nhật của thông tin mà họ công bố. Một dữ liệu sai có thể lan truyền qua hàng chục hệ thống và tạo ra hàng triệu quyết định sai.

- ***Tầng thứ hai là nền tảng (platform).*** Nền tảng chịu trách nhiệm về việc thu thập, xác thực, đồng bộ và phân phối dữ liệu. Họ quyết định dữ liệu nào được lập chỉ mục, dữ liệu nào được ưu tiên và dữ liệu nào được đưa vào tập lựa chọn của tác nhân. Trách nhiệm của nền tảng không chỉ là truyền tải thông tin mà còn là bảo đảm tính toàn vẹn của môi trường giao dịch số.

- ***Tầng thứ ba là mô hình (model).*** Đây là lớp diễn giải dữ liệu để tạo ra quyết định. Nhà phát triển mô hình phải chịu trách nhiệm về mục tiêu tối ưu, tiêu chí xếp hạng, khả năng giải thích và cơ chế xử lý rủi ro của hệ thống. Một mô hình không thể được miễn trách nhiệm chỉ vì nó sử dụng học máy hay mạng nơ-ron phức tạp. Càng tự chủ, mô hình càng phải minh bạch về logic ra quyết định.

- ***Tầng thứ tư là nhà cung cấp dịch vụ (merchant).*** Dù giao dịch được khởi tạo bởi tác nhân hay con người, doanh nghiệp vẫn chịu trách nhiệm đối với chất lượng dịch vụ thực tế mà khách hàng nhận được. Khách sạn phải đúng với mô tả, nhà hàng phải phục vụ đúng tiêu chuẩn, điểm tham quan phải vận hành đúng như thông tin đã công bố. AI không thể thay thế trách nhiệm nghề nghiệp của đơn vị trực tiếp cung cấp dịch vụ.

- ***Tầng thứ năm là cơ quan quản lý (government).*** Nhà nước không chịu trách nhiệm cho từng quyết định của tác nhân, nhưng chịu trách nhiệm thiết kế luật chơi. Điều này bao gồm xây dựng tiêu chuẩn dữ liệu, quy định về minh bạch thuật toán, cơ chế truy xuất nguồn gốc, quy trình giải quyết tranh chấp và các nguyên tắc phân bổ trách nhiệm giữa các chủ thể trong hệ sinh thái. Khi một quyết định được tạo ra bởi nhiều lớp công nghệ khác nhau, điều xã hội cần nhất không phải là thêm quy định cho AI, mà là một khuôn khổ quản trị giúp xác định rõ trách nhiệm của từng mắt xích.

Điểm quan trọng là năm tầng này không thay thế lẫn nhau. Trách nhiệm không dịch chuyển hoàn toàn từ doanh nghiệp sang nền tảng, từ nền tảng sang mô hình hay từ mô hình sang chính phủ. Trách nhiệm được xếp chồng. Mỗi tầng đều phải chịu trách nhiệm đối với phần quyền lực mà mình đang nắm giữ.

Đây cũng là một trong những thay đổi căn bản của nền kinh tế tác nhân. Trước đây, chuỗi giá trị chủ yếu mô tả dòng chảy của hàng hóa, dịch vụ và tiền tệ. Nay, bên cạnh chuỗi giá trị xuất hiện một chuỗi trách nhiệm. Mỗi quyết định của AI đồng thời kéo theo một chuỗi nghĩa vụ về dữ liệu, thuật toán, nền tảng, vận hành và quản trị. Chỉ khi chuỗi trách nhiệm này được thiết kế đầy đủ, xã hội mới có thể trao ngày càng nhiều quyền hành động cho các tác nhân mà vẫn duy trì được niềm tin, khả năng giải trình và công bằng trong nền kinh tế số.

Trong nền kinh tế tác nhân, câu hỏi không còn là "AI có chịu trách nhiệm hay không?" mà là "mỗi chủ thể trong chuỗi tạo ra quyết định của AI phải chịu trách nhiệm đối với phần quyền lực mà mình kiểm soát như thế nào?" Chính khả năng trả lời rõ ràng câu hỏi này sẽ quyết định mức độ đáng tin cậy của toàn bộ hệ sinh thái tác nhân.

Một chuỗi năm tầng sẽ làm loãng trách nhiệm. Khi ai cũng chịu trách nhiệm thì thực tế không ai chịu. Rủi ro loãng trách nhiệm là lý do chuỗi này cần một quy tắc kèm theo, không phải lý do bỏ chuỗi. Quy tắc gồm hai phần: mỗi tầng chịu trách nhiệm về phần mình kiểm soát và không được viện dẫn tầng khác để miễn trừ, đồng thời mỗi quyết định có hậu quả đáng kể phải có một chủ thể con người chịu trách nhiệm cuối, được xác định trước khi hệ thống chạy. Mô hình một chủ thể duy nhất chịu tất cả nghe gọn hơn nhưng trong thực tế sẽ dồn trách nhiệm lên mắt xích nhỏ nhất và ít quyền lực nhất của chuỗi, thường là doanh nghiệp trực tiếp phục vụ khách.

▪ ***Phân bổ trách nhiệm trong chuỗi tác nhân:***

• ***Mức giám sát tăng theo hậu quả, không theo độ phức tạp công nghệ.*** *Một tác nhân gợi ý quán ăn và một tác nhân thực hiện giao dịch lớn không thể chịu cùng một chế độ giám sát. Nguyên tắc này nay có căn cứ pháp lý ở Việt Nam: Luật Trí tuệ nhân tạo năm 2025 đặt yêu cầu áp dụng biện pháp quản lý tương xứng với mức độ rủi ro.*

• ***Mỗi thuộc tính quan trọng đi kèm ba thông tin:*** *nguồn, thời điểm cập nhật và chủ thể xác nhận. Thiếu ba thông tin ấy thì hệ thống không thể quy trách nhiệm khi dữ liệu sai.*

• ***Không viện dẫn độ phức tạp của mô hình để miễn trách nhiệm về mục tiêu tối ưu.*** *Chủ thể chọn mục tiêu chịu trách nhiệm về mục tiêu, kể cả khi không giải thích được từng bước suy luận.*

• ***Mọi ưu tiên có tính thương mại trong tập lựa chọn phải được công khai*** *cho cả tác nhân và người dùng và phải phân biệt rõ giữa không tồn tại và không truy cập được qua nền tảng.*

• ***Mỗi quyết định có hậu quả đáng kể có một chủ thể con người chịu trách nhiệm cuối,*** *được xác định trước khi hệ thống được phép chạy, không phải sau khi sự việc xảy ra.*

Quyền quyết định có thể tự động hóa. Trách nhiệm thì không.

## 9.6. MÔ TẢ SAI VÀ THIỆT HẠI THƯƠNG HIỆU

Trong nền kinh tế trải nghiệm, một thông tin sai về địa phương thường chỉ ảnh hưởng đến một nhóm người và còn có cơ hội được sửa chữa qua những trải nghiệm trực tiếp. Nhưng trong nền kinh tế tác nhân, một mô tả sai có thể nhanh chóng trở thành "sự thật mặc định" đối với hàng triệu lượt suy luận của máy. Khi các mô hình AI, công cụ tìm kiếm, nền tảng đặt dịch vụ và các tác nhân AI cùng học từ những nguồn dữ liệu giống nhau, sai sót không chỉ được lặp lại mà còn được khuếch đại qua nhiều lớp trung gian.

Điều đáng lo ngại là thương hiệu địa phương không bị tổn hại chỉ bởi những thông tin hoàn toàn sai. Một mô tả thiếu bối cảnh, một thuộc tính đã lỗi thời, một mối quan hệ thực thể bị gắn nhầm hay một kết luận suy diễn không chính xác cũng có thể làm thay đổi cách máy hiểu về cả một điểm đến. Từ đó, AI có thể giới thiệu sai đối tượng khách, xếp hạng thấp hơn, hoặc đơn giản là không còn đưa địa phương vào tập lựa chọn ban đầu. Những thiệt hại này diễn ra âm thầm, khó quan sát nhưng tác động trực tiếp đến dòng khách, doanh thu và hình ảnh thương hiệu.

Khác với thời đại truyền thông đại chúng, nơi một chiến dịch quảng bá lớn có thể nhanh chóng điều chỉnh nhận thức của công chúng, trong nền kinh tế tác nhân, việc sửa một mô tả sai thường chậm hơn rất nhiều so với tốc độ lan truyền của nó. Dữ liệu sai có thể đã được sao chép vào nhiều kho tri thức, nhiều mô hình và nhiều hệ thống khác nhau trước khi chủ thể liên quan kịp phát hiện. Vì vậy, bảo vệ thương hiệu không còn chỉ là tạo ra những câu chuyện hấp dẫn, mà còn là liên tục kiểm toán, xác minh và hiệu chỉnh cách máy đang mô tả mình.

Nói cách khác, trong nền kinh tế tác nhân, một câu mô tả sai không chỉ là một lỗi thông tin. Đó có thể là điểm khởi đầu của một chuỗi suy luận sai, một chuỗi quyết định sai và cuối cùng là một chuỗi tổn thất kinh tế kéo dài. Thương hiệu của địa phương vì thế phụ thuộc không chỉ vào những gì mình làm, mà còn vào cách các hệ thống AI đang kể lại câu chuyện ấy.

Chi phí hiệu chỉnh. Chi phí sửa một mô tả sai tăng theo số hệ thống đã sao chép nó, còn số hệ thống sao chép lại tăng theo thời gian mà sai lệch tồn tại mà không bị phát hiện. Vì vậy chi phí hiệu chỉnh tăng nhanh hơn thời gian trôi qua và mọi chậm trễ đều đắt hơn mức người quản lý cảm nhận tại thời điểm trì hoãn.

Một khu di tích mở lại mà máy không biết. Giả thiết một khu di tích đóng cửa mười tám tháng để trùng tu rồi mở lại. Nguồn chính thức của tỉnh cập nhật ngay, nhưng phần lớn các nguồn khác vẫn ghi là đóng cửa: bản đồ số, vài nền tảng đặt dịch vụ, một số bài viết cũ được nhiều nơi dẫn lại. Từ đó chuỗi sự việc diễn ra theo một vòng lặp. Tác nhân lập lịch trình loại khu di tích khỏi các phương án vì cho rằng nơi này không hoạt động. Số lượt tìm kiếm và giao dịch liên quan giảm, nên các nền tảng ghi nhận ít tín hiệu quan tâm hơn. Ít tín hiệu quan tâm lại củng cố đánh giá rằng địa điểm này không đáng đưa vào tập lựa chọn. Doanh nghiệp quanh khu di tích thấy khách giảm mà không hiểu vì sao, vì không có bài báo tiêu cực nào và không có khủng hoảng nào xảy ra. Vòng lặp có thể viết gọn: thông tin sai dẫn tới bị loại khỏi đề xuất, bị loại khỏi đề xuất dẫn tới ít giao dịch, ít giao dịch dẫn tới ít tín hiệu mới và ít tín hiệu mới làm cho thông tin sai ngày càng trông có lý. Bốn cơ chế được phân tích ở các tiểu mục sau, gồm ảo giác của mô hình, ký ức thuật toán, thiệt hại vô hình và trạng thái khóa, đều có thể nhìn thấy trong cùng một tình huống như vậy.

### 9.6.1. Khi máy tạo ra một sự thật không tồn tại

Một trong những khác biệt lớn nhất giữa công cụ tìm kiếm truyền thống và mô hình ngôn ngữ lớn là AI không chỉ truy xuất thông tin mà còn tạo ra câu trả lời. Chính khả năng sinh nội dung này khiến AI có thể tạo ra những thông tin nghe rất hợp lý nhưng hoàn toàn không đúng sự thật. Trong nghiên cứu về AI, hiện tượng này được gọi là ảo giác của mô hình (hallucination), khi mô hình tự tạo ra dữ liệu, sự kiện, nguồn tham khảo hoặc mối quan hệ không tồn tại nhưng vẫn trình bày với mức độ tự tin rất cao.

Đối với một địa phương, hậu quả của hiện tượng này có thể vượt xa một lỗi nhỏ về mặt ngôn từ. AI có thể gán nhầm một lễ hội cho sai tỉnh, mô tả sai lịch sử của một di tích, khẳng định một tuyến cáp treo đã hoạt động dù dự án còn đang xây dựng, hoặc giới thiệu một món ăn đặc sản vốn thuộc một vùng khác. Đối với người đọc, những thông tin ấy rất khó bị phát hiện vì chúng thường được viết mạch lạc, logic và phù hợp với ngữ cảnh.

Nguy hiểm hơn, phần lớn người dùng không có thói quen kiểm chứng lại từng câu trả lời của AI. Khi thông tin được trình bày bằng giọng văn chắc chắn, người dùng có xu hướng coi đó là tri thức đã được xác thực. Vì vậy, một lỗi suy luận của mô hình có thể nhanh chóng trở thành cơ sở cho các quyết định du lịch, đầu tư hoặc truyền thông.

Điều này đặt ra một yêu cầu mới đối với quản trị thương hiệu địa phương. Không thể chỉ tập trung xây dựng thông tin đúng mà còn phải chủ động phát hiện những thông tin sai do AI tạo ra. Kiểm toán định kỳ các câu trả lời của những hệ thống AI phổ biến sẽ dần trở thành một phần của công tác quản trị thương

hiệu, tương tự như việc theo dõi báo chí hay mạng xã hội trước đây. Ảo giác của mô hình vì thế không chỉ là vấn đề kỹ thuật của mô hình. Đó là một rủi ro truyền thông mới, nơi chính máy có thể vô tình tạo ra những câu chuyện chưa từng tồn tại.

### 9.6.2. Sai nhưng được lặp lại thành đúng

Trong xã hội, một lời đồn có thể trở thành "sự thật" nếu đủ nhiều người nhắc lại. Trong nền kinh tế tác nhân, cơ chế tương tự cũng xuất hiện nhưng với tốc độ lớn hơn nhiều. Một thông tin sai khi được nhiều hệ thống cùng lặp lại sẽ dần hình thành một dạng "ký ức thuật toán" (algorithmic memory), nơi các mô hình và nền tảng cùng củng cố lẫn nhau thay vì kiểm chứng độc lập.

Quá trình này thường diễn ra theo một vòng lặp. Một thông tin sai xuất hiện trên một trang web hoặc cơ sở dữ liệu. Hệ thống tìm kiếm lập chỉ mục thông tin đó. Một mô hình AI học hoặc tham chiếu từ nguồn ấy và tạo ra các câu trả lời tương tự. Người dùng tiếp tục sao chép câu trả lời của AI sang blog, mạng xã hội hoặc các bài viết mới. Những nội dung mới lại trở thành dữ liệu cho các mô hình tiếp theo. Sai sót ban đầu vì thế không mất đi mà ngày càng được khuếch đại.

Điều đáng lo ngại là sau một thời gian, gần như mọi nguồn đều nói giống nhau. Khi đó, việc xác định đâu là thông tin gốc trở nên rất khó khăn. Không phải vì thông tin ấy đúng hơn, mà vì nó đã được lặp lại quá nhiều lần. Trong môi trường số, tần suất xuất hiện đôi khi bị nhầm lẫn với mức độ tin cậy.

Đối với thương hiệu địa phương, đây là một rủi ro nghiêm trọng. Nếu một điểm đến bị gắn với một mô tả sai, một hình ảnh

lỗi thời hoặc một định kiến không chính xác, sự lặp lại liên tục của các hệ thống AI có thể khiến mô tả đó trở thành nhận thức mặc định của máy. Từ đó, mọi suy luận, gợi ý và khuyến nghị đều tiếp tục dựa trên nền tảng sai lệch ấy.

Quản trị thương hiệu trong thời đại AI không chỉ là sửa một bài viết hay một trang web. Mục tiêu là phá vỡ vòng lặp của algorithmic memory bằng cách liên tục cung cấp nguồn dữ liệu chính thống, cập nhật và có thẩm quyền cao hơn. Khi nguồn đúng đủ mạnh và đủ nhất quán, ký ức của máy mới có cơ hội được hiệu chỉnh. Đây cũng là lý do việc duy trì freshness, authority và consistency của dữ liệu không còn là hoạt động truyền thông, mà trở thành một phần của hạ tầng bảo vệ thương hiệu.

### 9.6.3. Danh tiếng bị hao mòn âm thầm

Không phải mọi thiệt hại đối với thương hiệu đều xuất hiện dưới dạng một cuộc khủng hoảng truyền thông. Trong nền kinh tế tác nhân, phần lớn tổn thất diễn ra lặng lẽ hơn nhiều. Không có bài báo chỉ trích, không có làn sóng phản đối trên mạng xã hội, cũng không có sự sụt giảm đột ngột về lượng khách. Thay vào đó là hàng nghìn quyết định nhỏ của các tác nhân AI, mỗi quyết định đều làm địa phương mất đi một cơ hội mà không ai nhận ra.

Một AI du lịch không đề xuất điểm đến vì hiểu sai khả năng tiếp cận. Một trợ lý số loại bỏ khách sạn vì cho rằng thông tin không đáng tin cậy. Một hệ thống lập lịch bỏ qua một lễ hội vì dữ liệu không được cập nhật. Mỗi lần như vậy chỉ là một giao dịch không diễn ra. Nhưng khi hàng triệu tác nhân cùng hoạt động mỗi ngày, những mất mát nhỏ ấy tích lũy thành thiệt hại kinh tế đáng kể.

Điều khiến loại thiệt hại này nguy hiểm là nó rất khó đo lường. Chính quyền có thể vẫn thấy lượng khách ổn định. Doanh nghiệp vẫn có doanh thu. Nhưng tốc độ tăng trưởng chậm hơn đối thủ, tỷ lệ xuất hiện trong các đề xuất AI giảm dần, số lượt được đưa vào hành trình của máy ít đi và cơ hội tiếp cận khách hàng mới ngày càng thu hẹp. Thương hiệu không sụp đổ trong một ngày, nó bị bào mòn từng chút một.

Khác với truyền thông truyền thống, nơi danh tiếng chủ yếu tồn tại trong nhận thức của con người, danh tiếng trong nền kinh tế tác nhân còn tồn tại dưới dạng tín hiệu dành cho máy. Nếu những tín hiệu ấy liên tục suy yếu, thương hiệu sẽ dần mất khả năng cạnh tranh ngay cả khi chất lượng trải nghiệm thực tế không hề giảm. Một trong những nhiệm vụ mới của quản trị thương hiệu là phát hiện invisible damage trước khi nó phản ánh vào các chỉ số kinh tế. Chờ đến khi lượng khách giảm mới hành động thường là quá muộn, bởi quá trình suy giảm đã bắt đầu từ rất lâu trong lớp hạ tầng dữ liệu mà con người không nhìn thấy.

### 9.6.4. Chi phí sửa sai tăng theo thời gian

Một thông tin sai vừa xuất hiện thường khá dễ sửa. Chỉ cần cập nhật nguồn chuẩn, đính chính nội dung hoặc điều chỉnh thực thể liên quan là nhiều hệ thống sẽ dần tiếp nhận phiên bản mới. Nhưng càng để lâu, sai sót càng lan rộng sang nhiều nền tảng, nhiều mô hình và nhiều cơ sở dữ liệu khác nhau. Khi đó, chi phí sửa chữa không còn tăng tuyến tính mà tăng theo cấp số nhân.

Đây là hiện tượng có thể xem như trạng thái khóa (lock-in) của thông tin. Sau khi một mô tả đã được tích hợp vào các chỉ mục tìm kiếm, kho tri thức, mô hình AI, ứng dụng du lịch và

hàng loạt nội dung do người dùng tạo ra, không còn tồn tại một "nút sửa" duy nhất. Muốn hiệu chỉnh, phải tác động đồng thời lên nhiều mắt xích trong toàn bộ hệ sinh thái.

Đối với địa phương, điều này đồng nghĩa với việc mỗi tháng chậm trễ đều làm chi phí quản trị tăng lên. Không chỉ cần cập nhật trang web chính thức, mà còn phải đồng bộ dữ liệu với nền tảng bản đồ, OTA, cơ sở dữ liệu mở, đối tác thương mại, công cụ tìm kiếm và các nhà cung cấp mô hình AI. Càng nhiều hệ thống đã học từ dữ liệu cũ, quá trình phục hồi càng kéo dài.

Trạng thái khóa còn tạo ra một nghịch lý khác. Ngay cả khi thông tin chính thức đã được sửa, các hệ thống AI vẫn có thể tiếp tục tạo ra câu trả lời cũ vì chúng được huấn luyện trên dữ liệu lịch sử hoặc tham chiếu từ những nguồn chưa được cập nhật. Người dùng vì thế vẫn gặp phiên bản sai trong một thời gian dài, dù chủ thể liên quan đã hoàn thành việc đính chính.

Bài học quản trị là rõ ràng: chi phí thấp nhất luôn nằm ở giai đoạn phòng ngừa. Kiểm soát chất lượng dữ liệu từ đầu, phát hiện sai sót sớm và hiệu chỉnh ngay khi chúng xuất hiện luôn hiệu quả hơn nhiều so với việc khôi phục danh tiếng sau khi hệ sinh thái AI đã hình thành một trạng thái khóa. Trong nền kinh tế tác nhân, tốc độ sửa sai không còn chỉ là một yêu cầu truyền thông mà đã trở thành một lợi thế cạnh tranh.

### 9.6.5. Quản trị thương hiệu trở thành quản trị tri thức

Thương hiệu mạnh là thương hiệu được máy hiểu đúng. Trong phần lớn lịch sử của marketing, quản trị thương hiệu chủ yếu là quản trị nhận thức của con người. Địa phương đầu tư vào logo,

khẩu hiệu, chiến dịch quảng bá, sự kiện và quan hệ báo chí để tạo nên một hình ảnh nhất quán trong tâm trí công chúng. Tuy nhiên, trong nền kinh tế tác nhân, thương hiệu còn phải tồn tại trong một không gian khác: không gian tri thức của máy. Vì vậy, quản trị thương hiệu ngày càng dịch chuyển thành quản trị tri thức (knowledge governance).

Điều này có nghĩa là câu hỏi quan trọng không còn chỉ là "người ta nghĩ gì về địa phương?" mà còn là "máy biết gì về địa phương?". Mọi mô tả, thực thể, quan hệ, dữ liệu giao dịch và tín hiệu cập nhật đều góp phần hình thành phiên bản số của địa phương trong các mô hình AI. Nếu lớp tri thức này sai lệch, thiếu nhất quán hoặc lỗi thời, mọi nỗ lực truyền thông phía trên đều bị suy giảm hiệu quả.

Quản trị tri thức không chỉ là xuất bản nhiều dữ liệu hơn. Quan trọng hơn là quản trị chất lượng của dữ liệu. Thực thể phải có định danh rõ ràng, quan hệ giữa các thực thể phải chính xác, dữ liệu phải được cập nhật liên tục, các nguồn chính thức phải thống nhất và mọi thay đổi phải được lan truyền nhanh tới toàn bộ hệ sinh thái số. Thương hiệu vì thế không còn được bảo vệ bằng những thông điệp đẹp, mà bằng những cấu trúc tri thức đáng tin cậy.

Điều này cũng làm thay đổi vai trò của chính quyền địa phương. Ngoài việc là cơ quan quản lý hành chính và phát triển kinh tế, chính quyền dần trở thành người quản trị tri thức chính thức (knowledge steward) của địa phương. Nhiệm vụ mới bao gồm xác lập thực thể chuẩn, duy trì dữ liệu gốc, điều phối việc cập nhật giữa các sở ngành và doanh nghiệp, đồng thời xây dựng các cơ chế để nền tảng, đối tác và hệ thống AI có thể tiếp nhận thông tin chính xác.

Có thể xem đây là bước chuyển lớn nhất của thương hiệu trong kỷ nguyên AI. Nếu thế kỷ XX là thời đại quản trị hình ảnh và đầu thế kỷ XXI là thời đại quản trị trải nghiệm, thì nền kinh tế tác nhân mở ra giai đoạn quản trị tri thức. Thương hiệu mạnh không còn là thương hiệu được nói đến nhiều nhất mà là thương hiệu được các hệ thống AI hiểu đúng, nhớ đúng và dùng đúng khi ra quyết định.

Thiệt hại vô hình không đo được thì không nên đưa vào quản trị. Quản trị dựa trên cảm nhận sẽ dẫn tới chi tiêu không có căn cứ. Không đo trực tiếp được không đồng nghĩa với không đo được. Loại tổn thất này có thể theo dõi qua các chỉ báo trung gian: độ phủ thực thể, tỷ lệ mô tả đúng khi truy vấn các trợ lý phổ biến, độ tươi của thuộc tính quan trọng, số thực thể trùng lặp, tỷ lệ xuất hiện trong tập đề xuất. Không chỉ báo nào trong số đó là thước đo hoàn hảo, nhưng một bộ chỉ báo được đo đều đặn vẫn tốt hơn việc chờ tới lúc lượng khách giảm mới hành động. Một sai lệch không được sửa sẽ trở thành ký ức của máy.

## 9.7. HẠ TẦNG KHIẾU NẠI ĐỐI VỚI HỆ SINH THÁI AI

Hầu hết các địa phương hiện nay chưa có cơ chế chính thức để yêu cầu sửa tri thức mà máy đang sử dụng. Trong nhiều thập niên, khi một thông tin sai xuất hiện trên báo chí, trang web hay bản đồ, địa phương thường biết phải liên hệ với ai để yêu cầu đính chính. Có tòa soạn, có nhà xuất bản, có cơ quan quản lý hoặc có đơn vị vận hành hệ thống. Quy trình có thể chậm, nhưng ít nhất vẫn tồn tại một địa chỉ chịu trách nhiệm.

Trong nền kinh tế tác nhân, tình hình trở nên khác hẳn. Một câu trả lời của AI có thể được tạo ra từ hàng chục nguồn dữ liệu khác nhau: đồ thị tri thức, dữ liệu mở, trang web doanh nghiệp, OTA, mạng xã hội, tài liệu số hóa và cả những mô hình đã được huấn luyện từ nhiều tháng trước. Khi thông tin sai xuất hiện, địa phương thường không biết sai bắt nguồn từ đâu, càng không biết phải gửi yêu cầu sửa cho ai.

Đây là một khoảng trống quản trị hoàn toàn mới. Vấn đề không chỉ là tri thức sai, mà là không tồn tại một cơ chế chuẩn để sửa tri thức. Trong nhiều trường hợp, ngay cả khi biết thông tin không chính xác, địa phương vẫn không có một kênh chính thức để yêu cầu cập nhật trên toàn bộ hệ sinh thái AI.

Vì vậy, trong nền kinh tế tác nhân, quản trị thương hiệu địa phương không thể dừng ở việc công bố thông tin đúng, mà còn phải mở rộng sang khả năng phát hiện, khiếu nại và hiệu chỉnh tri thức sai trong các hệ thống máy. Đây sẽ là một chức năng mới của quản trị công số, tương tự như việc bảo vệ tên miền hay quản lý dữ liệu công ngày nay. Một địa phương không chỉ cần quyền nói về mình, mà còn cần quyền yêu cầu các hệ thống AI sửa lại cách chúng đang hiểu về mình. Đây cũng là bước chuyển từ quản trị thông tin sang quản trị tri thức máy (machine knowledge governance).

Hội tụ kênh hiệu chỉnh. Khi tỷ trọng giao dịch do tác nhân khởi tạo đủ lớn, các nền tảng AI quy mô lớn sẽ hội tụ về một chuẩn tối thiểu cho việc tiếp nhận và hiệu chỉnh tri thức về thực thể, tương tự cách ngành tên miền từng hội tụ về một quy trình giải quyết tranh chấp thống nhất. Giả thuyết này có thể kiểm chứng bằng ba dấu hiệu quan sát được: sự xuất hiện của giao diện hiệu chỉnh công khai, việc công bố thời hạn phản hồi và mức độ giống nhau giữa biểu mẫu của các nền tảng lớn.

Trong hai năm 2024 và 2025, Việt Nam ban hành bốn văn bản luật trực tiếp liên quan đến lập luận của chương này. Luật Dữ liệu số 60/2024/QH15, hiệu lực từ ngày 01 tháng 7 năm 2025, lần đầu đặt một khung riêng cho dữ liệu số và cơ sở dữ liệu tổng hợp quốc gia. Luật Bảo vệ dữ liệu cá nhân số 91/2025/QH15, hiệu lực từ ngày 01 tháng 01 năm 2026 cùng với Nghị định số 356/2025/NĐ-CP hướng dẫn, xác lập quyền xem, chỉnh sửa và yêu cầu xóa dữ liệu cá nhân, đồng thời thay thế cho Nghị định số 13/2023/NĐ-CP. Luật Trí tuệ nhân tạo số 134/2025/QH15, hiệu lực từ ngày 01 tháng 3 năm 2026, yêu cầu hệ thống AI tương tác trực tiếp với con người phải để người dùng biết mình đang làm việc với máy, đặt nguyên tắc quản lý tương xứng với mức độ rủi ro và yêu cầu dấu nhận biết đối với nội dung do AI tạo ra. Luật Báo chí số 126/2025/QH15, hiệu lực từ ngày 01 tháng 7 năm 2026, bổ sung nghĩa vụ phối hợp gỡ bỏ và cải chính trên các nền tảng có sử dụng lại nội dung, tức là mở rộng việc cải chính xuống chuỗi lan truyền. Việc ban hành bốn văn bản này đã giải quyết nhiều khoảng trống pháp lý trước đó, nhưng vẫn chưa thể lấp hết khoảng trống của một số nội dung đề cập trong chương này. Chủ thể được bảo vệ trong ba văn bản đầu là cá nhân, doanh nghiệp và người dùng, còn trong văn bản thứ tư là cơ quan báo chí. Chưa văn bản nào trao cho một địa phương quyền yêu cầu hiệu chỉnh cách máy đang mô tả thực thể của mình, cũng chưa

quy định đầu mối tiếp nhận và thời hạn xử lý cho loại yêu cầu ấy. Nghĩa vụ phối hợp cải chính xuống nền tảng trong Luật Báo chí là tiền lệ gần nhất và đáng để tham chiếu khi thiết kế cơ chế cho tri thức về địa phương.

### 9.7.1. Khiếu nại trang web không đủ

Trong kỷ nguyên internet truyền thống, phần lớn tranh chấp thông tin đều xoay quanh một nội dung cụ thể trên một trang web cụ thể. Nếu địa chỉ, số điện thoại hay mô tả của một điểm đến bị sai, địa phương có thể liên hệ với đơn vị quản trị trang web để yêu cầu chỉnh sửa. Khi nội dung được cập nhật, người dùng truy cập sau đó sẽ nhìn thấy phiên bản mới.

Nhưng trong nền kinh tế tác nhân, trang web chỉ còn là một mắt xích trong chuỗi hình thành tri thức của máy. Một tác nhân AI có thể đồng thời tham chiếu dữ liệu từ đồ thị tri thức, dữ liệu mở của chính phủ, OTA, bản đồ số, mạng xã hội, tài liệu PDF, kho dữ liệu doanh nghiệp và nhiều nguồn đã được mô hình học từ trước. Vì vậy, sửa đúng một trang web không đồng nghĩa với việc sửa được tri thức mà AI đang sử dụng.

Thậm chí, nhiều mô hình AI còn vận hành trên các bản sao tri thức được huấn luyện từ nhiều tháng hoặc nhiều năm trước. Trang web đã được cập nhật nhưng tri thức trong mô hình vẫn giữ phiên bản cũ cho đến khi có chu kỳ thu thập dữ liệu và huấn luyện mới. Điều này tạo ra một khoảng trễ mà địa phương gần như không thể quan sát trực tiếp.

Quan trọng hơn, cùng một thông tin sai có thể đã được sao chép và lan truyền qua hàng chục hệ thống khác nhau. Một

trang web sai tạo ra nhiều bản sao sai, các bản sao đó lại trở thành nguồn dữ liệu cho những hệ thống AI tiếp theo. Khi ấy, việc sửa trang web chỉ giống như khóa một vòi nước trong khi dòng nước đã chảy vào cả mạng lưới.

### 9.7.2. Khiếu nại AI khác khiếu nại báo chí

Khi một tờ báo đăng sai, trách nhiệm thường khá rõ ràng. Có tác giả, biên tập viên, tổng biên tập và cơ quan chủ quản. Người bị ảnh hưởng biết phải gửi yêu cầu đính chính đến đâu và có thể viện dẫn những quy định pháp lý hiện hành về cải chính thông tin.

Đối với AI, chuỗi trách nhiệm phức tạp hơn rất nhiều. Một câu trả lời sai có thể không xuất phát từ một bài viết cụ thể mà là kết quả của quá trình tổng hợp từ hàng trăm nguồn dữ liệu khác nhau. Sai lệch có thể phát sinh từ dữ liệu gốc, từ đồ thị tri thức, từ thuật toán truy xuất, từ mô hình suy luận hoặc từ cách hệ thống kết hợp nhiều nguồn thông tin. Vì vậy, rất khó xác định chính xác "ai là người tạo ra lỗi".

Điều này khiến cơ chế khiếu nại cũng phải thay đổi. Địa phương không chỉ cần yêu cầu gỡ bỏ một nội dung, mà cần yêu cầu cập nhật thực thể, sửa quan hệ dữ liệu, bổ sung nguồn xác thực và bảo đảm rằng các hệ thống AI sẽ ưu tiên sử dụng phiên bản đã được hiệu chỉnh trong những lần suy luận tiếp theo.

Một khác biệt quan trọng nữa là báo chí chủ yếu truyền tải thông tin đến con người, còn AI sử dụng thông tin để đưa ra quyết định. Một bài báo sai có thể gây hiểu nhầm, một tác nhân AI hiểu sai có thể không giới thiệu điểm đến, không lập lịch

trình, không đặt dịch vụ hoặc không đưa địa phương vào tập lựa chọn của hàng triệu người dùng. Sai lệch vì thế không chỉ ảnh hưởng đến nhận thức mà còn trực tiếp ảnh hưởng đến dòng giao dịch và dòng doanh thu.

Do đó, cơ chế khiếu nại trong kỷ nguyên AI không thể chỉ là một phiên bản số của quyền yêu cầu cải chính báo chí. Nó phải trở thành một cơ chế quản trị tri thức máy, nơi địa phương có quyền yêu cầu rà soát, hiệu chỉnh và xác nhận lại cách các hệ thống AI đang biểu diễn mình. Đây sẽ là một thành phần quan trọng của hạ tầng quản trị số trong tương lai, tương tự như cơ chế bảo vệ tên miền, dữ liệu công hay sở hữu trí tuệ trong nền kinh tế internet.

### 9.7.3. Thiếu quy trình xác minh

Có dữ liệu đúng vẫn chưa đủ. Máy cần biết ai đã xác nhận. Một trong những điểm yếu lớn nhất của hệ sinh thái AI hiện nay không phải là khả năng tạo ra thông tin sai, mà là thiếu một quy trình chuẩn để xác minh khi có tranh chấp. Khi hai nguồn dữ liệu mâu thuẫn nhau, hệ thống thường tự đánh giá mức độ tin cậy dựa trên thuật toán, chứ chưa có một cơ chế tương tự như hoạt động kiểm chứng trong báo chí hay quy trình xác minh của cơ quan hành chính.

Đối với địa phương, đây là một khoảng trống đáng lo ngại. Nếu AI mô tả sai giờ mở cửa của một di tích, sai tên một lễ hội, nhầm đơn vị quản lý một khu bảo tồn hoặc gán nhầm một di sản cho địa phương khác, câu hỏi đầu tiên phải là: Ai có thẩm quyền xác nhận đâu là thông tin chính thức? Hiện nay, câu trả lời vẫn chưa rõ ràng.

Trong nhiều lĩnh vực, dữ liệu chính thức đã tồn tại nhưng chưa được tổ chức theo cách máy có thể ưu tiên nhận diện. Một quyết định của ủy ban nhân dân cấp tỉnh, một cơ sở dữ liệu của sở chuyên ngành hay một cổng dữ liệu mở có thể là nguồn có giá trị pháp lý cao nhất đối với con người, nhưng đối với AI, chúng có thể chỉ là một nguồn trong rất nhiều nguồn khác. Nếu không có tín hiệu rõ ràng về tính xác thực và thẩm quyền, máy khó phân biệt đâu là nguồn chính thức và đâu chỉ là bản sao được lan truyền rộng rãi.

Thiếu quy trình xác minh cũng đồng nghĩa với việc không có quy trình kết thúc tranh chấp. Địa phương có thể chứng minh thông tin của mình đúng, nhưng không biết bằng cách nào để hệ thống AI công nhận kết quả xác minh đó. Việc sửa chữa vì thế dễ rơi vào trạng thái kéo dài, mỗi nền tảng xử lý theo một cách khác nhau.

Trong nền kinh tế tác nhân, xác minh cần trở thành một phần của hạ tầng dữ liệu công. Không chỉ cần dữ liệu đúng, mà còn cần cơ chế để máy biết dữ liệu nào đã được xác minh, bởi cơ quan nào và còn hiệu lực đến khi nào. Chỉ khi đó, niềm tin của máy mới có thể được xây dựng trên những quy trình minh bạch thay vì chỉ trên mức độ phổ biến của thông tin.

### 9.7.4. Thiếu chuẩn phản hồi

Tiếp nhận là bước đầu. Lan truyền bản đã sửa mới là bước cuối. Ngay cả khi phát hiện thông tin sai và có thể chứng minh thông tin đúng, một câu hỏi khác lại xuất hiện: địa phương phải phản hồi theo chuẩn nào? Hiện nay, gần như mỗi nền tảng đều có cơ chế riêng, biểu mẫu riêng và tiêu chí riêng. Không tồn tại một chuẩn chung cho việc yêu cầu AI sửa tri thức.

Đối với AI, những thành phần đó vẫn còn thiếu. Địa phương không biết cần gửi những bằng chứng nào, theo định dạng nào, đến bộ phận nào và trong bao lâu sẽ nhận được phản hồi. Nhiều trường hợp, việc sửa dữ liệu phụ thuộc vào các biểu mẫu rời rạc hoặc các kênh hỗ trợ chung, không được thiết kế riêng cho các tranh chấp về tri thức máy.

Quan trọng hơn, phản hồi không nên chỉ dừng ở việc "đã nhận yêu cầu". Một hệ thống trưởng thành cần có cả vòng đời của việc hiệu chỉnh tri thức: tiếp nhận, xác minh, cập nhật, thông báo kết quả và theo dõi xem tri thức mới đã được lan truyền sang các hệ thống liên quan hay chưa. Nếu thiếu bước cuối cùng, thông tin đúng vẫn có thể bị che khuất bởi những bản sao cũ tiếp tục lưu hành.

Trong tương lai, khiếu nại và hiệu chỉnh có thể trở thành hai chức năng tiêu chuẩn của các nền tảng AI quy mô lớn. Cũng như internet đã hình thành các chuẩn về bảo mật, định danh và trao đổi dữ liệu, nền kinh tế tác nhân sẽ cần các chuẩn về khiếu nại tri thức và hiệu chỉnh tri thức. Đó không chỉ là vấn đề kỹ thuật, mà là nền tảng để bảo đảm tính minh bạch, trách nhiệm giải trình và quyền tự biểu diễn của mỗi địa phương trong môi trường số.

### 9.7.5. Đầu mối bảo vệ quyền biểu diễn số

Khi internet phát triển, nhiều quốc gia đã thành lập các cơ quan hoặc đầu mối chuyên trách để tiếp nhận khiếu nại về dữ liệu cá nhân, nội dung số, an toàn mạng và quyền của người tiêu dùng trên môi trường trực tuyến. Trong nền kinh tế tác nhân, một nhu cầu tương tự đang xuất hiện, cần có một đầu mối chịu trách nhiệm đối với các tranh chấp liên quan đến tri thức mà máy đang sử dụng.

Điều cần thiết ở đây là một chức năng, không nhất thiết là một bộ máy. Có thể hình dung ba cấp độ thực hiện, tăng dần theo quy mô và theo mức chín của pháp luật. Cấp một là một chức năng được giao thêm cho đơn vị đang quản lý dữ liệu, không thêm biên chế. Cấp hai là một đầu mối có tên gọi và thẩm quyền rõ ràng, đủ để làm việc với các nền tảng thay cho toàn bộ địa phương. Cấp ba là một thiết chế độc lập, chỉ lập khi khối lượng vụ việc và khuôn khổ pháp luật đòi hỏi. Trong sách này, tên gọi đầu mối bảo vệ quyền biểu diễn số (digital ombudsman) chỉ dùng để đặt tên cho chức năng, không hàm ý phải sinh thêm cơ quan.

Vai trò đầu tiên của đầu mối bảo vệ quyền biểu diễn số là tiếp nhận. Thay vì mỗi địa phương phải tự tìm cách liên hệ với từng nền tảng, từng mô hình hay từng nhà cung cấp dữ liệu, mọi yêu cầu có thể được gửi tới một đầu mối thống nhất. Cơ quan này sẽ kiểm tra tính đầy đủ của hồ sơ, xác định loại tranh chấp và hướng dẫn quy trình xử lý phù hợp.

Vai trò thứ hai là điều phối. Một thông tin sai thường không chỉ tồn tại ở một nơi mà lan truyền qua nhiều lớp dữ liệu và nhiều nền tảng. Đầu mối bảo vệ quyền biểu diễn số có thể đóng vai trò kết nối giữa chính quyền địa phương, cơ quan dữ liệu quốc gia, doanh nghiệp công nghệ và các nền tảng AI để việc hiệu chỉnh được thực hiện đồng bộ thay vì rời rạc. Điều này giúp giảm tình trạng "sửa ở nơi này nhưng sai vẫn tồn tại ở nơi khác".

Vai trò thứ ba là chuẩn hóa. Hiện nay, mỗi nền tảng sử dụng một biểu mẫu và quy trình khác nhau. Một đầu mối bảo vệ quyền biểu diễn số có thể xây dựng bộ tiêu chuẩn chung cho hồ sơ yêu cầu hiệu chỉnh tri thức: danh tính thực thể, bằng chứng pháp lý, nguồn dữ liệu chính thức, thời điểm có hiệu lực, mức độ ảnh hưởng và kết quả mong muốn. Khi mọi bên cùng sử dụng một

ngôn ngữ và một quy trình thống nhất, chi phí xử lý sẽ giảm và tính minh bạch sẽ tăng lên.

Quan trọng hơn, cơ quan này còn có thể công bố các chỉ số trách nhiệm của hệ sinh thái AI: số lượng yêu cầu được tiếp nhận, thời gian xử lý trung bình, tỷ lệ hiệu chỉnh thành công, các loại lỗi phổ biến và mức độ tuân thủ của từng nền tảng. Những chỉ số đó không nhằm tạo áp lực hành chính, mà nhằm xây dựng niềm tin và trách nhiệm giải trình trong một môi trường mà ngày càng nhiều quyết định được giao cho máy.

Trong dài hạn, khi AI trở thành hạ tầng thiết yếu của nền kinh tế số, một cơ chế như vậy có thể sẽ tự nhiên như sự tồn tại của cơ quan bảo vệ dữ liệu cá nhân ngày nay. Nếu quản trị tri thức là nguyên tắc thì đây là chỗ nguyên tắc ấy được thực thi. Đó cũng là ranh giới giữa hai mục: mục này bàn về hạ tầng khiếu nại hướng ra hệ sinh thái bên ngoài, còn mục tiếp theo bàn về năng lực mà chính quyền phải xây dựng bên trong bộ máy của mình.

Một đầu mối bảo vệ quyền biểu diễn số sẽ nhanh chóng trở thành cơ quan kiểm duyệt nội dung AI. Ranh giới giữa dữ kiện và ý kiến là ranh giới quyết định, nên nó cần được ghi vào thẩm quyền của thiết chế ngay từ đầu. Đầu mối này chỉ xử lý tranh chấp về những dữ kiện kiểm chứng được của thực thể: tên gọi chính thức, ranh giới, đơn vị quản lý, giờ mở cửa, quan hệ giữa các thực thể, tình trạng vận hành. Nó không xử lý ý kiến, đánh giá, xếp hạng hay quan điểm, không có quyền yêu cầu gỡ nội dung và công khai toàn bộ yêu cầu cùng kết quả xử lý. Một thiết chế không được giới hạn như vậy thì đúng là sẽ trượt sang kiểm duyệt. Một quyền không có kênh thực thi thì chỉ là một tuyên bố.

## 9.8. NĂNG LỰC QUẢN TRỊ NỘI BỘ CỦA ĐỊA PHƯƠNG

Chủ quyền thuật toán không được xây dựng bằng khẩu hiệu mà bằng năng lực quản trị thường xuyên. Trong các chương trước, cuốn sách đã đi từ câu hỏi AI đang nhìn địa phương như thế nào, đến ai chịu trách nhiệm về tri thức mà AI sử dụng và cuối cùng là chủ quyền của địa phương trong không gian số. Tuy nhiên, nhận diện đúng vấn đề mới chỉ là bước khởi đầu. Câu hỏi quan trọng hơn là: chính quyền địa phương cần làm gì để biến chủ quyền thuật toán thành năng lực quản trị thực tế?

Khác với những chiến dịch truyền thông có điểm bắt đầu và điểm kết thúc, chủ quyền thuật toán là một quá trình vận hành liên tục. Dữ liệu phải được cập nhật hằng ngày, thực thể phải được duy trì nhất quán, sai sót phải được phát hiện và sửa chữa nhanh chóng, còn các chuẩn dữ liệu phải được bảo đảm trong toàn bộ hệ sinh thái. Đây không phải là nhiệm vụ của riêng ngành công nghệ thông tin mà là một năng lực quản trị mới, đòi hỏi sự phối hợp giữa chính quyền, doanh nghiệp, nền tảng và cộng đồng.

Nếu mục trước bàn về hạ tầng khiếu nại hướng ra bên ngoài, mục này bàn về năm việc bên trong bộ máy, mỗi việc gắn với một đầu mối, một chu kỳ và một chỉ số đo. Chương trình nghị sự ấy không nên bắt đầu bằng việc mua thêm phần mềm hay triển khai thêm ứng dụng AI. Nó phải bắt đầu từ việc xây dựng các thể chế, quy trình và năng lực giúp địa phương luôn hiện diện chính xác, nhất quán và đáng tin cậy trong mọi hệ thống trí tuệ nhân tạo. Chỉ khi đó, chủ quyền thuật toán mới trở thành một năng lực bền vững thay vì một khẩu hiệu của chuyển đổi số.

Hạng mục ngân sách thường xuyên. Trong khoảng một thập niên, chi phí duy trì hiện diện số của địa phương sẽ chuyển từ chi

phí dự án sang hạng mục ngân sách thường xuyên và tỷ trọng của hạng mục này trong ngân sách xúc tiến sẽ tăng nhanh hơn tổng ngân sách. Có thể kiểm chứng bằng cách theo dõi cấu trúc chi của các cơ quan xúc tiến địa phương qua các năm, thay vì dựa vào tuyên bố chiến lược.

Từ ngày 01 tháng 7 năm 2025, chính quyền địa phương ở Việt Nam vận hành theo mô hình hai cấp gồm cấp tỉnh và cấp xã. Chương 8 đã lập luận rằng mô hình này làm thay đổi cách tổ chức dữ liệu về địa phương: quan sát thực địa thuộc cấp xã, còn hợp nhất, xác nhận và công bố thuộc cấp tỉnh. Ba việc trong chương này chịu ảnh hưởng trực tiếp. Thực thể chuẩn phải được xây ở cấp tỉnh, vì chỉ cấp tỉnh có đủ phạm vi để hợp nhất các biến thể tên. Việc phát hiện sai lệch phải bắt đầu từ cấp xã, vì đó là nơi biết một điểm đến đã dừng hoạt động hay chưa. Việc làm việc với nền tảng và nhà phát triển mô hình nên tập trung ở cấp tỉnh, vì không nền tảng nào duy trì được quan hệ song phương với hàng nghìn đơn vị cấp xã.

### 9.8.1. Giao trách nhiệm cho một đầu mối

Một trong những điểm yếu lớn nhất của quản trị tri thức địa phương hiện nay là mọi người đều sở hữu một phần dữ liệu nhưng không ai chịu trách nhiệm về tri thức mà AI đang sử dụng. Sở Văn hóa, Thể thao và Du lịch quản lý di sản và điểm đến, Sở Xây dựng quản lý địa chỉ và hạ tầng, chính quyền cấp xã nắm thực địa, doanh nghiệp quản lý dịch vụ, còn các nền tảng và mô hình AI lại tổng hợp thông tin từ tất cả các nguồn đó. Khi xuất hiện thông tin sai lệch, rất khó xác định đơn vị nào có trách nhiệm phát hiện, xác minh và điều phối việc sửa chữa.

Việc phải làm ở đây không phải lập thêm tổ chức mà là giao được trách nhiệm. Một đơn vị đang có, thường là đơn vị quản lý dữ liệu của tỉnh, được giao thêm bốn nhiệm vụ có thể kiểm tra được: giữ thực thể chuẩn (canonical entity), ban hành tiêu chuẩn dữ liệu dùng chung, tiếp nhận phản ánh sai lệch và làm đầu mối duy nhất khi làm việc với các nền tảng. Điểm khác biệt so với hiện trạng không nằm ở tên gọi đơn vị mà ở chỗ bốn nhiệm vụ ấy có địa chỉ, chứ không nằm rải rác ở mọi nơi và vì thế không ở đâu cả.

Đơn vị này không quản lý từng dữ liệu cụ thể mà quản lý tính nhất quán của toàn bộ hệ sinh thái tri thức. Một nguồn chuẩn không đồng nghĩa với một kho dữ liệu tập trung: mục tiêu là mọi hệ thống đều mô tả cùng một địa phương bằng cùng một tập thông tin đáng tin cậy, chứ không phải dồn mọi dữ liệu về một chỗ. Đầu mối là đơn vị quản lý dữ liệu của tỉnh, chu kỳ là thường xuyên, chỉ số đo là số thực thể trùng lặp và số nhiệm vụ chưa có địa chỉ.

### 9.8.2. Kiểm toán tri thức theo chu kỳ

Các địa phương đã quen với kiểm toán tài chính, kiểm toán đầu tư công hay kiểm toán an toàn thông tin. Tuy nhiên, trong nền kinh tế tác nhân, cần bổ sung một hình thức kiểm toán mới, kiểm toán tri thức. Mục tiêu của kiểm toán tri thức (knowledge audit) không phải kiểm tra dữ liệu có tồn tại hay không, mà đánh giá AI và các nền tảng số đang hiểu địa phương như thế nào.

Một cuộc kiểm toán tri thức định kỳ cần trả lời những câu hỏi rất cụ thể: AI có mô tả đúng tên gọi, vị trí và bản sắc của địa phương không? Các điểm đến quan trọng có xuất hiện đầy đủ

trên các nền tảng tìm kiếm và trợ lý AI không? Giờ mở cửa, giá vé, thông tin liên hệ và khả năng đặt dịch vụ có chính xác hay đã lỗi thời? Có bao nhiêu thực thể bị trùng lặp, bị thiếu quan hệ hoặc bị gắn nhầm vào một địa phương khác? Những sai lệch nào đang được nhiều mô hình AI lặp lại?

Điểm khác biệt của kiểm toán tri thức là đối tượng được kiểm toán không chỉ là cơ sở dữ liệu nội bộ, mà còn là hình ảnh số của địa phương trong toàn bộ hệ sinh thái AI. Kết quả kiểm toán sẽ giúp chính quyền ưu tiên sửa chữa những khoảng trống có tác động lớn nhất đến khả năng được máy đọc, đề xuất và giao dịch.

Cụ thể hơn: mỗi quý, đầu mối quản trị hiện diện số lấy một tập câu hỏi chuẩn, giữ nguyên qua các kỳ để so sánh được, kiểm tra trên các trợ lý và nền tảng phổ biến, rồi ghi nhận sai lệch theo loại thực thể, mức ảnh hưởng và thời gian tồn tại. Kết quả không dừng ở một báo cáo mà phải thành danh sách vụ việc, mỗi vụ có người chịu trách nhiệm và thời hạn đóng. Đầu mối là đơn vị quản lý dữ liệu, chu kỳ là hằng quý và bắt buộc sau mỗi lần thay đổi địa giới hoặc tên gọi, chỉ số đo là tỷ lệ mô tả đúng trên bộ câu hỏi ở Phụ lục 9A.

### 9.8.3. Xác lập nguồn chuẩn

Mọi hệ thống AI đều cần một điểm tham chiếu đáng tin cậy. Nếu không tồn tại một bộ dữ liệu chuẩn (canonical dataset) được công nhận chính thức thì mỗi nền tảng sẽ tự tổng hợp thông tin từ nhiều nguồn khác nhau và tạo ra những phiên bản khác nhau của cùng một địa phương. Theo thời gian, sự khác biệt này sẽ tích lũy thành sai lệch trong tìm kiếm, gợi ý và giao dịch.

Mỗi địa phương cần xây dựng một bộ dữ liệu nguồn thống nhất cho các thực thể quan trọng như địa danh, di tích, điểm du lịch, doanh nghiệp, sự kiện, tuyến giao thông và dịch vụ công. Bộ dữ liệu này phải xác định rõ mã định danh của từng thực thể, tên chuẩn, các tên gọi khác, quan hệ với các thực thể liên quan và nguồn xác thực chính thức. Quan trọng hơn, mọi thay đổi đều phải được cập nhật tại nguồn trước khi lan tỏa sang các hệ thống khác.

Bộ dữ liệu chuẩn không nhằm thay thế dữ liệu của doanh nghiệp hay các nền tảng số. Vai trò của nó là tạo ra một nguồn tham chiếu có thẩm quyền để các hệ thống khác có điều kiện đối chiếu và đồng bộ. Một lần nữa, một nguồn chuẩn không đồng nghĩa với một kho dữ liệu tập trung. Đầu mối là đơn vị quản lý dữ liệu cùng các sở chuyên ngành, chu kỳ là cập nhật ngay khi có thay đổi và rà soát toàn bộ mỗi năm, chỉ số đo là độ phủ thực thể ưu tiên và độ tươi của các thuộc tính quan trọng. Trong nền kinh tế tác nhân, địa phương không thể kiểm soát mọi mô hình AI, nhưng hoàn toàn có thể chủ động xây dựng bộ dữ liệu chuẩn mà các mô hình đó có cơ sở để tin cậy và sử dụng.

### 9.8.4. Sửa sai theo thời hạn

Sai sót trong tri thức số là điều không thể tránh khỏi. Điều quan trọng không phải là loại bỏ hoàn toàn sai sót, mà là địa phương có khả năng phát hiện, xác minh và sửa chữa chúng nhanh đến đâu. Một hệ thống không có cơ chế sửa sai sẽ khiến thông tin sai tiếp tục được sao chép, học lại và lan truyền qua nhiều thế hệ mô hình AI.

Vòng đời sáu bước của việc hiệu chỉnh đã trình bày trên. Việc của mục này là khác, đưa vòng đời ấy vào vận hành nội bộ. Nghĩa

là ban hành nó thành một quy trình hành chính có mã theo dõi, gắn thời hạn từng bước vào quy chế phối hợp giữa các sở và cấp xã và đưa số vụ việc quá hạn vào báo cáo công tác thay vì để nó nằm trong hộp thư của một chuyên viên.

Quy trình cũng cần lưu lịch sử thay đổi và công khai trạng thái xử lý, vì chính điều đó tạo ra uy tín của một nguồn dữ liệu. Đầu mối là đơn vị quản lý dữ liệu cùng các sở chuyên ngành, chu kỳ là liên tục, chỉ số đo là thời gian trung bình từ tiếp nhận đến cập nhật nguồn chuẩn và tỷ lệ vụ việc mà bản đã sửa xuất hiện trong các hệ thống chính sau chín mươi ngày.

### 9.8.5. Đo chất lượng hiện diện số thay vì đếm lượt khách

Từ quản lý nơi chốn sang quản trị hiện diện số. Trong nhiều thập niên, quản lý du lịch của địa phương chủ yếu tập trung vào quy hoạch điểm đến, xúc tiến quảng bá, phát triển sản phẩm và quản lý hoạt động kinh doanh. Những nhiệm vụ này vẫn giữ nguyên giá trị, nhưng trong nền kinh tế tác nhân, chúng không còn đủ để bảo đảm năng lực cạnh tranh. Một địa phương có thể được quản lý rất tốt trong thế giới thực nhưng vẫn gần như "vô hình" trong thế giới mà AI đang sử dụng để đưa ra quyết định.

Phạm vi quản trị cần được mở rộng từ quản lý du lịch sang quản trị hiện diện số (digital presence governance). Đối tượng quản lý không chỉ còn là tài nguyên, doanh nghiệp hay dòng khách, mà còn bao gồm các thực thể số, dữ liệu nguồn, đồ thị tri thức, API, khả năng giao dịch bằng máy và cách địa phương được các nền tảng, công cụ tìm kiếm và mô hình AI biểu diễn. Đây là lớp hạ tầng mới quyết định liệu địa phương có được máy nhìn thấy, hiểu đúng và lựa chọn hay không.

Ở cấp độ cao hơn, mục tiêu cuối cùng của quá trình này là chủ quyền thuật toán theo đúng nghĩa hẹp đã nêu ở mục 9.2.3. Chủ quyền ở đây không phải kiểm soát thuật toán của các nền tảng hay mô hình AI, mà là làm chủ chất lượng và tính xác thực của những tri thức mà các thuật toán sử dụng để mô tả địa phương. Một địa phương không thể quyết định AI sẽ xếp hạng mình ở vị trí nào, nhưng có thể quyết định liệu AI có đang học từ một nguồn dữ liệu đầy đủ, chính xác và được cập nhật hay không.

Đó là chuyển dịch căn bản trong tư duy quản trị. Thành công không chỉ được đo bằng số lượt khách và hiệu quả các chiến dịch quảng bá mà còn bằng khả năng duy trì một hiện diện số đáng tin cậy trong toàn hệ sinh thái AI. Đầu mối là ủy ban nhân dân cấp tỉnh, chu kỳ là hằng năm, chỉ số đo là bộ chỉ báo ở Bảng 9.3 được công bố để hội đồng nhân dân và cộng đồng cùng theo dõi.

Chương trình nghị sự này vượt quá nguồn lực của phần lớn địa phương, đặc biệt sau khi bộ máy được tinh giản. Phần lớn công việc trong mục này là chuẩn hóa một lần rồi duy trì và hầu hết có thể làm một lần ở cấp tỉnh cho toàn bộ các xã thay vì để mỗi xã tự làm. Sau khi bộ máy chuyển sang hai cấp từ giữa năm 2025, cách phân công tự nhiên là cấp xã quan sát và báo cáo thay đổi thực địa, cấp tỉnh hợp nhất, xác nhận và công bố. Việc đắt nhất không phải công nghệ mà là quy trình và quy trình thì dùng chung được. Chương 8 đã lập luận theo cùng hướng này khi bàn về năng lực cấp xã.

**Bảng 9.3.** Bộ chỉ báo quản trị hiện diện số

| Nhóm | Chỉ báo | Cách đo | Tần suất |
|---|---|---|---|
| **Độ phủ** | Tỷ lệ thực thể ưu tiên có mặt đầy đủ trong bộ dữ liệu chuẩn | Đối chiếu danh mục thực thể với lược đồ bắt buộc | Hằng quý |
| **Tính nhất quán** | Số thực thể trùng lặp hoặc lệch tên giữa các hệ thống chính | So sánh chéo giữa nguồn chuẩn và các nền tảng lớn | Hằng quý |
| **Độ tươi** | Tỷ lệ thuộc tính quan trọng được cập nhật trong kỳ | Thống kê từ dấu thời gian của bộ dữ liệu chuẩn | Hằng tháng |
| **Tỷ lệ mô tả đúng** | Tỷ lệ câu trả lời đúng khi truy vấn các trợ lý phổ biến | Bộ câu hỏi kiểm toán cố định, xem danh mục cuối sách | Hằng năm |
| **Hiện diện trong đề xuất** | Tỷ lệ xuất hiện của địa phương trong tập đề xuất với các truy vấn mẫu | Bộ truy vấn mẫu giữ nguyên qua các kỳ để so sánh được | Hằng năm |
| **Tốc độ hiệu chỉnh** | Thời gian trung bình từ tiếp nhận đến cập nhật nguồn chuẩn | Nhật ký quy trình hiệu chỉnh | Hằng quý |
| **Hiệu lực hiệu chỉnh** | Tỷ lệ vụ việc mà bản đã sửa xuất hiện trong các hệ thống chính sau 90 ngày | Kiểm tra lại theo mã theo dõi | Hằng quý |
| **Khả liên thông** | Tỷ lệ dịch vụ chính có thể giao dịch được bằng máy | Kiểm kê giao diện tích hợp của các dịch vụ chính | Hằng năm |

*Ghi chú: Bộ chỉ báo này là đề xuất của tác giả, chưa phải một chuẩn được công nhận. Giá trị của nó nằm ở việc đo cùng một cách qua nhiều kỳ, không nằm ở giá trị tuyệt đối của từng con số.*

▪ ***Hành động của chính quyền cấp tỉnh:***

• *Giao một đơn vị làm đầu mối tri thức số của tỉnh, kèm thẩm quyền xác nhận thực thể chuẩn, không lập thêm cơ quan mới nếu có thể giao cho đơn vị đang quản lý dữ liệu.*

• *Công bố bộ dữ liệu chuẩn cho các nhóm thực thể quan trọng, mỗi thực thể có mã định danh ổn định, tên chuẩn, danh sách tên gọi khác và nguồn xác thực.*

• *Đưa kiểm toán tri thức vào kế hoạch công tác định kỳ, tần suất tối thiểu một lần mỗi năm và bắt buộc sau mọi lần thay đổi địa giới hoặc tên gọi.*

• *Ban hành quy trình hiệu chỉnh có thời hạn xử lý và đơn vị chịu trách nhiệm cho từng bước.*

• *Yêu cầu tiêu chuẩn mở và điều khoản khả chuyển trong mọi hợp đồng công nghệ, gồm quyền lấy lại toàn bộ dữ liệu ở định dạng đọc được bằng máy khi kết thúc hợp đồng.*

• *Công bố định kỳ bộ chỉ báo hiện diện số để hội đồng nhân dân và cộng đồng cùng theo dõi.*

▪ ***Hành động của Sở/Ngành cấp tỉnh:***

• *Là đầu mối xác thực dữ liệu thuộc lĩnh vực quản lý.*

• *Chuẩn hóa danh mục thực thể, giấy phép, sản phẩm và dịch vụ.*

• *Công bố dữ liệu máy khả đọc, có nguồn và thời điểm cập nhật.*

• *Liên thông dữ liệu ngành với tỉnh, xã và các nền tảng*

• *Theo dõi, yêu cầu sửa thông tin sai lệch trên công cụ tìm kiếm và các nền tảng.*

• *Hướng dẫn doanh nghiệp, hộ kinh doanh xây dựng hồ sơ và bằng chứng số đáng tin cậy.*

• *Bảo đảm dữ liệu có thể truy xuất, chuyển đổi.*

▪ ***Hành động của chính quyền cấp xã:***

• *Báo cáo thay đổi thực địa theo một biểu mẫu duy nhất của tỉnh: tên gọi, tình trạng vận hành, giờ mở cửa, thay đổi đơn vị quản lý, điểm mới xuất hiện hoặc dừng hoạt động.*

- *Xác nhận danh sách thực thể trên địa bàn định kỳ, đây là việc chỉ cấp xã làm được vì phải quan sát trực tiếp.*
- *Ghi nhận và chuyển lên tỉnh mọi phản ánh của người dân và doanh nghiệp về việc máy mô tả sai địa bàn.*
- *Cùng cộng đồng xác định phần tri thức văn hóa và không gian tín ngưỡng cần giữ mờ.*
- *Không tự xây hệ thống dữ liệu riêng, dùng công cụ và lược đồ chung của tỉnh.*

▪ ***Hành động của doanh nghiệp và điểm đến:***

- *Coi việc cập nhật thuộc tính vận hành là công việc hằng ngày, không phải công việc truyền thông theo mùa.*
- *Dùng đúng tên chuẩn và mã định danh của tỉnh trong mọi hồ sơ trên các nền tảng, tránh tạo thêm biến thể tên.*
- *Bảo đảm khả năng đặt dịch vụ bằng máy đối với những gì đã công bố là có thể đặt được.*
- *Rà soát định kỳ cách các trợ lý phổ biến mô tả doanh nghiệp mình, coi đây là một phần của quản trị thương hiệu.*
- *Báo cáo sai lệch phát hiện được về đầu mối của tỉnh thay vì tự liên hệ riêng từng nền tảng.*

▪ ***Hành động của nền tảng và nhà phát triển mô hình:***

- *Công khai kênh tiếp nhận yêu cầu hiệu chỉnh tri thức về thực thể, kèm thời hạn phản hồi.*
- *Phân biệt trong kết quả trả về giữa không tồn tại và không truy cập được qua nền tảng.*
- *Công khai mọi ưu tiên có tính thương mại trong tập lựa chọn.*
- *Ghi nhận và ưu tiên tín hiệu thẩm quyền từ nguồn dữ liệu chính thức của địa phương khi các nguồn mâu thuẫn.*
- *Công bố định kỳ số liệu về yêu cầu hiệu chỉnh đã nhận, thời gian xử lý và tỷ lệ hoàn tất.*

## 9.9. KẾT LUẬN CHƯƠNG

Các chương trước đã lập luận rằng trong nền kinh tế tác nhân, khả đọc trở thành điều kiện để một địa phương được máy nhìn thấy, hiểu đúng và đưa vào tập lựa chọn. Chương 9 tiếp tục mở rộng lập luận đó bằng một câu hỏi căn bản hơn: Ai có quyền định nghĩa thực thể mà AI đang đọc và ai có quyền sửa chữa khi tri thức ấy sai lệch? Nếu khả đọc quyết định cơ hội được lựa chọn, thì chủ quyền quyết định khả năng duy trì cơ hội đó theo thời gian.

Vì vậy, chủ quyền thuật toán không phải là quyền kiểm soát tuyệt đối các hệ thống AI, cũng không phải nỗ lực ngăn cản dữ liệu lưu thông. Đó là năng lực bảo đảm rằng địa phương luôn được biểu diễn một cách chính xác, nhất quán, khả liên thông và có thể được cập nhật khi thực tế thay đổi. Chủ quyền trong nền kinh tế tác nhân không nằm ở việc sở hữu mọi dữ liệu, mà ở việc duy trì được quyền định nghĩa, quyền hiệu chỉnh và quyền bảo vệ thực thể số của chính mình.

Cuối cùng, mọi khái niệm về chủ quyền đều phải được kiểm chứng bằng năng lực vận hành. Nếu ngày mai một tác nhân AI mô tả sai địa phương, dẫn du khách đến nhầm nơi, công bố thông tin lỗi thời hoặc làm tổn hại danh tiếng của cộng đồng, liệu đã tồn tại một quy trình rõ ràng để phát hiện, xác minh, khiếu nại và sửa chữa hay chưa? Nếu câu trả lời vẫn là chưa, thì chủ quyền mới chỉ tồn tại trên giấy tờ. Chỉ khi địa phương có thể liên tục duy trì và bảo vệ sự hiện diện của mình trong không gian tri thức số, chủ quyền thuật toán mới thực sự trở thành một năng lực quản trị.

Đó cũng là thông điệp xuyên suốt của ba chương khép lại cuốn sách. Chương 7 lập luận rằng địa phương phải được thiết kế để

máy có thể nhận diện, truy vấn và giao dịch với các dịch vụ của mình, Chương 8 chỉ ra Việt Nam đang vướng ở đâu và Chương 9 xác lập ai có quyền định nghĩa, ai chịu trách nhiệm và bằng cách nào quyền ấy được thực thi. Trong nền kinh tế trải nghiệm, địa phương cạnh tranh bằng tài sản, bản sắc và trải nghiệm. Nhưng trong nền kinh tế tác nhân, địa phương còn phải cạnh tranh bằng khả đọc, khả hành động và chủ quyền tri thức. Ba năng lực này không thay thế những giá trị truyền thống của nơi chốn, mà trở thành lớp hạ tầng mới giúp những giá trị ấy tiếp tục được nhận biết, được lựa chọn và được chuyển hóa thành thịnh vượng trong kỷ nguyên AI. Chủ quyền của một địa phương dừng lại đúng ở giới hạn năng lực sửa sai của nó.

## PHỤ LỤC 9A. BỘ CÂU HỎI KIỂM TOÁN TRI THỨC

Bộ câu hỏi trong Bảng 9A.1 dùng cho một lần kiểm toán tri thức về địa phương. Có thể chạy trọn bộ mỗi năm một lần, hoặc chạy từng nhóm câu hỏi theo quý.

**Bảng 9A.1.** Bộ câu hỏi kiểm toán tri thức về địa phương

| Nhóm | Câu hỏi kiểm toán | Nguồn kiểm chứng |
|---|---|---|
| **Danh tính** | Máy có gọi đúng tên chuẩn hiện hành của địa phương và các đơn vị trực thuộc hay không | Văn bản về địa giới và tên gọi hiện hành |
| **Danh tính** | Có bao nhiêu biến thể tên đang tồn tại và biến thể nào đang được máy ưu tiên | Truy vấn nhiều trợ lý với cùng một câu hỏi |
| **Danh tính** | Tên gọi trước sắp xếp đơn vị hành chính có còn được dùng như thực thể riêng hay không | Bộ dữ liệu chuẩn và bản đồ số |
| **Ranh giới** | Máy có gán đúng thực thể vào đơn vị hành chính cấp trên hay không | Bản đồ hành chính hiện hành |
| **Thuộc tính** | Giờ mở cửa, giá vé, thông tin liên hệ của điểm đến chính có chính xác hay không | Nguồn của đơn vị quản lý trực tiếp |
| **Thuộc tính** | Thông tin về khả năng tiếp cận cho người cao tuổi và người khuyết tật có tồn tại hay không | Khảo sát thực địa của cấp xã |
| **Quan hệ** | Di sản, lễ hội và đặc sản có bị gán cho địa phương khác hay không | Hồ sơ di sản và văn bản công nhận |
| **Quan hệ** | Các điểm đến quan trọng có liên kết với thực thể địa phương hay đứng rời rạc | Đồ thị tri thức của các nền tảng chính |
| **Tình trạng** | Có điểm đến nào đã dừng hoạt động mà máy vẫn giới thiệu hay không | Báo cáo của cấp xã |
| **Tình trạng** | Có dự án chưa hoàn thành nào bị máy mô tả như đang vận hành hay không | Danh mục dự án của tỉnh |
| **Thẩm quyền** | Khi các nguồn mâu thuẫn, máy đang dẫn lại nguồn nào | Trích dẫn trong câu trả lời của trợ lý |
| **Thẩm quyền** | Nguồn chính thức của tỉnh có được máy dẫn lại hay không | Truy vấn có yêu cầu nêu nguồn |

***Ghi chú.*** *Bộ câu hỏi này nên được giữ nguyên qua các kỳ kiểm toán để kết quả so sánh được. Khi cần bổ sung câu hỏi mới, nên thêm vào cuối thay vì sửa câu hỏi cũ.*

# DANH MỤC THUẬT NGỮ

Danh mục này gồm hai nhóm. Nhóm thứ nhất là những thuật ngữ đã có trong học thuật hoặc kỹ thuật, chẳng hạn xuất xứ, phân giải thực thể, đồ thị tri thức, khả liên thông, khả chuyển dữ liệu và chi phí giao dịch. Nhóm thứ hai là những khái niệm hoặc cách dùng được phát triển trong cuốn sách này, đánh dấu bằng ký hiệu ♦ ở cột tiếng Việt. Với nhóm thứ hai, phần nghĩa nêu định nghĩa một câu, phạm vi sử dụng và điểm khác biệt so với khái niệm gần nhất.

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Bản thể luận | Ontology | Tập các loại thực thể và quan hệ giữa chúng |
| Bảng ánh xạ tên | Name mapping table | Bảng đối chiếu mọi tên cũ, tên viết tắt và biến thể chính tả về một mã định danh chuẩn duy nhất. |
| Bảng tên thay thế | Alias table | Danh sách mọi cách gọi khác trỏ về thực thể chuẩn, kèm ngữ cảnh và thời hiệu |
| Bánh đà dữ liệu | Data flywheel | Vòng lặp trong đó mỗi giao dịch vừa tạo doanh thu vừa tạo dữ liệu để cải thiện giao dịch tiếp theo |
| ♦ Bộ lọc khả hành động | Actionability filter | Bộ lọc thứ tư của mô hình bốn bộ lọc, hỏi máy có hoàn thành được nhiệm vụ hay không |
| ♦ Bốn bộ lọc | Four filters | Mô hình do sách này đề xuất: phân giải, chứng thực, xếp hạng và thực thi. Một địa phương phải đi qua cả bốn mới nằm trong tập lựa chọn của máy. |
| ♦ Cái van | The valve | Quyền giữ mờ: phần tri thức mà chủ thể chủ động không đưa vào lược đồ, đặt ở cuối chuỗi bốn bộ lọc như một cơ chế điều tiết. |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| ♦ Cái vòng | The loop | Vòng phản hồi trong đó kết quả của một lần được máy đọc trở thành dữ liệu đầu vào cho lần sau, làm lợi thế và bất lợi tự khuếch đại. |
| Cam kết có hiệu lực | Binding commitment | Kết quả mà máy nhận lại sau một lời gọi giao dịch, gồm giữ chỗ, xác nhận, chủ thể chịu trách nhiệm và cơ chế xử lý thất bại |
| Cân bằng Nash | Nash equilibrium | Trạng thái không chủ thể nào có lợi khi đơn phương thay đổi chiến lược. |
| Chi phí cạnh tranh phân phối | Distribution competition cost | Khoản chi để tăng khả năng hiển thị, thứ hạng hoặc tỷ lệ chuyển đổi sau khi đã đáp ứng điều kiện khả đọc và khả hành động |
| ♦ Chi phí khả hành động | Actionability cost | Chi phí để máy truy vấn tình trạng, kiểm tra tồn kho, đặt dịch vụ, xác nhận và hoàn tất giao dịch |
| ♦ Chi phí khả đọc | Legibility cost | Chi phí để máy nhận diện, phân giải thực thể và hiểu đúng địa phương |
| ♦ Chi phí tham gia nền kinh tế tác nhân | Participation cost | Khái niệm bao trùm ba lớp dưới đây, dùng khi nói về tổng chi phí một địa phương phải bỏ ra để có mặt và giao dịch được trong nền kinh tế tác nhân |
| Chỉ số dẫn dắt | Leading indicator | Chỉ số biến động trước kết quả kinh doanh |
| Chỉ số đi sau | Lagging indicator | Chỉ số chỉ phản ánh kết quả sau khi đã xảy ra |
| Chủ quyền dữ liệu | Data sovereignty | Quyền kiểm soát việc dữ liệu về một chủ thể được lưu ở đâu, do ai giữ và theo pháp luật nào. |
| Chủ quyền thuật toán | Algorithmic sovereignty | Khả năng của một địa phương giữ quyền quyết định cách mình được biểu diễn, suy luận và đề xuất trong các hệ thống thuật toán. |
| Chuẩn hóa Unicode | Unicode normalization | Quy mọi cách biểu diễn một chữ có dấu về cùng một dạng trước khi so sánh |
| Chuẩn mã hóa ký tự | Character encoding | Quy tắc chuyển ký tự thành chuỗi byte, chuyển đổi giữa các chuẩn gây mất dấu |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| ♦ Chuỗi giữ giá trị | Value capture chain | Cách nhìn chuỗi theo câu hỏi ai giữ lại bao nhiêu, thay vì theo câu hỏi ai tạo ra bao nhiêu |
| ♦ Chuỗi trích giá trị | Value extraction chain | Trường hợp giới hạn của chuỗi trên, khi một chủ thể rút giá trị mà không đóng góp phần tương ứng vào trải nghiệm |
| Có thể hành động | Actionable | Cấp độ trưởng thành thứ tư, khi máy làm thay đổi được trạng thái của thế giới thật chứ không chỉ mô tả nó |
| Giải quyết xung đột | Conflict Resolution | Cơ chế chọn một phiên bản khi các nguồn mâu thuẫn |
| Giao diện lập trình | API | Cách thức chuẩn để một hệ thống truy vấn hoặc giao dịch với hệ thống khác. |
| Hạ tầng khiếu nại | Redress infrastructure | Quy trình và đầu mối để một địa phương yêu cầu hệ sinh thái AI hiệu chỉnh một mô tả sai về mình. |
| Hành vi bầy đàn | Herd behavior | Xu hướng các chủ thể cùng đi đến một lựa chọn giống nhau. |
| ♦ Hiệu chỉnh sớm | Early Correction | Sửa sai khi phạm vi lan truyền còn nhỏ |
| Hiệu ứng Matthew | Matthew effect | Cơ chế trong đó lợi thế sẵn có tự sinh thêm lợi thế. |
| Hợp nhất thực thể | Entity consolidation | Ghép các bản ghi sau khi đã phân giải, kèm nối lịch sử tiền nhiệm vào kế nhiệm |
| Hợp nhất tri thức | Knowledge Fusion | Ghép nhiều mảnh bằng chứng thành một mô tả duy nhất |
| Hợp tác xã dữ liệu | Data cooperative | Mô hình nhiều cơ sở cùng góp dữ liệu vào một hạ tầng chung trong khi vẫn giữ quyền với dữ liệu gốc của mình |
| Khả chuyển | Portability | Khả năng mang dữ liệu sang một hệ thống khác mà không mất cấu trúc |
| Khả dụng | Availability | Tình trạng còn hay hết của một đơn vị tồn kho tại thời điểm truy vấn |
| Khả giao dịch bằng máy | Machine-transactable | Khả năng máy hoàn tất một giao dịch kinh tế. Hẹp hơn khả hành động, rộng hơn khả đặt bằng máy |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Khả hành động | Actionability | Mức độ mà một tác nhân có thể hoàn tất một chuỗi hành động có hiệu lực trên tài sản, dịch vụ hoặc hệ thống của địa phương. Đây là khái niệm rộng nhất trong ba khái niệm cùng họ |
| Khả đặt bằng máy | Machine-bookable | Trường hợp cụ thể của khả giao dịch bằng máy, áp dụng cho tồn kho theo thời gian: máy giữ chỗ, xác nhận và thanh toán được |
| Khả đọc bằng máy | Machine-readable | Trạng thái dữ liệu cho phép máy phân giải và hiểu |
| ♦ Khả đọc thuật toán | Algorithmic legibility | Mức độ tài sản và năng lực của địa phương có thể được máy phân giải, chứng thực, xếp hạng và hành động lên. |
| Khóa chặt | Lock-in | Trạng thái mô tả cũ tự duy trì dù thực tại đã đổi |
| ♦ Khoảng cách bản sắc | Identity gap | Chênh lệch giữa bản sắc mong muốn và bản sắc mà máy đang gán |
| Khởi đầu lạnh | Cold start | Tình trạng hệ thống thiếu dữ liệu để đánh giá một thực thể mới. |
| ♦ Ký ức thuật toán | Algorithmic memory | Phần quá khứ tiếp tục ảnh hưởng đến truy hồi, mô tả, xếp hạng và lựa chọn, phân tán trong nhiều lớp dữ liệu và hệ thống |
| Làm phẳng giao diện | Interface flattening | Hiện tượng khác biệt giữa các nơi chốn bị nén còn vài trường dữ liệu so sánh được |
| Liên kết thiếu | Missing link | Quan hệ cần thiết nhưng chưa được khai báo giữa hai thực thể |
| ♦ Liên minh dữ liệu địa phương | Local data alliance | Thoả thuận giữa chính quyền, doanh nghiệp và cộng đồng về ai giữ lớp dữ liệu nào và ai chịu trách nhiệm cập nhật. |
| Lợi nhuận nền tảng | Platform profit | Phần còn lại sau chi phí, có thể cao nhờ hiệu quả của mô hình phi tài sản mà không hàm ý quyền lực thị trường |
| Lớp bảo đảm | Assurance layer | Định danh chủ thể chịu trách nhiệm và cơ chế hoàn hủy, nằm ngoài bốn điều kiện giao dịch lõi |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Lược đồ dữ liệu | Schema | Quy định mỗi thực thể cần có những trường nào |
| Mã định danh chuẩn | Canonical ID | Điểm neo để mọi hệ thống quy chiếu về cùng một thực thể |
| Mạng lưới tham chiếu cũ | Legacy references | Toàn bộ địa chỉ web, mã định danh, liên kết và trích dẫn còn trỏ về tên cũ |
| ♦ Mất khả đọc | Loss of machine readability | Quá trình một địa phương dần biến mất khỏi tuyến suy luận của máy |
| ♦ Nền kinh tế tác nhân | Agent economy | Nền kinh tế trong đó một tỷ trọng ngày càng lớn các quyết định tìm kiếm, so sánh, lựa chọn và giao dịch được thực hiện thông qua tác nhân phần mềm thay cho con người. |
| Người thắng giành phần lớn | Winner takes most | Trạng thái thị trường trong đó một số ít chủ thể nhận phần lớn nhu cầu. |
| Nhà cung cấp phía sau | Backend provider | Vai trò của chủ thể vận hành dịch vụ thật nhưng không sở hữu giao diện và quan hệ khách hàng |
| ♦ OEM hóa địa phương | Locality OEM-isation | Trạng thái địa phương vẫn tạo ra trải nghiệm thật nhưng mất giao diện, mất quan hệ khách hàng và mất quyền định giá, trở thành nhà cung cấp phía sau. |
| Phân giải thực thể | Entity resolution | Xác định các bản ghi có trỏ tới cùng một đối tượng hay không, chưa phải bước ghép |
| Phân mảnh thực thể | Entity fragmentation | Tình trạng một địa phương bị máy nhận thành nhiều thực thể khác nhau do khác biệt về dấu thanh, viết tắt, phiên âm hoặc thay đổi địa giới. |
| Phí dịch vụ trung gian | Intermediation fee | Khoản thu tương ứng với giá trị mà trung gian thực sự tạo ra, gồm giảm chi phí tìm kiếm, gộp cầu, bảo lãnh giao dịch |
| Phụ thuộc đường dẫn | Path dependence | Thứ tự xuất hiện quyết định thứ hạng về sau |
| Quá tải du lịch | Overtourism | Tình trạng lượng khách vượt sức chịu tải của điểm đến và của cộng đồng cư dân. |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Quyền mờ | Right to opacity | Quyền của cộng đồng quyết định phần tri thức không đưa vào lược đồ, giới thiệu ở Chương 3 |
| Bản sao số | Digital twin | Bản sao dữ liệu của một địa phương, gồm lớp hành chính, vận hành và ngữ nghĩa |
| Sức chịu tải | Carrying capacity | Giới hạn số khách mà một điểm đến phục vụ được mà không suy giảm chất lượng. |
| ♦ Suy giảm khả hành động | Actionability decay | Quá trình một địa phương mất dần khả năng được máy chọn, gồm mất khả đọc, bị hiểu sai, mất khả giao dịch và dữ liệu hết hiệu lực |
| Tác nhân | Agent | Hệ thống phần mềm không chỉ trả lời mà còn thực hiện chuỗi hành động thay người dùng, gồm truy vấn, so sánh, quyết định và giao dịch. |
| Tài sản số dùng chung | Digital commons | Lớp dữ liệu được đầu tư và quản trị chung, mở cho mọi chủ thể khai thác, không thuộc sở hữu độc quyền |
| ♦ Tam giác bản sắc | Identity triangle | Ba đỉnh gồm bản sắc tự nhận, bằng chứng thực tế và bản sắc mà máy suy ra |
| Tập lựa chọn | Choice set | Nhóm phương án cuối cùng được trình bày cho người dùng hoặc được tác nhân thực thi. |
| Tập ứng viên | Candidate set | Nhóm phương án được hệ thống đưa vào vòng cân nhắc trước khi xếp hạng. |
| Thẩm quyền | Authority | Vị thế của một nguồn đối với một loại câu hỏi |
| Thất bại ở quy mô hệ thống | System-level failure | Kết cục xấu của toàn hệ thống sinh ra từ tổng các quyết định cục bộ đều hợp lý. |
| Thiên lệch theo mức độ phổ biến | Popularity bias | Xu hướng của hệ thống gợi ý ưu tiên phương án đã được nhiều người chọn. |
| Thực thể chuẩn | Canonical entity | Bản ghi trung tâm mà mọi dữ liệu về một đối tượng được gắn vào |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Thực thể tiền nhiệm và kế nhiệm | Predecessor, successor entity | Hai thực thể khác nhau nối bằng quan hệ kế thừa, không phải hai giai đoạn của một thực thể |
| Thực thể vỡ | Broken entity | Một đối tượng thật bị tách thành nhiều hồ sơ không nối được với nhau |
| Tính nhất quán | Consistency | Mức đồng bộ giữa các nguồn khi cùng mô tả một thực thể |
| ♦ Tô thuật toán | Algorithmic rent | Phần tô kinh tế phát sinh từ quyền kiểm soát điểm nghẽn trong quá trình đọc, xếp hạng, lựa chọn hoặc giao dịch của hệ thống thuật toán. Đây là thuật ngữ phân tích do cuốn sách này sử dụng, không phải một loại tô đã có định nghĩa pháp lý riêng |
| Tối ưu cục bộ | Local optimization | Việc tối ưu cho một người dùng tại một thời điểm mà không xét tổng thể hệ thống. |
| Tồn kho giao dịch được | Tradable inventory | Đơn vị năng lực phục vụ có định danh, có trạng thái thời gian và có thể giao dịch |
| Ủy nhiệm | Delegation | Việc con người chuyển một phần hoặc toàn bộ quyền ra quyết định cho tác nhân. |
| Vòng lặp phản hồi | Feedback loop | Cơ chế trong đó kết quả của một lần xếp hạng trở thành dữ liệu đầu vào của lần sau. |
| Xác minh chéo giữa các nguồn | Source corroboration | So sánh cùng một mệnh đề trên nhiều nguồn độc lập |
| Xếp hạng | Ranking | Bộ lọc thứ ba trong mô hình bốn bộ lọc, quyết định thứ tự ưu tiên của các phương án đã được chứng thực. |
| Xuất xứ | Provenance | Lý lịch của một dữ kiện: ai công bố, khi nào, dựa trên gì |
| Đa dạng hóa đề xuất | Recommendation diversification | Nguyên lý chủ động chèn phương án ngoài nhóm phổ biến nhất vào kết quả gợi ý. |
| Đầu mối chịu trách nhiệm dữ liệu | Data stewardship | Đơn vị chịu trách nhiệm về một trường dữ liệu, không nhất thiết là chủ sở hữu pháp lý |

| Tiếng Việt | Tiếng Anh | Nghĩa dùng trong sách |
|---|---|---|
| Điểm kiểm soát | Control point | Vị trí trong chuỗi mà phần lớn dòng quyết định hoặc dòng giao dịch phải đi qua |
| ♦ Điểm đến lắp ghép được | Composable destination | Điểm đến mà mọi dịch vụ đều có thể được máy lắp ghép thành hành trình |
| Định giá tắc nghẽn | Congestion pricing | Công cụ điều tiết cầu bằng cách tính phí vào thời điểm hoặc khu vực đông. |
| ♦ Cơ chế tái cấu trúc | Mechanism of restructuring | Phát biểu của sách: khi một bộ máy có quyền phân bổ bắt đầu đọc thế giới bằng một lược đồ, thế giới tự tổ chức lại theo lược đồ ấy, kể cả khi lược đồ mô tả sai. |
| Độ lệch trạng thái | State divergence | Khoảng cách giữa trạng thái thật của thực thể và trạng thái được ghi trong dữ liệu |
| Độ tươi | Freshness | Khoảng cách từ lần cập nhật cuối tới thời điểm truy hồi |
| Độ phủ | Coverage | Tỷ lệ tài sản thật đã hiện diện trong thế giới số ở dạng đọc và giao dịch được |
| Đồ thị giao dịch | Commerce graph | Lớp trên đồ thị tri thức, gồm các tài nguyên máy có thể trực tiếp giao dịch |
| Đồ thị sống | Living graph | Đồ thị mà trạng thái của các nút thay đổi theo thời gian thực |
| Đồ thị tri thức | Knowledge graph | Cấu trúc biểu diễn tri thức bằng nút và cạnh có ngữ nghĩa |
| ♦ Đồ thị địa phương | Place graph | Mạng lưới các thực thể của địa phương cùng quan hệ giữa chúng |
| Độ trễ cập nhật | Propagation latency | Thời gian để một thay đổi ngoài đời thật xuất hiện trên các nguồn máy đọc |
| Độ trễ phản hồi | Response latency | Thời gian hệ thống trả lời một truy vấn của máy |
| Được giao việc | Delegated task | Trạng thái mà tác nhân chọn một thực thể để thực hiện một nhiệm vụ thay người dùng, gồm cả phần thực thi |

# DANH MỤC TÀI LIỆU THAM KHẢO

## I. Văn bản pháp luật và tài liệu tiếng Việt

1. Ban Chấp hành Trung ương Đảng Cộng sản Việt Nam. (2025). Nghị quyết số 60-NQ/TW ngày 12 tháng 4 năm 2025, Hội nghị lần thứ 11 Ban Chấp hành Trung ương Đảng khóa XIII, về chủ trương tổ chức chính quyền địa phương hai cấp và số lượng 34 đơn vị hành chính cấp tỉnh.

2. Bộ Chính trị Ban Chấp hành Trung ương Đảng Cộng sản Việt Nam. (2024). Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia.

3. Bộ Tài nguyên và Môi trường. (2021). *Kịch bản biến đổi khí hậu - Phiên bản cập nhật năm 2020*. Nhà xuất bản Tài nguyên, Môi trường và Bản đồ Việt Nam, Hà Nội.

4. Bộ Văn hóa, Thể thao và Du lịch. (2026). Thông cáo họp báo thường kỳ quý II năm 2026. Hà Nội, tháng 7 năm 2026.

5. Chính phủ nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Nghị định số 356/2025/NĐ-CP ngày 31 tháng 12 năm 2025 quy định chi tiết một số điều của Luật Bảo vệ dữ liệu cá nhân, hiệu lực từ ngày 1 tháng 1 năm 2026, thay thế Nghị định số 13/2023/NĐ-CP.

6. Chương trình giảng dạy kinh tế Fulbright (2013-2016). Marketing địa phương: chiến lược phát triển vùng. Tài liệu giảng dạy. Thành phố Hồ Chí Minh.

7. Cục Du lịch Quốc gia Việt Nam và Cục Thống kê. (2026). Số liệu khách quốc tế, khách nội địa và tổng thu du lịch năm 2025. Hà Nội.

8. Cục Du lịch Quốc gia Việt Nam. (2026). Báo cáo hoạt động du lịch sáu tháng đầu năm 2026. Hà Nội.

9. Cục Thống kê (2021). Kết quả Tổng điều tra dân số và nhà ở: Di cư và đô thị hoá. Hà Nội: NXB Thống kê.

10. Cục Thống kê Thâm Quyến. Số liệu chi cho nghiên cứu và phát triển của thành phố, công bố thường niên.

11. Cục Thống kê Việt Nam. Các báo cáo về nhân khẩu học và già hóa dân số. https://www.nso.gov.vn/default/2025/12/du-bao-dan-so-viet-nam-giai-doan-2024-2074/?utm_source=chatgpt.com

12. Cục Thống kê, Bộ Tài chính. (2026). Báo cáo tình hình kinh tế xã hội sáu tháng đầu năm 2026. Hà Nội.

13. Hồ Đức Hùng (2004). Marketing địa phương của Thành phố Hồ Chí Minh. Đề tài nghiên cứu khoa học, Thành phố Hồ Chí Minh.

14. Liên minh châu Âu (2016). Regulation (EU) 2016/679, General Data Protection Regulation. Điều 16 về quyền yêu cầu chỉnh sửa và Điều 20 về quyền khả chuyển dữ liệu.

15. Liên minh châu Âu (2022). Regulation (EU) 2022/1925, Digital Markets Act.

16. Liên minh châu Âu (2022). Regulation (EU) 2022/2065, Digital Services Act.

17. Liên minh châu Âu (2024). Regulation (EU) 2024/1689, Artificial Intelligence Act.

18. Ngân hàng Nhà nước Việt Nam. Số liệu kiều hối chuyển về Việt Nam và Thành phố Hồ Chí Minh, cập nhật thường niên. https://tphcm.chinhphu.vn/kieu-hoi-ve-tphcm-dat-hon-1034-ty-usd-nam-2025-101260122225011 71.htm?utm_source=chatgpt.com

19. Nguyễn Khoa Điềm (1971). Đất Nước, trích trường ca Mặt đường khát vọng.

20. Nguyễn Thành Trung (2005). Phân loại liên kết thương hiệu mở dựa trên kinh nghiệm. Tạp chí Kinh tế và Phát triển, số 95, tr. 30-33.

21. Nguyễn Thành Trung (2006). Tiến tới một khuôn khổ lý thuyết lợi thế cạnh tranh bền vững cho doanh nghiệp vừa và nhỏ: cách tiếp cận dựa trên tri thức về đoán định tương lai. Tạp chí Nghiên cứu Kinh tế, số 10, tr. 71-77.

22. Nguyễn Thành Trung (2008). Hãy hiểu đúng nghĩa thương hiệu quốc gia. Thời báo Kinh tế Sài Gòn, số 29-30.

23. Nguyễn Thành Trung (2011). Khái luận thương hiệu quốc gia: Hình ảnh và bản sắc. Tạp chí Kinh tế và Dự báo, số 19, tr. 22-24.

24. Nguyễn Thành Trung (2020). Định vị thương hiệu tỉnh Bà Rịa - Vũng Tàu về thương mại, du lịch và đầu tư. Tạp chí Kinh tế Châu Á - Thái Bình Dương, số 571, tr. 43-44.

25. Nguyễn Thành Trung (2020). Tương tác và cải thiện trong mô hình thương hiệu địa phương. Tạp chí Kinh tế và Dự báo, số 24, tr. 78-82.

26. Nguyễn Thành Trung (2024). Thương hiệu địa phương: Hình ảnh và bản sắc. Hà Nội: NXB Chính trị quốc gia Sự thật.

27. Nguyễn Thành Trung (2026). Từ quản trị thương hiệu đến tạo dựng tài sản thương hiệu. Hà Nội: NXB Công Thương.

28. Nguyễn Thành Trung và Nguyễn Quốc Thịnh (2007). Thương hiệu với nhà quản lý. Tái bản lần 2, có sửa chữa và bổ sung. Hà Nội: NXB Lao động.

29. Nguyễn Thành Trung, Đinh Ngọc Hưởng, Nguyễn Kiều Trang, Nguyễn Thu Hương và Đoàn Minh Tân Trang (2011). Chỉ số hội nhập kinh tế quốc tế cấp địa phương. Báo cáo năng lực hội nhập kinh tế quốc tế cấp địa phương. Ủy ban quốc gia về hợp tác kinh tế quốc tế.

30. Nguyễn Thành Trung, Đinh Ngọc Hưởng, Nguyễn Thu Hương, Đoàn Minh Tân Trang, Nguyễn Kiều Trang, Đỗ Quang Thành và Nguyễn Cẩm Ly (2013). Báo cáo năng lực hội nhập kinh tế quốc tế tỉnh Bà Rịa - Vũng Tàu 2013. Ủy ban nhân dân tỉnh Bà Rịa - Vũng Tàu, Ban chỉ đạo Hội nhập tỉnh Bà Rịa - Vũng Tàu.

31. Nguyễn Thành Trung, Đinh Ngọc Hưởng, Nguyễn Thu Hương, Đoàn Minh Tân Trang, Nguyễn Kiều Trang, Đỗ Quang Thành và Nguyễn Cẩm Ly (2013). Chỉ số hội nhập kinh tế quốc tế cấp địa phương. Báo cáo năng lực hội nhập kinh tế quốc tế cấp địa phương 2013. Ủy ban quốc gia về hợp tác kinh tế quốc tế.

32. Phan Minh Đức (2016). Hình ảnh điểm đến, giá trị tâm lý xã hội tác động đến sự hài lòng và lòng trung thành của du khách đến Đà Lạt. Luận án tiến sĩ, Đại học Kinh tế Thành phố Hồ Chí Minh.

33. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2016). Luật Tiếp cận thông tin số 104/2016/QH13, thông qua ngày 6 tháng 4 năm 2016, hiệu lực từ ngày 1 tháng 7 năm 2018.

34. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2024). Luật Dữ liệu số 60/2024/QH15, thông qua ngày 30 tháng 11 năm 2024, hiệu lực từ ngày 1 tháng 7 năm 2025.

35. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Luật Báo chí số 126/2025/QH15, thông qua ngày 10 tháng 12 năm 2025, hiệu lực từ ngày 1 tháng 7 năm 2026. Điều 34 về cải chính và Điều 35 về gỡ bỏ thông tin vi phạm.

36. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Luật Bảo vệ dữ liệu cá nhân số 91/2025/QH15, thông qua ngày 26 tháng 6 năm 2025 tại Kỳ họp thứ 9 Quốc hội khóa XV, hiệu lực từ ngày 1 tháng 1 năm 2026.

37. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Luật Công nghiệp công nghệ số số 71/2025/QH15, thông qua ngày 14 tháng 6 năm 2025 tại Kỳ họp thứ 9 Quốc hội khóa XV, hiệu lực từ ngày 1 tháng 1 năm 2026, riêng các điều 11, 28 và 29 hiệu lực từ ngày 1 tháng 7 năm 2025. Khoản 9 Điều 3, khoản 7 Điều 4, khoản 6 Điều 12, điểm đ khoản 2 Điều 34 và Chương IV đã bị bãi bỏ theo Điều 33 Luật Trí tuệ nhân tạo số 134/2025/QH15.

38. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Luật Tổ chức chính quyền địa phương số 72/2025/QH15, thông qua ngày 16 tháng 6 năm 2025 tại Kỳ họp thứ 9 Quốc hội khóa XV, hiệu lực từ ngày thông qua. Luật xác lập mô hình chính quyền địa phương hai cấp gồm cấp tỉnh và cấp xã.

39. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Luật Trí tuệ nhân tạo số 134/2025/QH15, thông qua ngày 10 tháng 12 năm 2025 tại Kỳ họp thứ 10 Quốc hội khóa XV, hiệu lực từ ngày 1 tháng 3 năm 2026, trừ các nội dung quy định tại Điều 35. Luật gồm 8 chương và 35 điều.

40. Quốc hội nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2025). Nghị quyết số 202/2025/QH15 ngày 12 tháng 6 năm 2025 về việc sắp xếp đơn vị hành chính cấp tỉnh, hiệu lực từ ngày ký. https://vanban.chinhphu.vn/?pageid=27160&docid=213930. Truy cập ngày 6 tháng 8 năm 2026.

41. Vũ Trí Dũng và Nguyễn Đức Hải (2011). Marketing lãnh thổ. Hà Nội: NXB Đại học Kinh tế quốc dân.

## II. Tài liệu tiếng nước ngoài

42. Abuelsaad, T., Akkil, D., Dey, P., Jagmohan, A., Vempaty, A. và Kokku, R. (2024). Agent-E: From Autonomous Web Navigation to Foundational Design Principles in Agentic Systems. Emergence AI. arXiv:2407.13032. https://arxiv.org/abs/2407.13032. Truy cập ngày 6 tháng 8 năm 2026.

43. Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics, 84(3), tr. 488-500. https://doi.org/10.2307/1879431.

44. Anderson, C. (2006). The Long Tail: Why the Future of Business Is Selling Less of More. Hyperion.

45. Anholt, S. (2007). Competitive Identity: The New Brand Management for Nations, Cities and Regions. Palgrave Macmillan.

46. Apple Inc. (2025). What happens to app prices when developers pay lower commission fees. Nghiên cứu về Đạo luật Thị trường Số, tháng 11 năm 2025.

47. Arthur, W. B. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events. The Economic Journal, 99(394), tr. 116-131. https://doi.org/10.2307/2234208.

48. Aven Hospitality và h2c. (2026). The New Search Economy: Rewriting the Rules of Direct Booking with AI. Sách trắng, công bố dưới dạng nội dung tài trợ trên Skift ngày 22 tháng 6 năm 2026. https://skift.com/2026/06/22/ai-search-economy-hospitality-white-paper/. Truy cập ngày 6 tháng 8 năm 2026.

49. Barabási, A. L. và Albert, R. (1999). Emergence of Scaling in Random Networks. Science, 286(5439), tr. 509-512.

50. Bender, E. M., Gebru, T., McMillan-Major, A. và Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, tr. 610-623.

51. Bikhchandani, S., Hirshleifer, D. và Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. Journal of Political Economy, 100(5), tr. 992-1026. https://doi.org/10.1086/261849.

52. Booking Holdings Inc. (2026). Fourth quarter and full year 2025 financial results. Norwalk, Connecticut, 18/22026. https://www.sec.gov/Archives/edgar/data/1075531/000107553126000008/q4-25bkngearningsrelease.htm.

53. Borges, J. L. (1952). El idioma analítico de John Wilkins. Trong Otras inquisiciones. Sur. Đăng lần đầu năm 1942.

54. Bowker, G. C. và Star, S. L. (1999). Sorting Things Out: Classification and Its Consequences. MIT Press.

55. Brookings Institution. (2024). Mobile app stores may need price regulation. Washington, DC.

56. Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. New Media and Society, 14(7), tr. 1164-1180.

57. Bucher, T. (2018). If...Then: Algorithmic Power and Politics. Oxford University Press.

58. Callimachus - Cyrene. Pinakes. Alexandria Library..

59. Carnovalini, F., Rodà, A. và Wiggins, G. A. (2025). Popularity Bias in Recommender Systems: The Search for Fairness in the Long Tail. Information, 16(2), bài 151. DOI 10.3390/info16020151.

60. Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L. và cộng sự (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, tập 19, số 1, bài 43, tr. 1-12. DOI 10.5334/dsj-2020-043.

61. Carroll, S. R., Herczog, E., Hudson, M., Russell, K. và Stall, S. (2021). Operationalizing the CARE and FAIR Principles for Indigenous Data Futures. Scientific Data, tập 8, bài 108. DOI 10.1038/s41597-021-00892-0.

62. Christen, P. (2012). Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Springer.

63. Cloudbeds. (2026). *A guide to OTA commission rates in 2026*. Retrieved August 6, 2026. https://www.cloudbeds.com/online-travel-agencies/commissions/

64. Coase, R. H. (1937). The Nature of the Firm. Economica, 4(16), tr. 386-405.

65. Comune di Venezia. (2026). *Regolamento e Delibere CDA (Contributo di Accesso)*. Venezia. https://www.comune.venezia.it/it/content/regolamento-e-delibere-contributo-accesso

66. Competition and Markets Authority. (2025). Strategic market status designation for mobile platforms. London.

67. Couldry, N. & Mejias, U. A. (2019). The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford University Press.

68. David, P. A. (1985). Clio and the Economics of QWERTY. American Economic Review, 75(2), tr. 332-337. https://www.jstor.org/stable/1805621.

69. Dewey, M. (1876). A Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library. Amherst, Massachusetts.

70. Diakopoulos, N. (2016). Accountability in Algorithmic Decision Making. Communications of the ACM, volume 59, no.2.

71. Digital Applied. (2026). 1,000 AI Overviews Analyzed: Citation Pattern Study. 26/4/2026. https://www.digitalapplied.com/blog/we-analyzed-1000-ai-overviews-citation-pattern-study.

72. DMG Media. (2025, 25 February). *CMA designation investigation into Google Search: Invitation to Comment – Response from DMG Media*. Submitted to the Competition and Markets Authority (CMA), London.

73. European Commission. (2024). Digital Markets Act: designation decisions and compliance obligations for gatekeepers. Brussels.

74. Expedia Group. (2026). *AI Solutions*. Expedia Group Developer Hub. https://developers.expediagroup.com/docs/ai-solutions?locale=en_US&utm_source=chatgpt.com

75. Fellegi, I. P. và Sunter, A. B. (1969). A Theory for Record Linkage. Journal of the American Statistical Association, 64(328), tr. 1183-1210. https://doi.org/10.1080/01621459.1969.10501049.

76. First Page Sage. (2026). Google Click-Through Rates (CTRs) by Ranking Position. https://firstpagesage.com/reports/google-click-through-rates-ctrs-by-ranking-position/.

77. Fishkin, R. (2026). In 2026, Less than One Third of Google Searches Still Send a Click. SparkToro. https://sparktoro.com/blog/in-2026-less-than-one-third-of-google-searches-still-send-a-click/.

78. Floridi, L. (2020). The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU. Philosophy and Technology, tập 33.

79. Foucault, M. (1966). Les mots et les choses. Gallimard. English version: The Order of Things, 1970.

80. Garfield, E. (1955). Citation Indexes for Science. Science, 122(3159), tr. 108-111.

81. Gillespie, T. (2018). Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. Yale University Press.

82. Glissant, É. (1990). Poétique de la Relation. Gallimard. English version: Glissant, Édouard. (1997). *Poetics of Relation*. Translated by Betsy Wing. Ann Arbor: University of Michigan Press.

83. Goodhart, C. A. E. (1975). Problems of Monetary Management: The UK Experience. Trong Papers in Monetary Economics, tập 1. Reserve Bank of Australia, Sydney.

84. Google Cloud. (2025). Announcing Agent Payments Protocol (AP2). https://cloud.google.com/blog/products/ai-machine-learning/announcing-agents-to-payments-ap2-protocol.

85. Google. (2026, January 11). *New tech and tools for retailers to succeed in an agentic shopping era*. Google Blog. https://blog.google/products/ads-commerce/agentic-commerce-ai-tools-protocol-retailers-platforms/

86. Hardin, G. (1968). The Tragedy of the Commons. Science, 162(3859), tr. 1243-1248.

87. Hayek, F. A. (1945). The Use of Knowledge in Society. American Economic Review, 35(4), tr. 519-530.

88. He, H. và cộng sự (2024). WebVoyager: Building an End-to-End Web Agent with Large Multimodal Models. arXiv:2401.13919.

89. Heidegger, M. (1954). Die Frage nach der Technik. Trong Vorträge und Aufsätze. Günther Neske. Enlíh version: The Question Concerning Technology, Harper and Row, 1977.

90. Holston, J. (1989). The Modernist City: An Anthropological Critique of Brasília. University of Chicago Press.

91. Lopez, P. (2026, April 30). *The Independent Hotelier's Playbook for Reducing OTA Commission*. Lighthouse.

92. ICANN (1999). Uniform Domain-Name Dispute-Resolution Policy.

93. Jacobs, J. (1961). The Death and Life of Great American Cities. Random House.

94. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A. và Fung, P. (2023). Survey of Hallucination in Natural Language Generation. ACM Computing Surveys, 55(12), paper 248, pg. 1-38. https://doi.org/10.1145/3571730.

95. Kant, I. (1781, B 1787). Kritik der reinen Vernunft. Johann Friedrich Hartknoch, Riga. English version: Critique of Pure Reason, Guyer, P. và Wood, A. W. dịch, Cambridge University Press, 1998.

96. Kenney, M. & Zysman, J. (2016). The rise of the platform economy. Issues in Science and Technology, 32(3), tr. 61-69.

97. Khan, L. M. (2017). Amazon's antitrust paradox. Yale Law Journal, 126(3), tr. 710-805.

98. Kitchin, R. (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. SAGE.

99. Kotler, P., Haider, D. & Rein, I. (1993). Marketing Places: Attracting Investment, Industry, and Tourism to Cities, States, and Nations. Free Press.

100. Koutsoupias, E. & Papadimitriou, C. (1999). Worst-case Equilibria. Trong STACS 1999. Springer.

101. Kukutai, T. & Taylor, J. (chủ biên, 2016). Indigenous Data Sovereignty: Toward an Agenda. Australian National University Press.

102. Kyoto City. (2025, March 25). *Kyoto City Assembly approves revision of the Kyoto City Accommodation Tax Ordinance*. Kyoto. Effective 1 March 2026.

103. LeCun, Y. (2022, June 27). *A Path Towards Autonomous Machine Intelligence* (Version 0.9.2). OpenReview.

104. Levinson, M. (2006). The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger. Princeton University Press.

105. Linnaeus, C. (1758). *Systema Naturae per Regna Tria Naturae* (10th ed., Vol. 1). Holmiae: Laurentius Salvius.

106. Linux Foundation. (2026). Launching the x402 Foundation with the contribution of the x402 protocol from Coinbase.

107. Marcus, G. (2020). The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence. arXiv:2002.06177.

108. Merton, R. K. (1968). The Matthew Effect in Science. Science, 159(3810), tr. 56-63. https://doi.org/10.1126/science.159.3810.56.

109. Mews. (2026). OTA commission rates: cut costs and boost direct revenue. https://www.mews.com/en/blog/minimizing-ota-hotel-commissions

110. Moreau, L. & Missier, P. (chủ biên) (2013). PROV-DM: The PROV Data Model. W3C Recommendation. https://www.w3.org/TR/2013/REC-prov-dm-20130430/.

111. Nash, J. F. (1950). Equilibrium Points in N-Person Games. Proceedings of the National Academy of Sciences, 36(1), tr. 48-49.

112. Nissenbaum, H. (2010). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.

113. OpenAI & Stripe. (2026). *Agentic Commerce Protocol (ACP) Specification*. Agentic Commerce Protocol.

114. Parker, G. G., Van Alstyne, M. W. và Choudary, S. P. (2016). Platform revolution. W. W. Norton.

115. Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.

116. Pearl, J. và Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect. Basic Books.

117. Pew Research Center. (2025). Google users are less likely to click on links when an AI summary appears in the results. Washington, DC. https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/.

118. Phocuswright. (2026). Budgets, Barriers and the Race to Agentic AI. https://www.phocuswright.com/Travel-Research/Consumer-Trends/Budgets-Barriers-and-the-Race-to-Agentic-AI.

119. Pine, B. J. và Gilmore, J. H. (1999). The experience economy: Work is theatre and every business a stage. Harvard Business School Press.

120. Plato. (1992). *Republic* (G. M. A. Grube, Trans.; C. D. C. Reeve, Rev.). Indianapolis, IN: Hackett Publishing Company.

121. Polanyi, M. (1966). The Tacit Dimension. Doubleday. Bản tái bản: University of Chicago Press, 2009.

122. Porter, M. E. (1985). Competitive advantage: Creating and sustaining superior performance. Free Press.

123. Rahman, K. S. & Thelen, K. (2019). The rise of the platform business model and the transformation of twenty-first-century capitalism. Politics and Society, 47(2), tr. 177-204.

124. Reid, L. (2025). AI in Search is driving more queries and higher quality clicks. https://blog.google/products-and-platforms/products/search/ai-search-driving-more-queries-higher-quality-clicks/.

125. Rochet, J.-C. & Tirole, J. (2003). Platform competition in two-sided markets. Journal of the European Economic Association, 1(4), tr. 990-1029.

126. Salganik, M. J., Dodds, P. S. & Watts, D. J. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. Science, 311(5762), tr. 854-856. https://doi.org/10.1126/science.1121066.

127. Scott, J. C. (1998). Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press.

128. Shapiro, C. và Varian, H. R. (1999). Information rules: A strategic guide to the network economy. Harvard Business School Press.

129. Shepard, F. (1873). Shepard's Citations. Chicago.

130. Simon, H. A. (1971). *Designing Organizations for an Information-Rich World*. In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest* (pp. 37–72). Baltimore, MD: Johns Hopkins Press.

131. SISTRIX. (2020). Why (almost) everything you knew about Google CTR is no longer valid. https://www.sistrix.com/blog/why-almost-everything-you-knew-about-google-ctr-is-no-longer-valid/.

132. Skift. (2026). Expedia to Launch Agentic AI Tools for B2B Partners.

133. Sojern. (2026). Hotel direct bookings versus OTA fees. https://www.sojern.com/blog/how-to-get-more-hotel-bookings-on-otas-and-why-you-might-not-need-to

134. Spence, M. (1973). Job Market Signaling. Quarterly Journal of Economics, 87(3), tr. 355-374.

135. Srnicek, N. (2017). Platform capitalism. Polity Press.

136. Taylor, L. (2017). What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally. Big Data and Society, volume 4, no.2.

137. Trip.com Group. (2026, March 4). *Trip.com Group Reinforces AI Innovation and European Collaboration at ITB Berlin*. PR Newswire.

138. Vickrey, W. (1969). Congestion Theory and Transport Investment. American Economic Review, 59(2), tr. 251-260.

139. Vrandečić, D. & Krötzsch, M. (2014). Wikidata: A Free Collaborative Knowledgebase. Communications of the ACM, volume 57, no.10.

140. Web Data Commons. (2025). https://webdatacommons.org/structureddata/2024-12/stats/stats.html.

141. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. và cộng sự (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. Scientific Data, volume 3, 160018. DOI 10.1038/sdata.2016.18.

142. Zhou, S. và cộng sự (2023). WebArena: A Realistic Web Environment for Building Autonomous Agents. arXiv:2307.13854.

143. Ziegler, C. N., McNee, S. M., Konstan, J. A. và Lausen, G. (2005). Improving Recommendation Lists Through Topic Diversification. Trong Proceedings of the 14th International Conference on World Wide Web, tr. 22-32. DOI 10.1145/1060745.1060754.

144. Zuboff, S. (2019). The Age of Surveillance Capitalism. Public Affairs.

**ĐỊA PHƯƠNG**
**TRONG NỀN KINH TẾ**
**TÁC NHÂN**

**Từ thương hiệu**
**đến chủ quyền thuật toán**

Chịu trách nhiệm xuất bản

GIÁM ĐỐC - TỔNG BIÊN TẬP
TRƯƠNG THU HIỀN

| | |
|---|---|
| Biên tập: | TÔN NỮ THANH BÌNH |
| Chế bản: | TRẦN HOÀNG LONG - THU THẢO |
| Trình bày bìa: | BÙI QUANG HIẾU - HÀ LAN |
| Sửa bản in: | ĐỖ THỊ TÌNH |

Đối tác liên kết: TS. NGUYỄN THÀNH TRUNG

NHÀ XUẤT BẢN CÔNG THƯƠNG
Trụ sở: Số 655 Phạm Văn Đồng, Phường Nghĩa Đô, Hà Nội
Điện thoại: 024 3 934 1562 - Fax: 024 3 938 7164
Website: http://nxbcongthuong.vn
Email: nxbct@moit.gov.vn

---

In 500 cuốn, khổ 16 x 24 cm tại Công ty TNHH In Đại Thành.
Địa chỉ: Lô S5-18, S5-19 Cụm sản xuất làng nghề tập trung
Thôn Triều Khúc, Phường Thanh Liệt, Hà Nội
Đăng ký kế hoạch xuất bản số: 3166-2026/CXBIPH/5-137/CT
Quyết định xuất bản số: 740/QĐ-NXBCT, cấp ngày 05/09/2026
Mã số ISBN: 978-632-630-955-3
In xong và nộp lưu chiểu năm 2026.

*Internet giúp con người tìm kiếm thông tin. Các tác nhân trí tuệ nhân tạo đang tiến thêm một bước: đọc, so sánh, suy luận và thay con người thực hiện ngày càng nhiều quyết định. Khi quyền lựa chọn được ủy nhiệm cho máy, một câu hỏi mới xuất hiện: một địa phương sẽ được nhìn thấy, được hiểu và được lựa chọn như thế nào trong quá trình suy luận của các tác nhân AI?*

*Trong nền kinh tế trải nghiệm, địa phương cần đủ khác biệt để con người ghi nhớ. Trong nền kinh tế tác nhân, địa phương đồng thời phải đủ khả đọc để máy hiểu. Dữ liệu phân tán, thiếu cấu trúc hay không thể liên kết có thể khiến những tài sản thực sự của một nơi không trở thành tri thức đối với máy.*

*Cuốn sách khám phá sự chuyển dịch ấy và những hệ quả đối với thương hiệu, du lịch, đầu tư, dòng khách, giá trị và năng lực cạnh tranh địa phương. Đồng thời, sách đặt ra một câu hỏi mang tính chiến lược: Làm thế nào để một địa phương vừa giữ được bản sắc trước con người, vừa hiện diện được trong trí nhớ của máy và bảo vệ chủ quyền của mình?*

ISBN: 978-632-630-955-3

**Giá: 395.000đ**