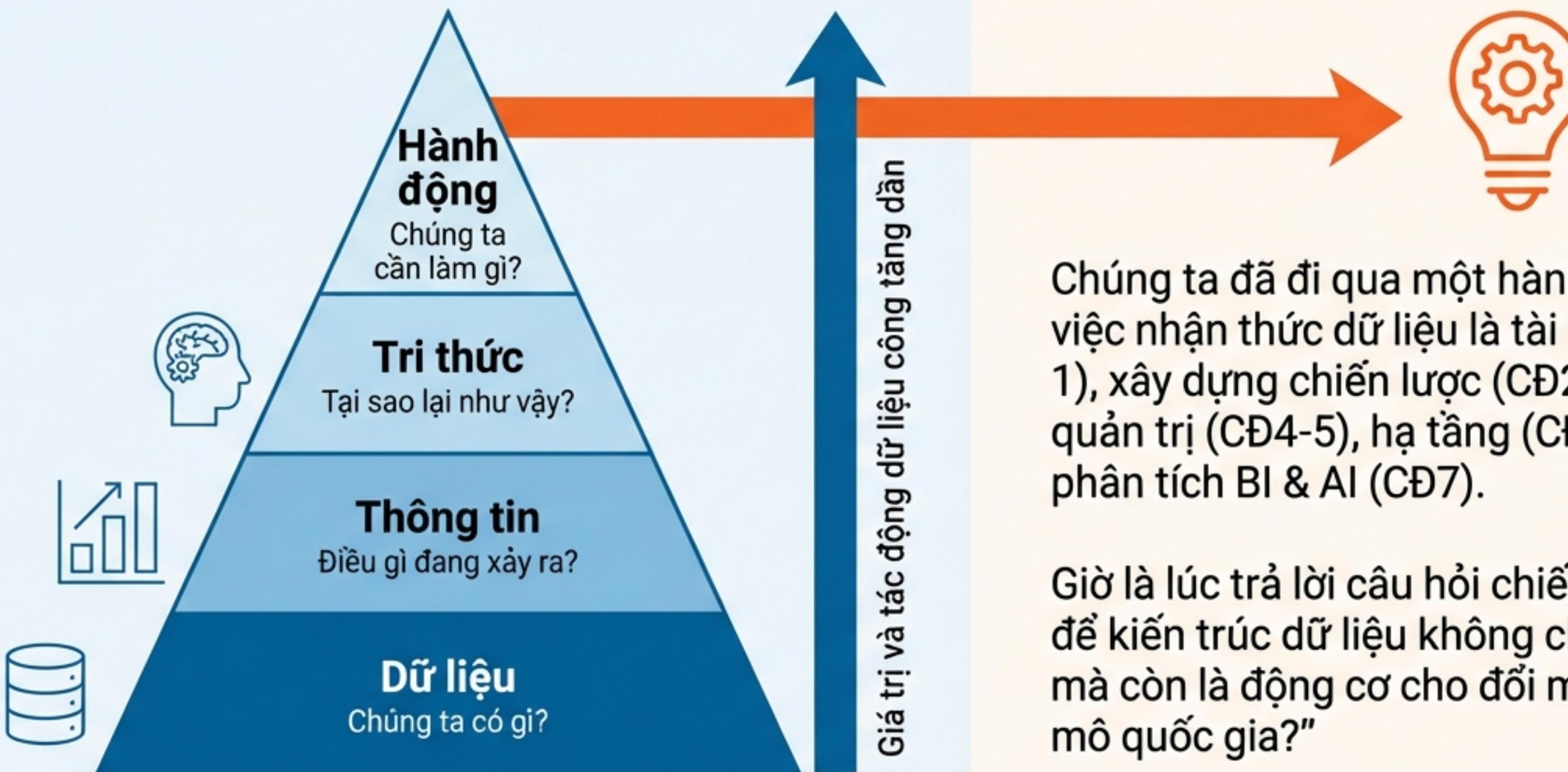


**CHUYÊN ĐỀ 8**

# **Kiến tạo Tương lai từ Dữ liệu: Kiến trúc Miền và Đổi mới Sáng tạo**

Khám phá mô hình kiến trúc giúp khai phóng tiềm năng dữ liệu, thúc đẩy đổi mới và tạo ra giá trị công vượt trội.

# Hành trình giá trị: Từ Dữ liệu thô đến Quyết định và Đổi mới



Chúng ta đã đi qua một hành trình nền tảng: từ việc nhận thức dữ liệu là tài nguyên (Chuyên đề 1), xây dựng chiến lược (CĐ2), kiến trúc (CĐ3), quản trị (CĐ4-5), hạ tầng (CĐ6), đến ứng dụng phân tích BI & AI (CĐ7).

Giờ là lúc trả lời câu hỏi chiến lược: "Làm thế nào để kiến trúc dữ liệu không chỉ hỗ trợ phân tích, mà còn là động cơ cho đổi mới sáng tạo ở quy mô quốc gia?"

Chuyển dịch từ **phân tích** dữ liệu sang **kiến tạo giá trị** từ dữ liệu.

# Thách thức của Mô hình Tập trung: Khi Trung tâm trở thành Điểm nghẽn

Các mô hình dữ liệu tập trung hoàn toàn, dù cần thiết ở giai đoạn đầu, sẽ bộc lộ những hạn chế khi quy mô và độ phức tạp tăng lên, đặc biệt là khi mục tiêu là đổi mới sáng tạo.

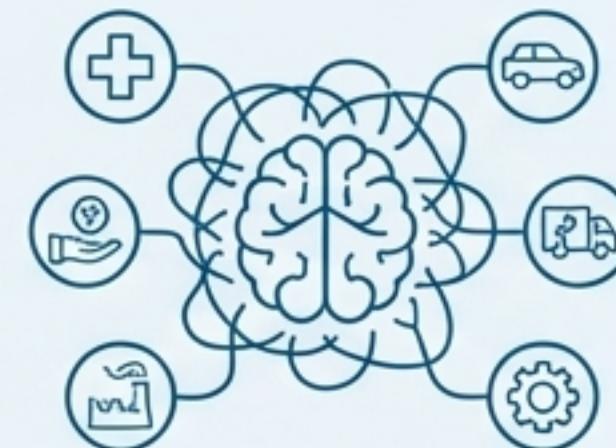
## Mô hình Tập trung (Mô hình cũ)



**Điểm nghẽn xử lý (Bottleneck):** Mọi yêu cầu dữ liệu đều phải đi qua một đội ngũ trung tâm, gây chậm trễ.

### Thiếu ngũ cảnh chuyên

**ngành:** Đội ngũ trung tâm khó có thể hiểu sâu sắc nhu cầu đặc thù của mọi lĩnh vực (y tế, giao thông, tài chính...).



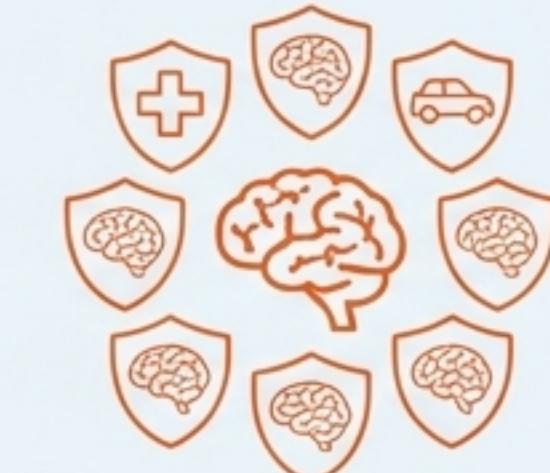
**Quản trị cứng nhắc, cản trở thử nghiệm:** Quy trình quản trị tập trung gây khó khăn cho việc triển khai các sáng kiến mới, các dự án thử nghiệm (sandbox).

## Yêu cầu cho Đổi mới (Mô hình cần hướng tới)



**Linh hoạt và Tự phục vụ (Agility & Self-Serve):** Các đơn vị chuyên ngành cần khả năng tự chủ động truy cập và khai thác dữ liệu.

**Chuyên môn hóa theo lĩnh vực:** Dữ liệu được sở hữu và quản lý bởi các chuyên gia am hiểu nhất về nó.



**Quản trị liên kết, cho phép sáng tạo:** Một khung quản trị chung cho phép các miền tự do thử nghiệm trong một hành lang an toàn.

# Mô hình Mới: Kiến trúc Miền dữ liệu (Data Domain Architecture)

Kiến trúc Miền dữ liệu là một phương pháp tiếp cận **kinh tế - kỹ thuật (socio-technical)**, tổ chức kiến trúc và quyền sở hữu dữ liệu xoay quanh các miền (domain) nghiệp vụ hoặc lĩnh vực quản lý nhà nước cụ thể (ví dụ: Miền Y tế, Miền Giao thông, Miền Tài chính).



# Bốn Nguyên tắc Trụ cột của Kiến trúc Miền Dữ liệu

Mô hình này vận hành dựa trên 4 nguyên tắc cơ bản, lấy cảm hứng từ mô hình Data Mesh và được điều chỉnh cho phù hợp với khu vực công.

## 1. Sở hữu Phân tán theo Miền (Decentralized Domain Ownership)

Mỗi miền dữ liệu (ví dụ: y tế, giáo dục) chịu trách nhiệm hoàn toàn về dữ liệu của mình, từ thu thập, chất lượng đến cung cấp. Dữ liệu được quản lý bởi những người hiểu nó nhất.

## 2. Dữ liệu là một Sản phẩm (Data as a Product)

Mỗi miền coi dữ liệu của mình không phải là một tài sản kỹ thuật, mà là một sản phẩm hoàn chỉnh, sẵn sàng để các miền khác "tiêu dùng" (sử dụng). Sản phẩm này phải tin cậy, dễ truy cập và có tài liệu rõ ràng.

## 3. Nền tảng Dữ liệu Tự phục vụ (Self-service Data Platform)

Một nền tảng công nghệ chung (như NDOP) cung cấp các công cụ để các miền có thể tự quản lý, xử lý và chia sẻ sản phẩm dữ liệu của mình một cách dễ dàng mà không cần phụ thuộc vào đội ngũ trung tâm.

## 4. Quản trị Tính toán Liên kết (Federated Computational Governance)

Một cơ chế quản trị chung được thiết lập ở cấp quốc gia (ví dụ: các tiêu chuẩn, chính sách bảo mật), nhưng việc thực thi được tự động hóa và tích hợp vào nền tảng, cho phép các miền tuân thủ một cách tự nhiên trong khi vẫn giữ được sự tự chủ.

# Nhân tố then chốt: Vai trò của Kiến trúc sư Miền dữ liệu



Để kiến trúc miền vận hành hiệu quả, cần một vai trò mới: **Kiến trúc sư Miền dữ liệu (Data Domain Architect)**. Đây là người chịu trách nhiệm chiến lược và kỹ thuật cho toàn bộ dữ liệu trong miền của mình.

- **Thiết kế Mô hình Dữ liệu của Miền:** Xây dựng và duy trì kiến trúc, mô hình dữ liệu và luồng dữ liệu cho toàn bộ miền (ví dụ: Miền Y tế).
- **Tạo ra các 'Sản phẩm Dữ liệu':** Đóng gói dữ liệu thành các sản phẩm có giá trị, dễ sử dụng, an toàn và được tài liệu hóa rõ ràng để các miền khác có thể khai thác.
- **Đảm bảo Chất lượng và Tuân thủ:** Chịu trách nhiệm về chất lượng, độ chính xác và tính tuân thủ của dữ liệu trong miền theo các quy định chung của quốc gia.
- **Cầu nối Hợp tác:** Làm việc với các Kiến trúc sư Miền khác và Ban chỉ đạo Dữ liệu Quốc gia để đảm bảo sự liên thông, tránh trùng lặp và giải quyết các vấn đề liên-miền.

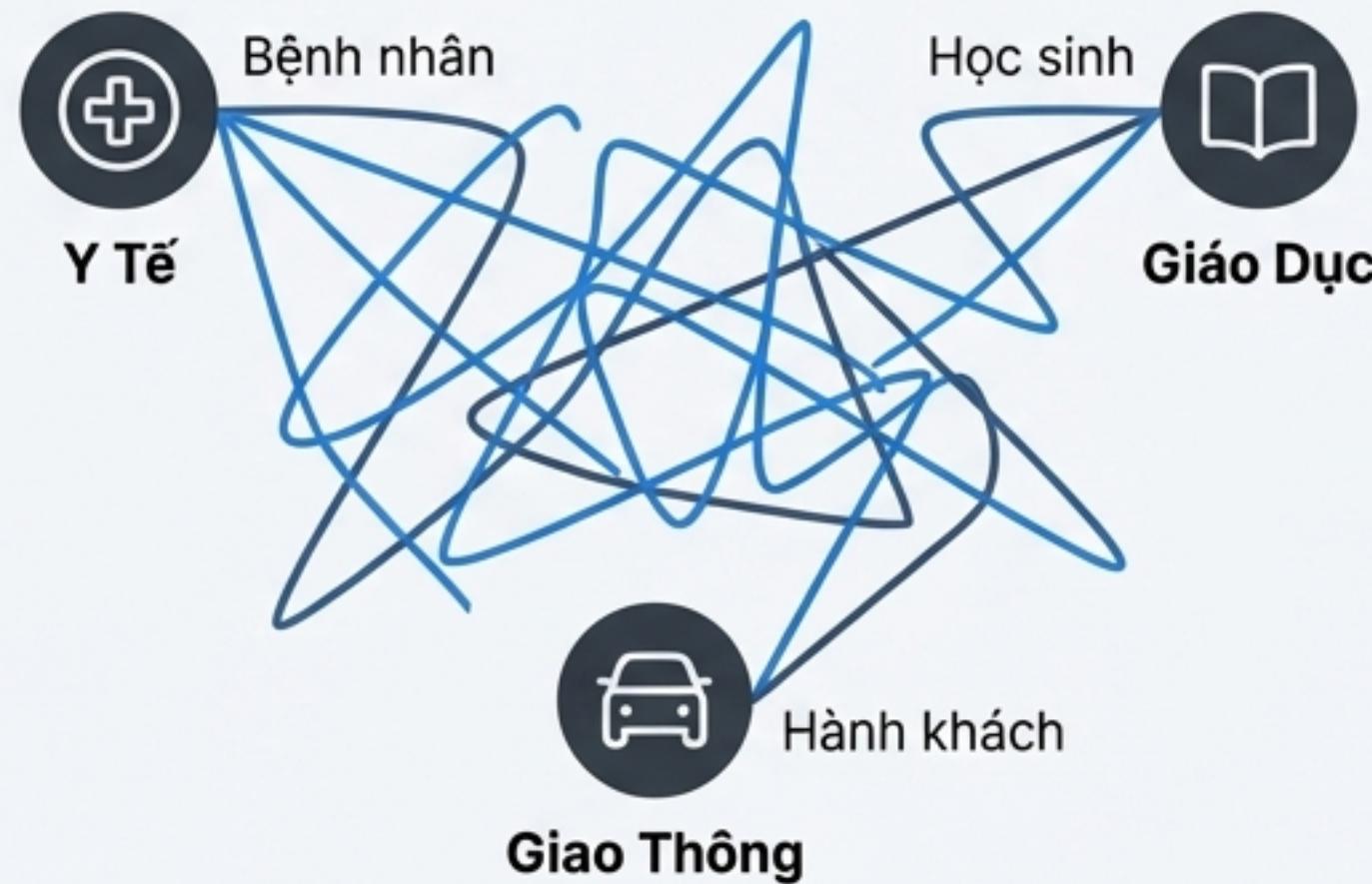
# Thách thức của Phân tán: Làm thế nào để các Miền "hiểu" nhau?

Khi mỗi miền tự chủ quản lý dữ liệu, một rủi ro lớn xuất hiện: mỗi nơi có thể định nghĩa các khái niệm theo cách riêng, dẫn đến sự hỗn loạn và không thể tích hợp.

\*Ví dụ: “Đối tượng ưu tiên” trong Miền An sinh xã hội có giống với “Đối tượng chính sách” trong Miền Y tế không?\*

## Trước

Thiếu ngôn ngữ chung



## Sau

Liên thông ngữ nghĩa



# Ontology: Không chỉ là Định nghĩa, mà là Mối quan hệ và Ngữ cảnh

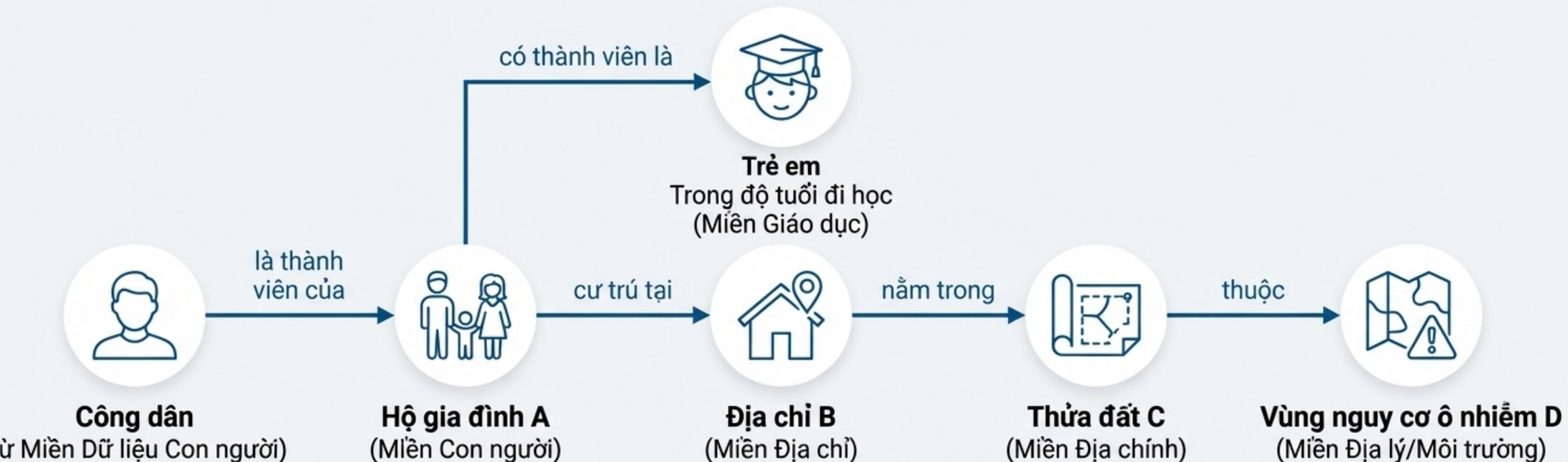
**Ontology** là một hệ thống định nghĩa chính thức và tường minh về các **khái niệm, thuộc tính** của chúng, và quan trọng nhất là **mối quan hệ** giữa các khái niệm đó trong một miền kiến thức. Nó cung cấp khả năng liên thông về mặt ngữ nghĩa (semantic interoperability).

Tiêu chí	Từ điển Dữ liệu (Data Dictionary)	Ontology
Mục đích	Mô tả <b>kỹ thuật</b> của dữ liệu (tên cột, kiểu dữ liệu, định nghĩa ngắn).	Mô tả ngữ nghĩa của một miền kiến thức (concepts, relationships).
Phạm vi	Tập trung vào ' <b>CÁI GÌ</b> ' (What). Ví dụ: 'CCCD' là một chuỗi 12 ký tự.	Tập trung vào 'TẠI SAO' và 'NHƯ THẾ NÀO' (Why & How). Ví dụ: 'CCCD' <b>định danh</b> cho một 'Công dân', và 'Công dân' <b>sở hữu</b> một 'Thửa đất'.
Khả năng	Giúp <b>con người hiểu</b> dữ liệu.	Giúp cả <b>con người và máy tính (AI) hiểu và suy luận</b> từ dữ liệu.
Ví dụ	Danh sách các trường trong CSDL Dân cư.	Mô hình hóa <b>mối quan hệ</b> giữa Công dân, Hộ khẩu, Tài sản, và An sinh xã hội.

Từ điển dữ liệu là một **danh sách**. Ontology là một **bản đồ tri thức**.

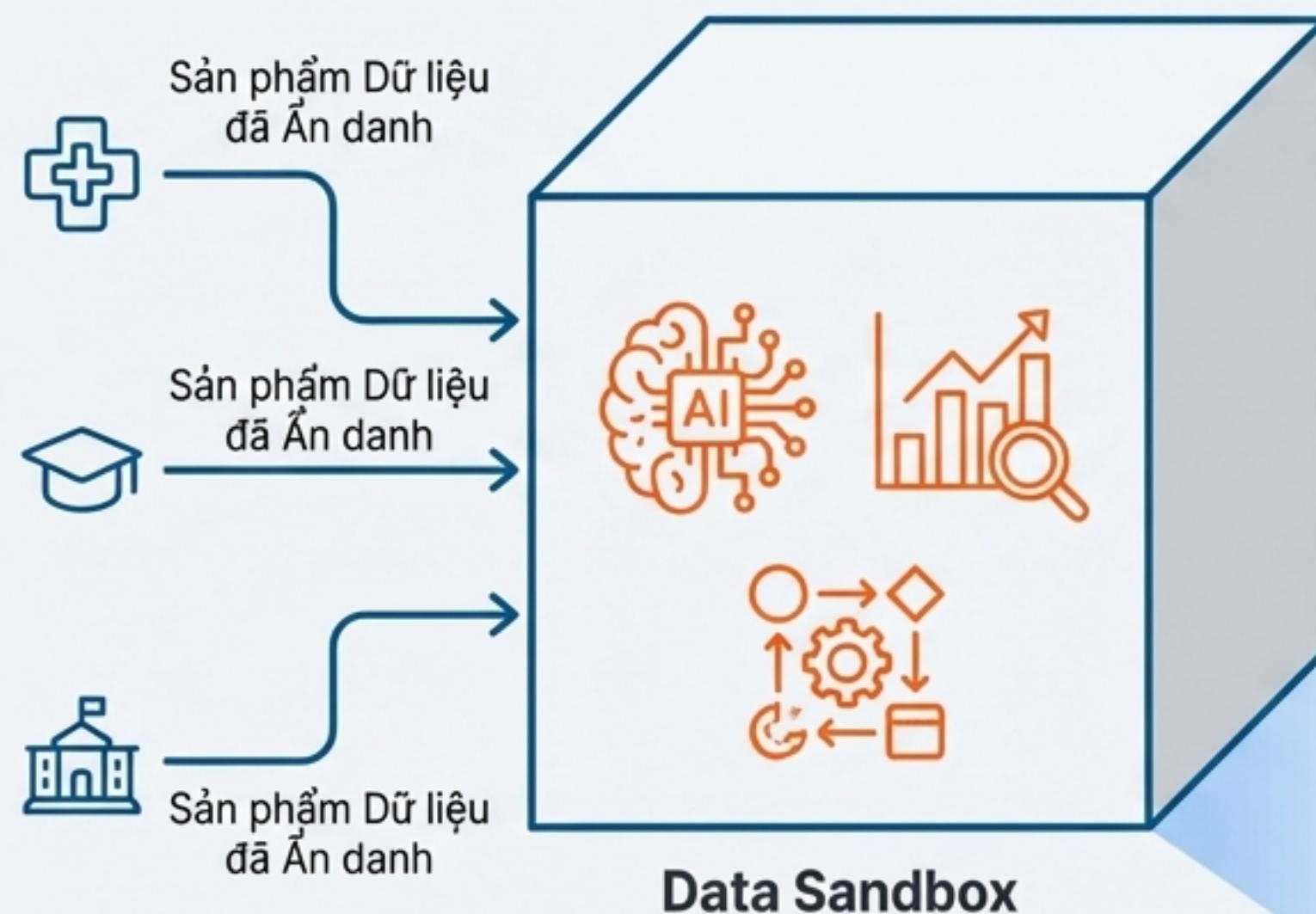
# Sức mạnh của Ontology trong Thực tiễn: Kết nối đa miền

Làm thế nào để trả lời câu hỏi phức tạp: “Xác định các hộ gia đình có trẻ em trong độ tuổi đi học sống trong khu vực có nguy cơ ô nhiễm môi trường cao?”



Nếu không có Ontology để định nghĩa các mối quan hệ này (“là thành viên của”, “cư trú tại”, “nằm trong”), việc kết nối dữ liệu từ 5-6 CSDL khác nhau sẽ cực kỳ phức tạp và tốn thời gian. Ontology giúp tự động hóa quá trình suy luận này.

# Động cơ Đổi mới Sáng tạo: Môi trường Sandbox Dữ liệu

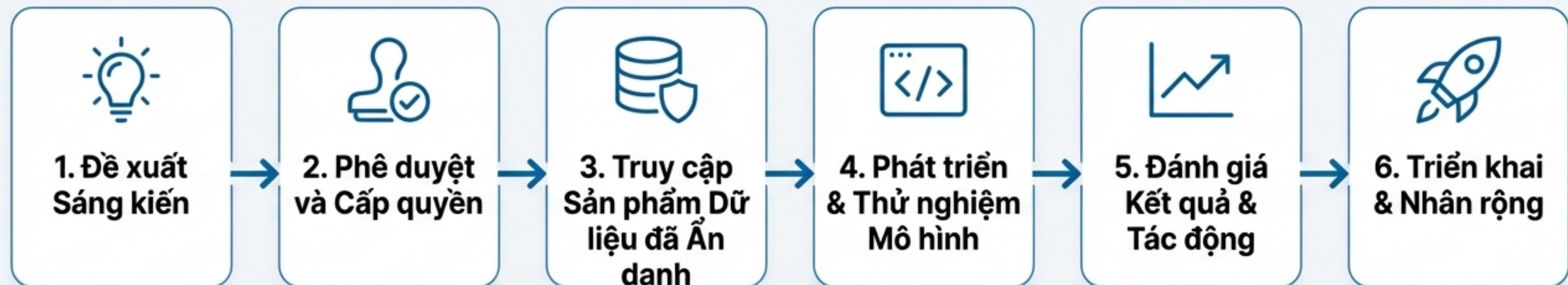


**Sandbox Dữ liệu** là một môi trường an toàn, được kiểm soát chặt chẽ, cho phép các cơ quan, viện nghiên cứu, và đối tác tin cậy thử nghiệm, phân tích và xây dựng các mô hình/ứng dụng mới dựa trên dữ liệu công đã được ẩn danh hoặc tổng hợp.

## Tại sao Kiến trúc Miền lại là nền tảng cho Sandbox?

- **Cung cấp 'Sản phẩm Dữ liệu' sẵn sàng:** Các miền dữ liệu có thể cung cấp các bộ dữ liệu chất lượng cao, đã được chuẩn hóa và ẩn danh (data products) cho môi trường sandbox.
- **Phân quyền và kiểm soát truy cập rõ ràng:** Kiến trúc sư Miền có thể cấp quyền truy cập có thời hạn vào các bộ dữ liệu cụ thể, đảm bảo an toàn.
- **Thúc đẩy thử nghiệm liên-miền:** Cho phép kết hợp dữ liệu từ nhiều miền khác nhau (ví dụ: Y tế + Giao thông) để giải quyết các bài toán phức tạp.

# Vòng đời của một Sáng kiến trong Sandbox: Từ Ý tưởng đến Tác động



# Quản trị Sandbox: Sáng tạo trong một Hành lang An toàn

Để sandbox trở thành một công cụ hiệu quả, cần một khung quản trị chặt chẽ, cân bằng giữa sự tự do sáng tạo và yêu cầu bảo mật, tuân thủ.



## 1. Quy tắc Truy cập Rõ ràng (Clear Access Rules)

Ai được truy cập? Truy cập dữ liệu gì? Trong bao lâu? Mục đích là gì? Mọi quyền truy cập đều được ghi log và kiểm toán.



## 2. Ưu tiên Bảo vệ Dữ liệu Cá nhân (Privacy by Design)

Mọi dữ liệu cá nhân phải được ẩn danh hóa (anonymization), mã hóa (tokenization) hoặc tổng hợp hóa (aggregation) trước khi đưa vào sandbox. Tuyệt đối không sử dụng dữ liệu cá nhân thô.



## 3. Môi trường Cách ly (Isolated Environment)

Môi trường sandbox phải được cách ly hoàn toàn về mặt kỹ thuật với các hệ thống vận hành chính thức để tránh mọi rủi ro về an ninh. Dữ liệu không thể bị sao chép ra ngoài.



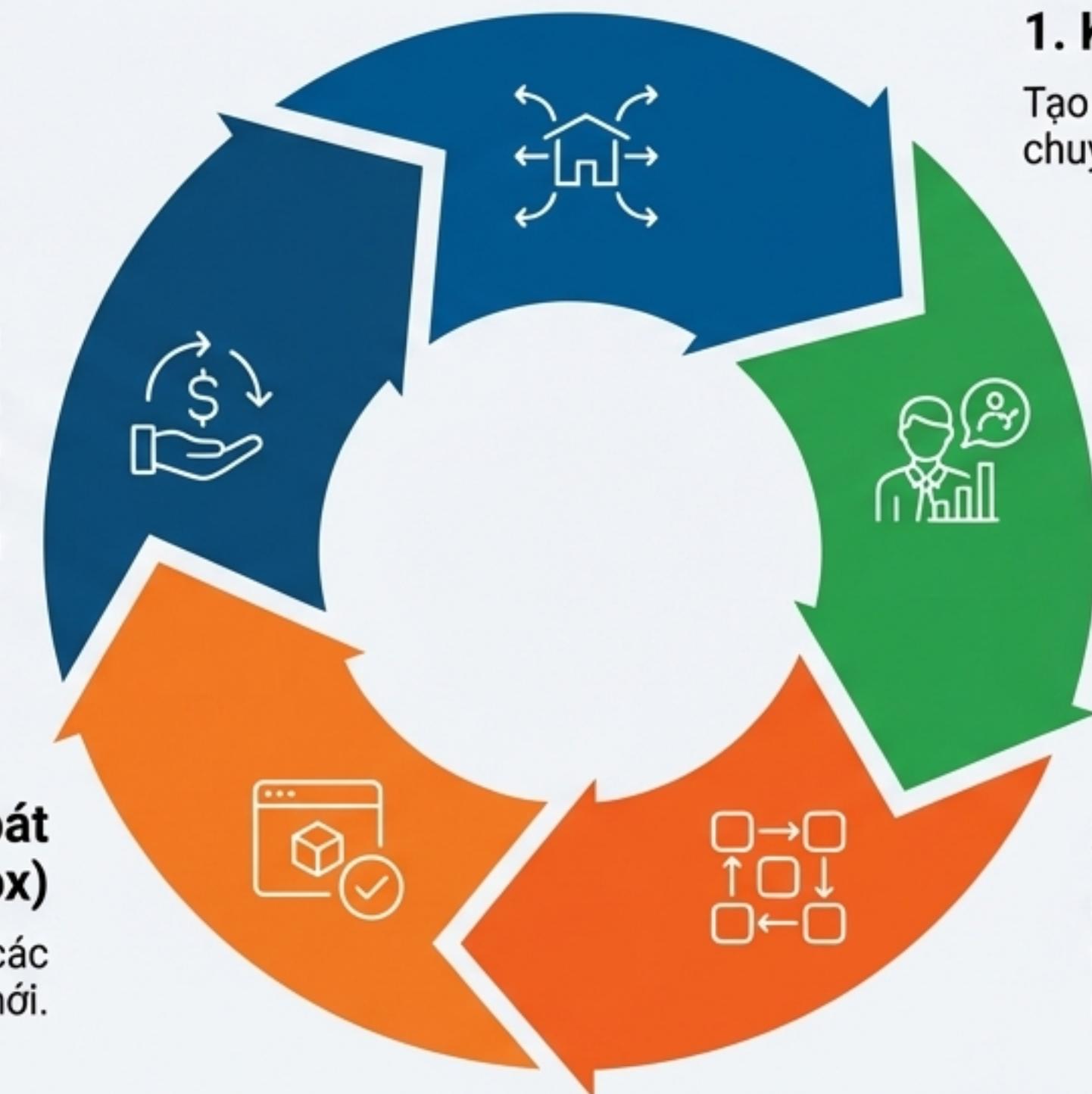
## 4. Tiêu chí Thành công & Lối ra (Success & Exit Criteria)

Mỗi dự án trong sandbox phải có mục tiêu và tiêu chí đánh giá thành công rõ ràng. Sau khi kết thúc, quyền truy cập sẽ bị thu hồi và môi trường được làm sạch.

Sandbox là môi trường **rủi ro thấp, tác động cao** để thúc đẩy các quan hệ đối tác **công-tư và mô phỏng chính sách**.

# Vòng tuần hoàn Giá trị: Từ Kiến trúc đến Đổi mới bền vững

Kiến trúc Miền, vai trò Kiến trúc sư, Ontology và Sandbox không phải là các yếu tố độc lập. Chúng cùng nhau tạo thành một vòng tuần hoàn giá trị, nơi mỗi thành phần cung cấp cho các thành phần còn lại, tạo ra động lực phát triển liên tục.



## 5. Giá trị Công được tạo ra

Các giải pháp thành công giúp cải thiện dịch vụ công, tối ưu chính sách, và tạo ra dữ liệu mới, lại tiếp tục làm giàu cho các Miền Dữ liệu, khởi động lại vòng tuần hoàn ở một cấp độ cao hơn.

## 1. Kiến trúc Miền linh hoạt

Tạo ra nền tảng cho sự tự chủ và chuyên môn hóa.

## 2. Kiến trúc sư được trao quyền

Dẫn dắt việc tạo ra các sản phẩm dữ liệu chất lượng cao.

## 4. Đổi mới được kiểm soát (Sandbox)

Tận dụng dữ liệu liên-miền để tạo ra các giải pháp và tri thức mới.

## 3. Ngữ nghĩa rõ ràng (Ontology)

Cho phép các sản phẩm dữ liệu từ nhiều miền khác nhau được kết nối và tích hợp một cách thông minh.

# Ba Thông điệp Then chốt

1

## KIẾN TRÚC MIỀN LÀ NỀN TẢNG CHO SỰ LINH HOẠT.

Để vượt qua các điểm nghẽn của mô hình tập trung, cần trao quyền sở hữu và trách nhiệm dữ liệu cho các miền chuyên môn.

2

## KIẾN TRÚC SƯ & ONTOLOGY LÀ CHẤT KẾT DÍNH.

Kiến trúc mới đòi hỏi vai trò lãnh đạo kỹ thuật (Kiến trúc sư Miền) và một ngôn ngữ chung (Ontology) để đảm bảo sự liên thông và nhất quán.

3

## SANDBOX BIẾN KIẾN TRÚC THÀNH GIÁ TRỊ.

Giá trị thực sự được tạo ra trong các môi trường thử nghiệm an toàn, nơi dữ liệu được biến thành các giải pháp đổi mới sáng tạo phục vụ người dân và xã hội.

Chuyên đề này đặt nền móng kiến trúc cho sự đổi mới. Tiếp theo, Chuyên đề 9 sẽ tổng kết toàn bộ **hành trình xây dựng Chính phủ dữ liệu**.

→ **Chuyên đề 9: Tổng kết và  
Bài tập tổng hợp**