

Số: **1973**/QĐ-BKHCN

Hà Nội, ngày **31** tháng **3** năm 2026

QUYẾT ĐỊNH

**Về ban hành Khung kiến trúc dữ liệu, Khung quản trị,
quản lý dữ liệu Bộ Khoa học và Công nghệ (Phiên bản 1.0)**

BỘ TRƯỞNG BỘ KHOA HỌC VÀ CÔNG NGHỆ

Căn cứ Luật Tổ chức Chính phủ ngày 18 tháng 02 năm 2025;

Căn cứ Luật Chuyển đổi số ngày 11 tháng 12 năm 2025;

Căn cứ Luật An ninh mạng ngày 10 tháng 12 năm 2025;

Căn cứ Luật Công nghiệp công nghệ số ngày 14 tháng 6 năm 2025;

Căn cứ Luật Giao dịch điện tử ngày 22 tháng 06 năm 2023;

Căn cứ Luật Viễn thông ngày 24 tháng 11 năm 2023;

Căn cứ Luật Dữ liệu ngày 30 tháng 11 năm 2024;

Căn cứ Luật Bảo vệ dữ liệu cá nhân ngày 26 tháng 06 năm 2025;

Căn cứ Luật Bảo vệ bí mật nhà nước ngày 10 tháng 12 năm 2025;

Căn cứ Nghị định số 55/2025/NĐ-CP của Chính phủ quy định chức năng, nhiệm vụ, quyền hạn và cơ cấu tổ chức của Bộ Khoa học và Công nghệ;

Căn cứ Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia;

Căn cứ Kế hoạch số 02-KH/BCĐTW ngày 19 tháng 6 năm 2025 của Ban Chỉ đạo Trung ương về phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số về thúc đẩy chuyển đổi số liên thông, đồng bộ, nhanh, hiệu quả đáp ứng yêu cầu sắp xếp tổ chức bộ máy của hệ thống chính trị;

Căn cứ Nghị định số 42/2022/NĐ-CP ngày 24 tháng 6 năm 2022 của Chính phủ cung cấp thông tin dịch vụ công trực tuyến trên môi trường mạng;

Căn cứ Nghị định số 165/2025/NĐ-CP ngày 30 tháng 6 năm 2025 của Chính phủ quy định chi tiết và biện pháp thi hành Luật Dữ liệu;



Căn cứ Nghị định số 194/2025/NĐ-CP ngày 03 tháng 7 năm 2025 của Chính phủ quy định chi tiết một số điều của Luật Giao dịch điện tử về cơ sở dữ liệu quốc gia, kết nối và chia sẻ dữ liệu, dữ liệu mở phục vụ giao dịch điện tử của cơ quan nhà nước;

Căn cứ Nghị định số 278/2025/NĐ-CP ngày 22 tháng 10 năm 2025 của Chính phủ quy định về kết nối, chia sẻ dữ liệu bắt buộc giữa các cơ quan thuộc hệ thống chính trị;

Căn cứ Nghị định số 356/2025/NĐ-CP ngày 31 tháng 12 năm 2025 của Chính phủ quy định chi tiết một số điều và biện pháp thi hành Luật Bảo vệ dữ liệu cá nhân;

Căn cứ Nghị quyết số 71/NQ-CP ngày 01 tháng 4 năm 2025 của Chính phủ về sửa đổi, bổ sung cập nhật Chương trình hành động của Chính phủ thực hiện Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia;

Căn cứ Nghị quyết số 11/NQ-CP ngày 14 tháng 01 năm 2026 của Chính phủ về sửa đổi, bổ sung, cập nhật một số nội dung tại Chương trình hành động của Chính phủ thực hiện Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia;

Căn cứ Quy định số 05-QĐ/BCĐTW ngày 27 tháng 8 năm 2025 của Ban Chỉ đạo Trung ương về phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số ban hành quy định về Mô hình liên thông số thống nhất, hiệu quả và quản trị dựa trên quản trị dữ liệu trong hệ thống chính trị;

Căn cứ Nghị quyết số 214/NQ-CP ngày 23 tháng 7 năm 2025 của Chính phủ về thúc đẩy tạo lập dữ liệu phục vụ chuyển đổi số toàn diện;

Căn cứ Quyết định số 2439/QĐ-TTg ngày 04 tháng 11 năm 2025 của Thủ tướng Chính phủ ban hành Khung kiến trúc dữ liệu quốc gia, Khung quản trị, quản lý dữ liệu quốc gia, Từ điển dữ liệu dùng chung (Phiên bản 1.0);

Căn cứ Quyết định số 2629/QĐ-TTg ngày 01 tháng 12 năm 2025 của Thủ tướng Chính phủ về việc phê duyệt Chương trình phát triển Chính phủ số;

Căn cứ Quyết định số 11/QĐ-BCĐCP ngày 27 tháng 02 năm 2026 của Ban chỉ đạo Chính phủ về Kế hoạch hoạt động năm 2026 Ban Chỉ đạo của Chính phủ về phát triển khoa học, công nghệ, đổi mới sáng tạo, chuyển đổi số và Đề án 06;

Căn cứ Quyết định số 3090/QĐ-BKHCN ngày 08 tháng 10 năm 2025 của Bộ trưởng Bộ Khoa học và Công nghệ về ban hành Khung kiến trúc tổng thể quốc gia số;



Căn cứ Quyết định số 176/QĐ-BKHCN ngày 13 tháng 01 năm 2026 của Bộ trưởng Bộ Khoa học và Công nghệ về Ban hành Kế hoạch hành động năm 2026 của Bộ Khoa học và Công nghệ thực hiện Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia;

Căn cứ Quyết định số 1262/QĐ-BKHCN ngày 18 tháng 02 năm 2026 của Bộ trưởng Bộ Khoa học và Công nghệ về Điều chỉnh, bổ sung Kế hoạch hành động năm 2026 của Bộ Khoa học và Công nghệ thực hiện Nghị quyết số 57-NQ/TW ngày 22 tháng 12 năm 2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia;

Theo đề nghị của Giám đốc Trung tâm Công nghệ thông tin,

QUYẾT ĐỊNH:

Điều 1. Ban hành Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ Khoa học và Công nghệ (Phiên bản 1.0) (chi tiết Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu của Bộ kèm theo).

Điều 2. Quyết định này có hiệu lực kể từ ngày ký ban hành.

Điều 3. Chánh Văn phòng, Giám đốc Trung tâm Công nghệ thông tin, Thủ trưởng các đơn vị thuộc Bộ và các cơ quan, tổ chức có liên quan chịu trách nhiệm thi hành Quyết định này. /.

Nơi nhận:

- Như Điều 3;
- Bộ trưởng (để b/c);
- Các Thứ trưởng;
- Ban Chỉ đạo dữ liệu của Bộ;
- Các cơ quan, đơn vị thuộc Bộ;
- Công TTĐT của Bộ;
- Lưu: VT, TTCNTT.

**KT. BỘ TRƯỞNG
THỨ TRƯỞNG**



Bùi Hoàng Phương



KHUNG KIẾN TRÚC DỮ LIỆU, KHUNG QUẢN TRỊ, QUẢN LÝ DỮ LIỆU BỘ KHOA HỌC VÀ CÔNG NGHỆ (Phiên bản 1.0)

(Kèm theo Quyết định số **1973**/QĐ-BKHCN ngày **31** tháng **3** năm 2026
của Bộ trưởng Bộ Khoa học và Công nghệ)

I. TẦM NHÌN, MỤC TIÊU VÀ PHẠM VI

1. Tầm nhìn

- Xây dựng Khung kiến trúc dữ liệu Bộ nhằm hình thành một nền tảng dữ liệu thống nhất, đồng bộ và có khả năng kết nối, chia sẻ dữ liệu trong phạm vi Bộ KH&CN và phạm vi quốc gia theo yêu cầu.

- Hệ thống dữ liệu của Bộ được định hướng phát triển theo hướng trở thành tài sản chiến lược, phục vụ hiệu quả công tác quản lý nhà nước, hoạch định chính sách, chỉ đạo điều hành và cung cấp dịch vụ công. Dữ liệu được quản lý tập trung, chuẩn hóa và có khả năng liên thông, chia sẻ với các hệ thống dữ liệu quốc gia, các bộ, ngành và địa phương, qua đó hình thành hệ sinh thái dữ liệu mở, phục vụ phát triển chính phủ số, kinh tế số và xã hội số.

- Khung kiến trúc dữ liệu của Bộ hướng tới việc bảo đảm mọi dữ liệu được quản lý theo nguyên tắc “đúng, đủ, sạch, sống, thống nhất, dùng chung”, nâng cao chất lượng dữ liệu, tăng cường khả năng phân tích, khai thác và sử dụng dữ liệu trong hoạt động quản lý và nghiên cứu của Bộ KH&CN.

2. Mục tiêu

2.1. Mục tiêu tổng quát

- Thiết lập quy hoạch tổng thể về dữ liệu của Bộ, bảo đảm sự thống nhất, đồng bộ và liên thông dữ liệu giữa các đơn vị trực thuộc Bộ, đồng thời kết nối, chia sẻ hiệu quả với các hệ thống dữ liệu quốc gia, các Bộ, ngành và địa phương. Qua đó, hình thành nền tảng dữ liệu vững chắc phục vụ công tác chỉ đạo, điều hành, phát triển chính phủ số, kinh tế số, xã hội số của Bộ KH&CN, đồng thời cung cấp dịch vụ công và các dịch vụ dữ liệu chất lượng cao cho người dân và doanh nghiệp.

- Kiến trúc dữ liệu của Bộ được xây dựng phù hợp với Khung kiến trúc dữ liệu quốc gia theo quy định của Chính phủ, bảo đảm sự nhất quán về cấu trúc, phân tầng và tiêu chuẩn kỹ thuật. Đây là cơ sở để tăng cường khả năng kết nối, tích hợp và liên thông giữa các cơ sở dữ liệu quốc gia, cơ sở dữ liệu chuyên ngành và hệ thống dữ liệu của các Bộ, ngành, địa phương, phục vụ hiệu quả công tác quản lý nhà nước, chỉ đạo điều hành và hoạch định chính sách.

- Trên nền tảng đó, kiến trúc dữ liệu là cơ sở để thiết lập Khung quản trị, quản lý dữ liệu và từ điển dữ liệu dùng chung của Bộ, bảo đảm dữ liệu được quản lý chặt chẽ, an toàn, minh bạch và khai thác hiệu quả. Đồng thời, tạo điều kiện để Bộ chủ động tham gia vào quá trình xây dựng quốc gia số, nâng cao hiệu lực, hiệu quả quản lý nhà nước và cung cấp dịch vụ công trực tuyến chất lượng cao.

2.2. Mục tiêu cụ thể

- Chuẩn hóa và thống nhất mô hình dữ liệu của Bộ; xác định rõ các nhóm dữ liệu dùng chung và dữ liệu chuyên ngành.

- Hình thành hệ thống dữ liệu thống nhất trong toàn Bộ, bảo đảm các dữ liệu được chuẩn hóa về cấu trúc, định dạng và nội dung. Các cơ sở dữ liệu và hệ thống thông tin của các đơn vị thuộc Bộ được kết nối, liên thông, giảm thiểu tình trạng phân tán, cát cứ dữ liệu. Dữ liệu được tổ chức theo danh mục dùng chung và từ điển dữ liệu thống nhất, bảo đảm khả năng tích hợp và khai thác trên phạm vi toàn Bộ.

- Thiết lập đầy đủ khung quản trị dữ liệu, bao gồm các quy định, quy trình và tiêu chuẩn quản lý dữ liệu. Trong đó, xác định rõ trách nhiệm của các đơn vị trong việc tạo lập, cập nhật, quản lý và chia sẻ dữ liệu; đồng thời thiết lập cơ chế giám sát, đánh giá việc tuân thủ các quy định về quản trị dữ liệu. Dữ liệu được quản lý xuyên suốt vòng đời, từ khâu tạo lập đến khai thác và lưu trữ.

- Nâng cao chất lượng dữ liệu thông qua việc áp dụng các tiêu chuẩn và công cụ kiểm soát chất lượng dữ liệu. Bảo đảm dữ liệu có độ chính xác, đầy đủ, nhất quán và kịp thời, phục vụ hiệu quả cho công tác quản lý nhà nước, nghiên cứu khoa học và đổi mới sáng tạo. Từng bước xây dựng cơ chế đánh giá và xếp hạng chất lượng dữ liệu đối với các hệ thống thông tin trong Bộ.

- Xây dựng và vận hành hiệu quả nền tảng tích hợp, chia sẻ dữ liệu của Bộ, bảo đảm kết nối thông suốt giữa các hệ thống thông tin nội bộ và với các hệ thống

bên ngoài. Dữ liệu được chia sẻ theo cơ chế kiểm soát, phân quyền rõ ràng, đáp ứng yêu cầu khai thác dữ liệu của các đơn vị, tổ chức và cá nhân theo quy định.

- Đảm bảo an toàn thông tin và an ninh dữ liệu, tuân thủ Khung kiến trúc dữ liệu quốc gia và các quy định pháp luật có liên quan, bảo vệ dữ liệu trong toàn bộ vòng đời khai thác và sử dụng.

- Hiện đại hóa hạ tầng và nền tảng dữ liệu, ứng dụng các công nghệ mới trong quản lý và khai thác dữ liệu như dữ liệu lớn, trí tuệ nhân tạo và phân tích dữ liệu nâng cao, phù hợp với định hướng phát triển Chính phủ số.

3. Phạm vi áp dụng

- Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ được áp dụng thống nhất đối với các đơn vị thuộc Bộ KH&CN. Khung này đồng thời là căn cứ để tổ chức phối hợp, liên thông và chia sẻ dữ liệu giữa Bộ KH&CN với các bộ, ngành, địa phương và cơ quan nhà nước khác theo quy định của pháp luật.

- Phạm vi quản lý bao gồm toàn bộ vòng đời dữ liệu số trong Bộ KH&CN, từ thu thập, tạo lập, cập nhật, chuẩn hóa, phân loại, dán nhãn, lưu trữ, bảo vệ, xử lý, chia sẻ, khai thác, điều phối đến lưu trữ lịch sử và hủy bỏ dữ liệu.

II. NGUYÊN TẮC, YÊU CẦU

1. Nguyên tắc

a) Về tuân thủ Khung kiến trúc tổng thể quốc gia số, Khung kiến trúc dữ liệu quốc gia, Khung quản trị, quản lý dữ liệu quốc gia và Khung kiến trúc số Bộ

- Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ được xây dựng trên cơ sở tuân thủ Khung kiến trúc dữ liệu quốc gia, Khung quản trị và quản lý dữ liệu quốc gia, Từ điển dữ liệu dùng chung, bảo đảm thống nhất về nguyên tắc, cấu trúc và phương pháp tiếp cận.

- Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ là một hợp phần của Khung kiến trúc số Bộ, được xây dựng phù hợp và thống nhất với Khung kiến trúc tổng thể quốc gia số.

- Mọi thành phần Khung kiến trúc dữ liệu Bộ phải bảo đảm khả năng kết nối, liên thông và tích hợp với các hệ thống dữ liệu quốc gia, các bộ, ngành và địa phương khác, tránh trùng lặp, chồng chéo trong đầu tư và xây dựng dữ liệu.

- Việc thiết kế Khung kiến trúc dữ liệu Bộ phải bảo đảm đồng bộ giữa các lớp kiến trúc (nghiệp vụ, dữ liệu, ứng dụng, hạ tầng), hỗ trợ hiệu quả cho hoạt động chuyển đổi số, quản lý, điều hành và cung cấp dịch vụ số của Bộ.

b) Về thu thập, tạo lập và cập nhật dữ liệu

- Thu thập dữ liệu một lần: dữ liệu chỉ được thu thập một lần duy nhất tại cơ quan có thẩm quyền tạo lập, bảo đảm tính chính xác, đề hạn chế việc thu thập trùng lặp giữa các hệ thống thông tin, giảm chi phí xử lý và có độ tin cậy cao.

- Một nguồn dữ liệu gốc duy nhất: mỗi loại dữ liệu phải xác định rõ cơ quan chủ quản chịu trách nhiệm tạo lập, cập nhật, quản lý, chịu trách nhiệm bảo đảm chất lượng và giá trị pháp lý của dữ liệu tạo sự rõ ràng trong kiểm soát, bảo mật và chia sẻ giữa các hệ thống.

- Phân loại dữ liệu và bảo mật: bảo đảm phân loại, gán nhãn dữ liệu theo mức độ bảo mật, phân loại theo đối tượng và cấp độ khai thác dữ liệu; áp dụng các biện pháp bảo vệ phù hợp về kỹ thuật và tổ chức bảo đảm tuân thủ quy định về quốc phòng, an ninh, bí mật nhà nước và quyền, lợi ích hợp pháp của tổ chức, cá nhân.

- Cập nhật và đồng bộ liên tục: dữ liệu phải được cập nhật kịp thời, chính xác và đảm bảo được khả năng liên thông, tích hợp, đồng bộ, chia sẻ dữ liệu giữa các hệ thống thông tin có liên quan thông qua Nền tảng tích hợp, chia sẻ dữ liệu của Bộ kết nối CSDL chuyên ngành, CSDL quốc gia.

- Tuân thủ quy định quản lý dữ liệu chủ quốc gia: đối với các dữ liệu thuộc danh mục dữ liệu chủ quốc gia, việc cập nhật, thay đổi phải tuân thủ quy định và được đồng bộ với các hệ thống dữ liệu quốc gia theo quy định hiện hành.

c) Về lưu trữ và quản lý dữ liệu

- Lưu trữ dữ liệu ở mức cần thiết: các cơ quan, đơn vị có trách nhiệm lưu trữ dữ liệu ở mức cần thiết, chỉ lưu trữ dữ liệu phù hợp với chức năng, nhiệm vụ được giao và mục tiêu sử dụng, tránh lưu trữ dư thừa, trùng lặp, đồng thời bảo đảm tiết kiệm tài nguyên hệ thống.

- Quản lý lưu trữ dữ liệu theo vòng đời: dữ liệu phải được quản lý xuyên suốt vòng đời từ khi tạo lập, lưu trữ, khai thác/chia sẻ, đến lưu trữ lịch sử hoặc hủy bỏ theo quy định.

- Bảo đảm chất lượng dữ liệu: dữ liệu phải đáp ứng các yêu cầu về tính đúng, đủ, sạch, sống, nhất quán và cập nhật kịp thời, đáp ứng yêu cầu quản lý, sẵn sàng khai thác và phục vụ công việc phân tích, ra quyết định.

d) Về chia sẻ và kết nối dữ liệu

- Minh bạch về dữ liệu chia sẻ: công khai thông tin về nguồn gốc, nội dung và phạm vi sử dụng của dữ liệu chia sẻ, nâng cao mức độ tin cậy và tăng cường giám sát.

- Chia sẻ dữ liệu chủ động: việc chia sẻ cần được triển khai bằng phương thức chủ động, sẵn sàng kết nối và cung cấp dữ liệu, phục vụ dịch vụ công, nhiệm vụ quản lý nhà nước, điều hành và ra quyết định.

- Tối thiểu hóa trung gian kết nối dữ liệu: chia sẻ dữ liệu phải được thực hiện thông qua Nền tảng tích hợp, chia sẻ dữ liệu của Bộ, tối thiểu hóa các khâu trung gian trong quá trình kết nối, bảo đảm dữ liệu được truyền tải nhanh chóng, chính xác và tiết kiệm chi phí hạ tầng.

đ) Về khai thác, sử dụng và kết thúc sử dụng dữ liệu

- Sử dụng dữ liệu đúng mục đích: chỉ được khai thác dữ liệu phục vụ hoạt động quản lý nhà nước, bảo đảm tuân thủ các quy định về quản lý dữ liệu, an toàn thông tin và bảo vệ dữ liệu theo quy định của pháp luật.

- Minh bạch và trách nhiệm: cơ quan khai thác phải công khai phạm vi sử dụng và chịu trách nhiệm về mục đích, kết quả sử dụng dữ liệu;

- Việc cung cấp, chia sẻ dữ liệu ra bên ngoài Bộ phải tuân thủ các quy định về phân loại dữ liệu, bảo mật thông tin và quyền tiếp cận dữ liệu; chỉ thực hiện đối với các dữ liệu được phép chia sẻ theo quy định.

- Dữ liệu được khai thác, sử dụng đúng mục đích, đúng thẩm quyền và phù hợp với chức năng, nhiệm vụ của từng đơn vị, cá nhân. Việc truy cập và sử dụng dữ liệu phải được kiểm soát thông qua cơ chế phân quyền, xác thực và ghi nhận nhật ký, bảo đảm khả năng theo dõi, giám sát và truy vết trong quá trình sử dụng.

- Khi kết thúc nhu cầu sử dụng, dữ liệu phải được lưu trữ lịch sử hoặc hủy bỏ dữ liệu theo quy định của pháp luật.

- Khi kết thúc mục đích sử dụng, dữ liệu phải được xử lý theo quy định về lưu trữ, bảo quản hoặc hủy bỏ. Việc lưu trữ dữ liệu phải bảo đảm khả năng tra

cứu, khai thác khi cần thiết; việc hủy bỏ dữ liệu phải tuân thủ quy trình, bảo đảm không làm lộ, lọt thông tin và không ảnh hưởng đến tính toàn vẹn của hệ thống.

- Việc quản lý khai thác, sử dụng và kết thúc sử dụng dữ liệu phải được thực hiện xuyên suốt vòng đời dữ liệu, bảo đảm dữ liệu được sử dụng hiệu quả, an toàn và đúng quy định.

2. Yêu cầu

Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ cần bảo đảm một số yêu cầu dưới đây:

a) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ là cơ sở tham chiếu thống nhất cho việc xây dựng, nâng cấp, vận hành các cơ sở dữ liệu và hệ thống thông tin có quản lý dữ liệu của các đơn vị thuộc Bộ.

b) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ phải kế thừa và tuân thủ Khung kiến trúc dữ liệu quốc gia, Khung quản trị, quản lý dữ liệu quốc gia, Từ điển dữ liệu dùng chung và các quy định pháp luật có liên quan.

c) Khung kiến trúc dữ liệu Bộ phải là một hợp phần thống nhất trong Khung kiến trúc số của Bộ, bảo đảm đồng bộ với các lớp kiến trúc nghiệp vụ, ứng dụng và hạ tầng.

d) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ phải quy định rõ phạm vi lưu trữ dữ liệu của từng đơn vị, phù hợp với chức năng, nhiệm vụ được giao, tránh lưu trữ trùng lặp, phân tán.

đ) Khung kiến trúc dữ liệu Bộ phải xác định rõ điểm tạo lập dữ liệu đối với từng nhóm dữ liệu, mối quan hệ giữa các nhóm dữ liệu, bảo đảm mỗi loại dữ liệu có một nguồn dữ liệu gốc duy nhất chịu trách nhiệm pháp lý.

e) Khung kiến trúc dữ liệu Bộ phải áp dụng thống nhất các tiêu chuẩn, quy chuẩn kỹ thuật về dữ liệu, siêu dữ liệu (metadata), mã định danh theo quy định quốc gia; Chia sẻ dữ liệu phải được thực hiện thông qua Nền tảng tích hợp, chia sẻ dữ liệu Bộ, bảo đảm khả năng liên thông dữ liệu.

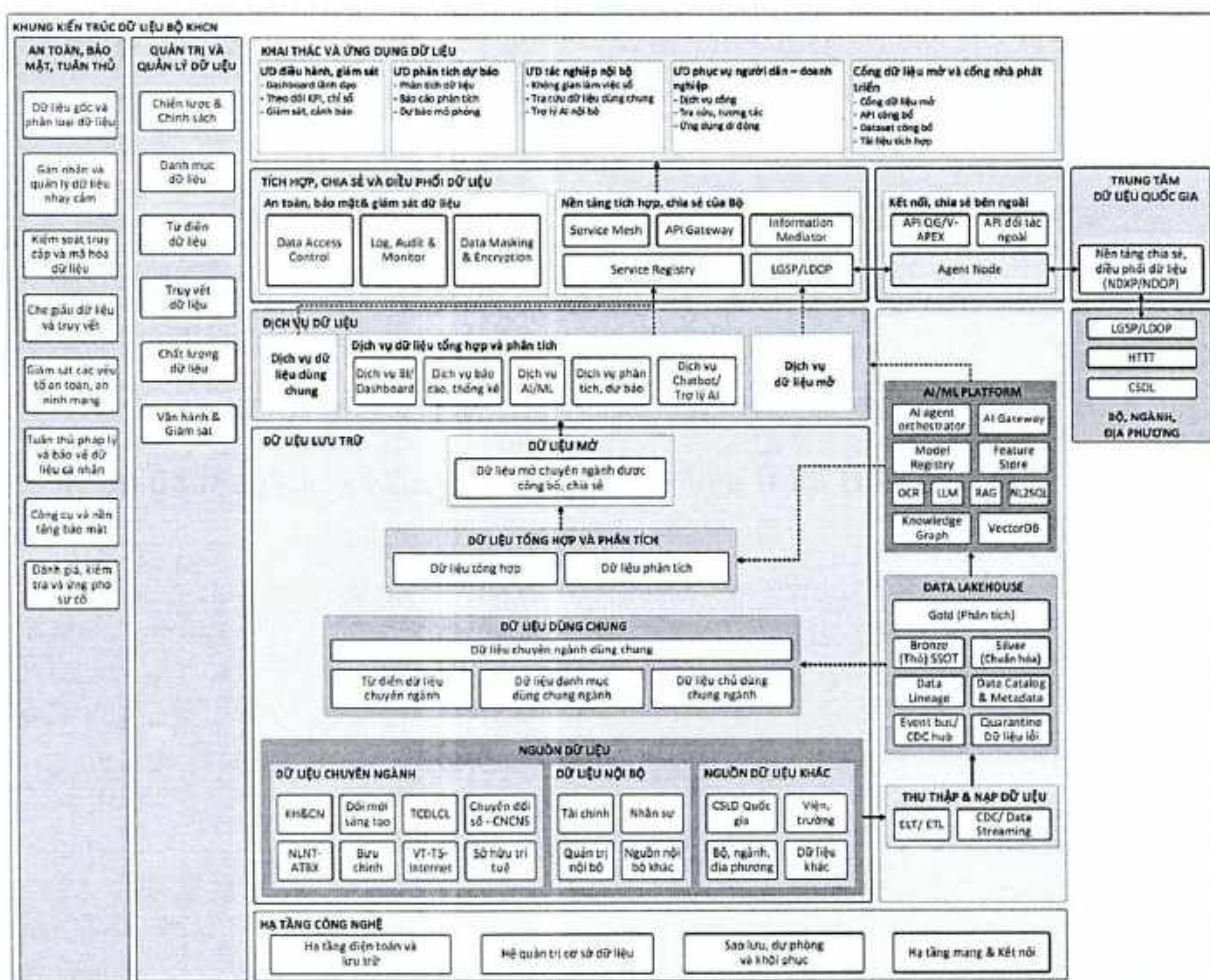
g) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ đảm bảo tuân thủ Luật An ninh mạng, Luật Bảo vệ dữ liệu cá nhân, Luật Bảo vệ bí mật nhà nước, Luật Dữ liệu và các văn bản quy phạm pháp luật khác có liên quan.

h) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ phải bảo đảm khả thi trong điều kiện hạ tầng, nguồn lực và tổ chức hiện có của Bộ, là căn cứ để xây dựng lộ trình, kế hoạch và các dự án triển khai dữ liệu theo từng giai đoạn, bảo đảm triển khai thống nhất, tránh đầu tư trùng lặp.

i) Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ được đánh giá, rà soát và cập nhật định kỳ theo quy định về quản trị, quản lý dữ liệu và theo yêu cầu thực tiễn triển khai của Bộ.

III. KHUNG KIẾN TRÚC DỮ LIỆU BỘ KH&CN

1. Sơ đồ tổng quát Khung kiến trúc dữ liệu Bộ KH&CN



Hình 1. Mô hình tổng thể Khung kiến trúc dữ liệu Bộ KH&CN

Khung kiến trúc dữ liệu Bộ KH&CN được tổ chức theo các lớp chức năng chính và 2 thành phần xuyên suốt, đồng bộ với Khung Kiến trúc số của Bộ KH&CN. Thứ tự phân tầng tuân thủ logic: Hạ tầng → Dữ liệu → Dịch vụ → Tích hợp → Ứng dụng.

Lớp	Tên lớp	UGP	Mô tả chức năng
XS1	An toàn, Bảo mật, Tuân thủ	Xuyên suốt	Áp dụng xuyên suốt tất cả lớp; phân cấp bảo mật theo cấp độ; tuân thủ đầy đủ hệ thống pháp luật về dữ liệu, an ninh mạng, trí tuệ nhân tạo.
XS2	Quản trị và quản lý dữ liệu	Xuyên suốt	6 trụ cột: Chiến lược & Chính sách; Danh mục dữ liệu; Từ điển dữ liệu; Truy vết dữ liệu; Chất lượng dữ liệu; Vận hành & Giám sát
L5	Khai thác và ứng dụng dữ liệu	Lớp 5	Giao diện và ứng dụng đầu cuối phục vụ người dùng, bao gồm ứng dụng điều hành giám sát, ứng dụng phân tích và dự báo, ứng dụng tác nghiệp nội bộ, ứng dụng phục vụ người dân và doanh nghiệp, cổng dữ liệu mở và cổng dành cho nhà phát triển nhằm khai thác hiệu quả dữ liệu của Bộ.
L4	Tích hợp, chia sẻ, điều phối	Lớp 4	Thực hiện thu thập, tích hợp, chia sẻ, kết nối và điều phối dữ liệu trong nội bộ Bộ và với các hệ thống bên ngoài; kết nối NDXP/Trung tâm dữ liệu quốc gia, đồng thời bảo đảm kiểm soát truy cập, giám sát, ghi log, bảo mật và an toàn dữ liệu trong quá trình trao đổi.
L3	Dịch vụ dữ liệu	Lớp 3	Cung cấp các dịch vụ dữ liệu và dịch vụ AI dùng chung dưới dạng service/API, bao gồm dịch vụ truy xuất dịch vụ BI/dashboard, dịch vụ báo cáo – thống kê, dịch vụ phân tích – dự báo, dịch vụ AI/ML, dịch vụ chatbot/trợ lý AI và các dịch vụ dữ liệu mở phục vụ các hệ thống, ứng dụng khai thác phía trên.

Lớp	Tên lớp	UGP	Mô tả chức năng
L2	Dữ liệu, lưu trữ	Lớp 2	Hình thành và quản lý tài sản dữ liệu của Bộ, bao gồm dữ liệu chuyên ngành dùng chung, từ điển dữ liệu chuyên ngành, dữ liệu danh mục dùng chung ngành, dữ liệu chủ chuyên ngành, dữ liệu tổng hợp và phân tích, dữ liệu mở.
L1	Hạ tầng công nghệ	Lớp 1	Hạ tầng điện toán tập trung; Trung tâm dữ liệu chính và dự phòng; Hệ quản trị cơ sở dữ liệu; Hạ tầng mạng và Kết nối; Hệ thống sao lưu và phục hồi.

1.1. Hạ tầng công nghệ

Phần hạ tầng công nghệ cung cấp năng lực tính toán, lưu trữ và kết nối vật lý cho toàn bộ hệ thống, bảo đảm tính sẵn sàng cao, an toàn và khả năng mở rộng linh hoạt.

a) Hạ tầng điện toán và lưu trữ:

- Triển khai tại Trung tâm dữ liệu dùng chung của Bộ (on-premise), phục vụ lưu trữ và xử lý các dữ liệu nhạy cảm, dữ liệu chuyên ngành.

- Kết hợp sử dụng hạ tầng điện toán đám mây đối với các hệ thống cần mở rộng linh hoạt như phân tích dữ liệu lớn, AI và dịch vụ công trực tuyến...

b) Hệ quản trị cơ sở dữ liệu:

- Hệ quản trị CSDL quan hệ cho dữ liệu vận hành nghiệp vụ của Bộ và các đơn vị.

- Hệ thống xử lý luồng sự kiện tập trung (Event Bus): tiếp nhận, phân phối và xử lý dữ liệu thay đổi theo thời gian thực từ mọi nguồn.

- Hệ thống lưu trữ đối tượng (Object Storage): lưu trữ vật lý cho toàn bộ kho dữ liệu tập trung, bảo đảm khả năng mở rộng và hiệu năng cao.

c) Sao lưu, dự phòng và khôi phục:

- Thực hiện sao lưu dữ liệu định kỳ và liên tục cho các hệ thống và cơ sở dữ liệu.

- Xây dựng phương án phục hồi hệ thống và dữ liệu khi xảy ra sự cố; xác định rõ mục tiêu thời gian phục hồi (RTO) và mục tiêu điểm phục hồi dữ liệu (RPO) phù hợp với từng cấp độ hệ thống.

- Dữ liệu mật được sao lưu theo phương thức ngoại tuyến, bảo đảm tuyệt đối an toàn.

d) Hạ tầng mạng và kết nối:

- Sử dụng Mạng truyền số liệu chuyên dùng của Đảng và Nhà nước cho các kết nối nội bộ giữa các cơ quan, bảo đảm tách biệt với mạng công cộng.

- Áp dụng mô hình bảo mật không tin tưởng mặc định (Zero Trust): mọi kết nối đều phải được xác thực và ủy quyền, không phân biệt nguồn gốc trong hay ngoài hệ thống.

- Quản lý khóa mã hóa, bí mật và chứng chỉ số tập trung.

1.2. Dữ liệu, lưu trữ

Tầng này quản lý tập trung tài sản dữ liệu của Bộ, đồng thời cung cấp nền tảng lưu trữ, xử lý, phân tích và trí tuệ nhân tạo phục vụ khai thác dữ liệu.

a) Nguồn dữ liệu: Nguồn dữ liệu của Bộ được chia thành ba nhóm chính:

- Dữ liệu chuyên ngành: Bao gồm dữ liệu phát sinh từ các hệ thống thông tin và hoạt động chuyên môn của Bộ KH&CN như: Dữ liệu về khoa học, công nghệ và đổi mới sáng tạo; Dữ liệu sở hữu trí tuệ (sáng chế, kiểu dáng, nhãn hiệu...); Dữ liệu tiêu chuẩn, đo lường, chất lượng; Dữ liệu về chuyển đổi số, công nghệ và các lĩnh vực quản lý chuyên ngành khác... Các dữ liệu chủ yếu ở dạng hồ sơ nghiệp vụ, tài liệu chuyên ngành và các tập dữ liệu phục vụ quản lý nhà nước.

- Dữ liệu nội bộ: Bao gồm dữ liệu phát sinh trong quá trình điều hành và hoạt động nội bộ của Bộ, như: Dữ liệu tài chính, kế toán; Dữ liệu nhân sự; Dữ liệu quản trị nội bộ; Nguồn dữ liệu nội bộ khác, ...Nhóm dữ liệu này phục vụ công tác quản trị nội bộ và hỗ trợ điều hành.

- Nguồn dữ liệu khác: Bao gồm dữ liệu thu thập, chia sẻ từ các tổ chức bên ngoài Bộ, như: Dữ liệu từ CSDL quốc gia, Dữ liệu từ các bộ, ngành, địa phương thông qua các nền tảng tích hợp và chia sẻ dữ liệu (LGSP/LDOP, NDXP/NDOP);

Dữ liệu từ viện nghiên cứu, trường đại học, doanh nghiệp và đối tác và các nguồn dữ liệu khác. Nhóm dữ liệu này giúp bổ sung, làm giàu và đối chiếu với dữ liệu nội bộ của Bộ.

b) Dữ liệu dùng chung:

- Dữ liệu chuyên ngành dùng chung là các tập dữ liệu đã được chuẩn hóa, tích hợp và cung cấp phục vụ nhiều hệ thống, nhiều đơn vị trong Bộ, như dữ liệu tổng hợp về nhiệm vụ KH&CN, dữ liệu thống kê, báo cáo dùng chung. Các tập dữ liệu này thường được khai thác thông qua các dịch vụ dữ liệu hoặc API để phục vụ điều hành, phân tích và chia sẻ.

- Dữ liệu danh mục dùng chung ngành (Reference data) là các bảng mã, danh mục phân loại do cơ quan nhà nước ban hành (ví dụ: danh mục lĩnh vực KH&CN, loại hình tổ chức, phân loại nhiệm vụ...), được sử dụng thống nhất trong toàn Bộ nhằm bảo đảm tính tương thích và liên thông dữ liệu giữa các hệ thống.

- Từ điển dữ liệu ngành đóng vai trò làm công cụ chuẩn hóa, chứa các định nghĩa, thuật ngữ nghiệp vụ, kiểu dữ liệu, mã tham chiếu và cấu trúc dữ liệu để bảo đảm tính nhất quán khi trao đổi dữ liệu giữa các hệ thống.

- Dữ liệu chủ dùng chung ngành (Master data) là các thông tin cốt lõi mô tả các đối tượng quản lý chính của Bộ như tổ chức KH&CN, cá nhân nghiên cứu, nhiệm vụ KH&CN, đối tượng sở hữu trí tuệ... Đây là dữ liệu có tính ổn định cao, được sử dụng làm chuẩn tham chiếu trong nhiều hệ thống khác nhau.

Việc quản lý tập trung và chuẩn hóa các loại dữ liệu này giúp giảm trùng lặp, nâng cao chất lượng dữ liệu, đồng thời tạo nền tảng cho việc tích hợp, chia sẻ và khai thác dữ liệu trong toàn Bộ.

c) Dữ liệu tổng hợp và phân tích:

- Dữ liệu tổng hợp là tập hợp dữ liệu được hình thành thông qua quá trình xử lý, thống kê và gộp nhóm từ các dữ liệu giao dịch (Dữ liệu phát sinh từ hoạt động nghiệp vụ, xử lý hồ sơ hằng ngày của Bộ).

- Dữ liệu phân tích là các thông tin phân tích nâng cao (Inferred data), tức là những tri thức được suy diễn hoặc khai phá từ dữ liệu gốc thông qua việc áp dụng các mô hình Trí tuệ nhân tạo và Học máy (AI/ML).

d) Dữ liệu mở:

Dữ liệu mở là các tập dữ liệu được cơ quan nhà nước có thẩm quyền công bố rộng rãi, cho phép tổ chức, cá nhân khai thác, sử dụng và tái sử dụng theo quy định. Trong Bộ KH&CN, dữ liệu mở có thể bao gồm dữ liệu về tiêu chuẩn, sáng chế, kết quả nghiên cứu, cũng như các dữ liệu thống kê đã được xử lý.

Trước khi công bố, dữ liệu được thực hiện phân loại, kiểm tra và chuẩn hóa nhằm bảo đảm tuân thủ các quy định về bảo mật, quyền riêng tư và an toàn thông tin. Việc cung cấp dữ liệu mở góp phần tăng cường minh bạch, thúc đẩy đổi mới sáng tạo và hỗ trợ hình thành hệ sinh thái khai thác dữ liệu trong lĩnh vực khoa học và công nghệ.

đ) Thu thập và nạp dữ liệu:

- ETL/ELT Pipelines: Thực hiện trích xuất, chuyển đổi và nạp dữ liệu theo lô (batch), phục vụ tổng hợp dữ liệu định kỳ (ví dụ: tổng hợp báo cáo hoạt động KH&CN theo ngày/tháng).

- CDC (Change Data Capture): Ghi nhận và đồng bộ các thay đổi dữ liệu từ các hệ thống nguồn theo thời gian gần thực (ví dụ: khi hồ sơ nhiệm vụ KH&CN hoặc dữ liệu sở hữu trí tuệ được cập nhật, dữ liệu sẽ được đồng bộ về nền tảng dữ liệu).

- Streaming/Event-based processing: Hỗ trợ tiếp nhận và xử lý dữ liệu theo sự kiện nhằm phục vụ các nhu cầu giám sát và phân tích theo thời gian gần thực.

e) Data Lakehouse

- Đóng vai trò là phần trung tâm của mô hình quản trị dữ liệu, nơi trực tiếp tiếp nhận dữ liệu được tích hợp từ các hệ thống chuyên ngành, dữ liệu nội bộ và các nguồn dữ liệu khác của Bộ sau khi đi qua lớp thu thập và tích hợp (LGSP/ETL/CDC).

- Tại đây, dữ liệu được lưu trữ, xử lý làm sạch, chuẩn hóa và tổ chức phân tầng chặt chẽ thành các lớp để phục vụ các mục tiêu tổng hợp, khai thác và chia sẻ khác nhau.

- Các thành phần cấu trúc cốt lõi bao gồm:

- + Dữ liệu thô (Bronze),

- + Dữ liệu chuẩn hóa (Silver),

+ Dữ liệu phân tích (Gold), giúp bảo đảm chất lượng dữ liệu và tối ưu khả năng khai thác.

+ Event bus/CDC hub: Trục sự kiện và trung tâm ghi nhận dữ liệu thay đổi, hỗ trợ đồng bộ dữ liệu từ các hệ thống nguồn theo thời gian gần thực.

+ Data Catalog & Metadata: Quản lý danh mục tài sản dữ liệu và siêu dữ liệu, hỗ trợ khám phá, tra cứu và hiểu rõ cấu trúc dữ liệu toàn Bộ.

+ Data Lineage: Hệ thống truy vết dữ liệu, cho phép theo dõi nguồn gốc và toàn bộ quá trình biến đổi của dữ liệu từ nguồn đến đầu ra.

g) AI/ML Platform

- Được tích hợp trực tiếp trên nền tảng dữ liệu dùng chung, bảo đảm kho dữ liệu của Bộ không chỉ là nơi lưu trữ tĩnh mà còn là nền tảng sinh ra giá trị gia tăng từ dữ liệu.

- Nền tảng thực hiện chức năng tổ chức và chuẩn bị dữ liệu chuyên sâu để phục vụ các bài toán phân tích nâng cao, dự báo, và triển khai các mô hình công nghệ hiện đại như RAG (Retrieval-Augmented Generation), NL2SQL, Knowledge Graph, VectorDB.

- Các thành phần cấu trúc cốt lõi bao gồm:

+ LLM (Mô hình ngôn ngữ lớn): Cung cấp năng lực xử lý ngôn ngữ tự nhiên, phân tích văn bản và tự động hóa các tác vụ liên quan đến dữ liệu văn bản chuyên ngành.

+ AI agent orchestrator (Trình điều phối tác tử AI): Quản lý, định tuyến và điều phối các luồng công việc tự động được thực hiện bởi các trợ lý ảo/tác tử trí tuệ nhân tạo.

+ AI Gateway: Cổng giao tiếp an toàn, làm nhiệm vụ quản lý, kiểm soát và tối ưu hóa các yêu cầu truy xuất đến các dịch vụ trí tuệ nhân tạo và mô hình học máy.

1.3. Dịch vụ dữ liệu

Dịch vụ dữ liệu là lớp trung gian cung cấp các dịch vụ dữ liệu và trí tuệ nhân tạo dùng chung, đóng vai trò kết nối giữa nền tảng dữ liệu tại và các ứng dụng khai thác dữ liệu. Dịch vụ dữ liệu không trực tiếp lưu trữ dữ liệu, cũng không phải là nơi người dùng cuối tương tác, mà tập trung đóng gói các năng lực dữ liệu,

phân tích và AI thành các dịch vụ chuẩn hóa để các hệ thống và ứng dụng có thể khai thác thông qua API hoặc các cơ chế dịch vụ dùng chung.

a) Dịch vụ dữ liệu tổng hợp và phân tích:

Khối này cung cấp các dịch vụ khai thác dữ liệu ở mức cao hơn, phục vụ phân tích, tổng hợp, dự báo và ứng dụng trí tuệ nhân tạo. Các dịch vụ như BI/Dashboard, báo cáo thống kê, dịch vụ AI/ML, phân tích dữ liệu, dự báo và chatbot/trợ lý AI đều được triển khai trong khối này trên cơ sở dữ liệu và nền tảng AI của Tầng 2.

Thông qua các dịch vụ này, dữ liệu được chuyển hóa thành các chỉ số, báo cáo, mô hình phân tích và công cụ hỗ trợ ra quyết định, phục vụ lãnh đạo Bộ, các đơn vị chuyên môn và cán bộ trong quá trình xử lý công việc. Đây là lớp giúp biến dữ liệu thành các dịch vụ có giá trị khai thác thực tế, đồng thời tạo điều kiện để các ứng dụng phía trên có thể tích hợp các năng lực phân tích và AI một cách thống nhất.

b) Dịch vụ dữ liệu mở:

Khối dịch vụ dữ liệu mở cung cấp các dịch vụ và giao diện truy cập đối với các tập dữ liệu được phép công bố ra bên ngoài. Các dữ liệu này được cung cấp dưới dạng dịch vụ để bảo đảm khả năng khai thác, tái sử dụng và tích hợp cho các tổ chức, doanh nghiệp, nhà nghiên cứu và nhà phát triển.

Khối này là cầu nối giữa nền tảng dữ liệu của Bộ với hệ sinh thái bên ngoài, góp phần thúc đẩy minh bạch, đổi mới sáng tạo và phát triển các ứng dụng dựa trên dữ liệu khoa học và công nghệ.

1.4. Tích hợp, chia sẻ và điều phối dữ liệu

a) Tích hợp, chia sẻ dữ liệu của Bộ (LGSP/LDOP):

- Là trực tích hợp trung tâm của Bộ KH&CN, thực hiện kết nối, chia sẻ và điều phối dữ liệu giữa các hệ thống thông tin và cơ sở dữ liệu trong nội bộ Bộ KH&CN.

- Triển khai các thành phần như API Gateway, Service Mesh, Service Registry, Information Mediator để quản lý, định tuyến và chuẩn hóa các dịch vụ dữ liệu.

- Quản lý tập trung danh mục API, bảo đảm việc cung cấp và khai thác dịch vụ dữ liệu theo cơ chế thống nhất, kiểm soát được truy cập và sử dụng

b) An toàn, bảo mật và giám sát dữ liệu:

- Data Access Control: Kiểm soát truy cập dữ liệu theo vai trò, phân quyền theo hệ thống và người dùng.

- Log, Audit & Monitoring: Ghi nhận nhật ký truy cập, giám sát và kiểm toán việc khai thác dữ liệu nhằm phát hiện sớm các bất thường.

- Masking & Encryption: Áp dụng các kỹ thuật che giấu và mã hóa dữ liệu trong quá trình trao đổi, bảo đảm an toàn thông tin và tuân thủ quy định về bảo vệ dữ liệu.

c) Kết nối, chia sẻ bên ngoài:

- Thực hiện kết nối với các bộ, ngành, địa phương và các hệ thống đối tác thông qua các giao diện API và cơ chế tích hợp tiêu chuẩn.

- Kết nối với Nền tảng tích hợp, chia sẻ dữ liệu quốc gia (NDXP) và Nền tảng chia sẻ, điều phối dữ liệu quốc gia (NDOP) để trao đổi và khai thác dữ liệu liên ngành.

- Thành phần Agent Node (điểm kết nối bảo mật) được triển khai tại Bộ KH&CN để bảo đảm an toàn, kiểm soát và chuẩn hóa việc trao đổi dữ liệu giữa hệ thống của Bộ KH&CN với các nền tảng quốc gia, Trung tâm dữ liệu quốc gia.

1.5. Khai thác và ứng dụng dữ liệu

Khai thác và ứng dụng dữ liệu là lớp tiếp xúc trực tiếp với người dùng cuối, cung cấp các ứng dụng và công cụ khai thác dữ liệu phục vụ điều hành, tác nghiệp và cung cấp dịch vụ. Tầng này sử dụng các dịch vụ dữ liệu và AI để triển khai các ứng dụng phục vụ lãnh đạo, cán bộ, người dân và doanh nghiệp.

a) Ứng dụng điều hành, giám sát:

- Dashboard điều hành: Cung cấp các bảng điều khiển tổng hợp phục vụ lãnh đạo Bộ theo dõi tình hình hoạt động KH&CN, đổi mới sáng tạo, chuyển đổi số.

- Theo dõi KPI, chỉ số: Tổng hợp và hiển thị các chỉ số đánh giá hiệu quả hoạt động của các đơn vị, chương trình, nhiệm vụ.

- Giám sát và cảnh báo: Thiết lập các cơ chế cảnh báo tự động dựa trên dữ liệu nhằm phát hiện sớm các vấn đề trong quá trình thực hiện nhiệm vụ.

b) Ứng dụng phân tích và dự báo:

- Phân tích dữ liệu: Cung cấp công cụ phân tích chuyên sâu trên dữ liệu tổng hợp, hỗ trợ đánh giá xu hướng trong lĩnh vực khoa học và công nghệ.
- Báo cáo phân tích: Xây dựng các báo cáo đa chiều phục vụ công tác quản lý, hoạch định chính sách.
- Dự báo, mô phỏng: Ứng dụng các mô hình phân tích và AI để dự báo xu hướng công nghệ, đổi mới sáng tạo và hỗ trợ ra quyết định.

c) Ứng dụng tác nghiệp nội bộ:

- Không gian làm việc số: Tích hợp các công cụ hỗ trợ xử lý công việc, quản lý hồ sơ và trao đổi thông tin trong nội bộ Bộ.
- Tra cứu dữ liệu dùng chung: Cho phép cán bộ truy cập, khai thác dữ liệu tập trung phục vụ công việc chuyên môn.
- Trợ lý AI nội bộ: Hỗ trợ tìm kiếm, tổng hợp thông tin, tra cứu văn bản và hỗ trợ xử lý công việc dựa trên dữ liệu.

d) Ứng dụng phục vụ người dân và doanh nghiệp:

- Dịch vụ công trực tuyến: Cung cấp các dịch vụ công trong lĩnh vực KH&CN.
- Tra cứu và tương tác: Cho phép người dân, doanh nghiệp tra cứu thông tin, gửi yêu cầu và tương tác với cơ quan quản lý.
- Ứng dụng di động: Hỗ trợ truy cập dịch vụ và thông tin trên các nền tảng di động.

đ) Công dữ liệu mở và công nhà phát triển:

- Công dữ liệu mở: Công bố các tập dữ liệu mở của Bộ, hỗ trợ tra cứu và tải dữ liệu.
- API công bố: Cung cấp các API để các tổ chức, doanh nghiệp và nhà phát triển khai thác dữ liệu.
- Dataset và tài liệu tích hợp: Cung cấp tài liệu hướng dẫn, mô tả dữ liệu và công cụ hỗ trợ tích hợp, phục vụ xây dựng các ứng dụng dựa trên dữ liệu KH&CN.

1.6. Quản trị và quản lý dữ liệu

Thành phần quản trị gồm 6 trụ cột, áp dụng xuyên suốt toàn bộ các lớp dữ liệu (không phải một tầng riêng trong kiến trúc phân tầng mà là khung quản trị bao trùm):

a) Chiến lược và Chính sách:

Ban hành và duy trì chiến lược dữ liệu của Bộ KH&CN, khung quản trị dữ liệu, các chính sách, tiêu chuẩn kỹ thuật và quy trình liên quan.

Thiết lập vai trò Cán bộ bảo vệ dữ liệu (DPO) và kiến trúc sư trưởng doanh nghiệp (Enterprise Architect) trong Bộ.

b) Danh mục dữ liệu (Data Catalog):

- Xây dựng và duy trì danh mục tài sản dữ liệu toàn Bộ; đăng ký quyền sở hữu và trách nhiệm quản lý cho từng tập dữ liệu.

- Công cụ khám phá và tra cứu dữ liệu giúp cán bộ tìm kiếm và hiểu rõ dữ liệu sẵn có trong toàn hệ thống.

c) Từ điển dữ liệu

- Xây dựng và duy trì từ điển dữ liệu dùng chung, chuẩn hóa định nghĩa và thuật ngữ nghiệp vụ.

- Quản lý Đăng ký cấu trúc dữ liệu (Schema Registry) đảm bảo tính nhất quán khi trao đổi dữ liệu giữa các hệ thống.

d) Truy vết dữ liệu (Data Lineage)

- Theo dõi nguồn gốc và quá trình biến đổi dữ liệu từ nguồn đến đầu ra.

- Khả năng truy vấn trạng thái dữ liệu tại bất kỳ thời điểm trong quá khứ (Time Travel), phục vụ kiểm toán, thanh tra và giải trình.

đ) Chất lượng dữ liệu

- Thiết lập hợp đồng dữ liệu (Data Contract) bắt buộc: mọi CSDL đưa vào kho tập trung phải có hợp đồng định nghĩa tiêu chuẩn chất lượng.

- Cổng kiểm tra chất lượng (DQ Gates) tự động: kiểm tra dữ liệu theo các luật chất lượng trước khi dữ liệu được lên tầng Chuẩn hóa và Phân tích.

- Đo lường chỉ số chất lượng dữ liệu định kỳ; quy trình ứng phó và xử lý sự cố chất lượng dữ liệu rõ ràng.

e) Vận hành và Giám sát:

- Giám sát hạ tầng và pipeline dữ liệu 24/7; cảnh báo sự cố tự động; báo cáo tình trạng vận hành định kỳ.

- Kiểm soát chi phí tài nguyên tính toán (FinOps): theo dõi và tối ưu hóa chi phí vận hành hạ tầng dữ liệu.

- Giám sát chất lượng dịch vụ AI: theo dõi độ chính xác, độ trễ và tỷ lệ lỗi của các ứng dụng trí tuệ nhân tạo.

g) Mô hình quản lý vòng đời dữ liệu:

Mô hình 8 giai đoạn đảm bảo dữ liệu được khai thác hiệu quả, an toàn xuyên suốt vòng đời:

GD	Tên giai đoạn	Nội dung	Ví dụ KH&CN
1	Thu thập / Tạo lập	Xác định nguồn, thu thập hoặc tạo lập từ hệ thống, cảm biến, TTHC	Nhà KH đăng ký đề tài qua Cổng MST; cảm biến bức xạ gửi dữ liệu IoT
2	Xử lý trước lưu trữ	Làm sạch, chuẩn hóa, loại bỏ trùng lặp, chuyển đổi định dạng	Chuẩn hóa tên tổ chức KH&CN theo mã MDM; loại bỏ bản ghi trùng nhà KH
3	Lưu trữ và quản lý	Lưu vào CSDL/Lakehouse, phân quyền truy cập, phân tầng Bronze-Silver-Gold	Hồ sơ đề tài vào Bronze; sau DQ Gates lên Silver; tổng hợp vào Gold
4	Tích hợp dữ liệu	Kết nối từ nhiều nguồn, đồng bộ giữa các hệ thống qua CDC/ETL	Đồng bộ DL SHTT từ Cục SHTT về Lakehouse; tích hợp DL 34 Sở

GD	Tên giai đoạn	Nội dung	Ví dụ KH&CN
5	Phân tích và sử dụng	Khai thác cho báo cáo, Dashboard, AI/ML dự đoán	Dashboard KPI nhiệm vụ KH&CN cho lãnh đạo; AI dự báo xu hướng công nghệ
6	Khai thác và Chia sẻ	Xuất báo cáo, API, chia sẻ qua LGSP/NDOP	Cung cấp API thống kê KH&CN cho Bộ KH-ĐT; chia sẻ DL chủ cho TTDL QG
7	Lưu trữ dài hạn	Lưu trữ lịch sử theo chính sách, tuân thủ pháp luật	Lưu trữ hồ sơ đề tài đã nghiệm thu 10 năm theo quy định
8	Hủy bỏ dữ liệu	Xóa bỏ DL hết giá trị, bảo đảm không thể khôi phục	Xóa DL giao dịch quá thời hạn; hủy DL cá nhân theo yêu cầu rút đồng ý

2. Mô hình tham chiếu dữ liệu

DỮ LIỆU BỘ KHOA HỌC VÀ CÔNG NGHỆ												
	DỮ LIỆU DỊCH VỤ NGƯỜI ĐÀN, DOANH NGHIỆP	DỮ LIỆU NGHIỆP VỤ KH&CN	DỮ LIỆU NGHIỆP VỤ ĐMST	DỮ LIỆU NGHIỆP VỤ CBS & CN CNS	DỮ LIỆU NGHIỆP VỤ SỞ HỮU TRÍ TUỆ	DỮ LIỆU NGHIỆP VỤ TC&CL	DỮ LIỆU NGHIỆP VỤ VT, TS, INTERNET	DỮ LIỆU NGHIỆP VỤ NLNT, ATBX	DỮ LIỆU NGHIỆP VỤ BƯU CHÍNH	THẺ CHẾ, CHÍNH SÁCH, NGUỒN LỰC...	DỮ LIỆU DỊCH VỤ QL, CHỈ ĐẠO ĐIỀU HÀNH	LỚP DỮ LIỆU
DỮ LIỆU (DUNG CHUNG) QUỐC GIA TẠI TTDLQG	Danh sách TTHC, DVC quốc gia											
	Danh mục thủ tục hành chính, Danh mục dịch vụ công...	Danh mục nhiệm vụ KH&CN, Kết quả nghiên cứu công bố...	Danh mục thời gian nghiệp vụ sáng tạo, Danh mục tổ chức hỗ trợ ĐMST...	Danh mục nền tảng số, Danh mục doanh nghiệp công nghệ số...	Dữ liệu tra cứu đơn đăng ký, Dữ liệu tra cứu văn bằng bản quyền...	Danh mục TC, QD, Danh mục tổ chức đánh giá và phí hợp...	Danh mục tài nguyên Internet, Danh mục tên miền...	Danh mục cơ sở được công khai, Danh mục giấy phép công khai...	Danh mục doanh nghiệp bưu chính, Danh mục điểm phục vụ...	Văn bản quy phạm pháp luật, Văn bản chỉ đạo điều hành...	Chỉ tiêu công khai, Báo cáo công khai, Thông kê kết quả thực hiện nhiệm vụ...	DỮ LIỆU MỞ
DL tổng hợp quốc gia, DL phục vụ chỉ đạo điều hành của CP, TTg	Tỷ lệ giải quyết đúng hạn, Tỷ lệ hồ sơ trực tuyến, Mức độ hài lòng...	Thông kê nhiệm vụ KH&CN, Phân tích kết quả nghiên cứu	Thông kê đối tượng, DN thời gian nghiệp vụ và DN ĐMST, Báo cáo hệ thống thời ĐMST	Chỉ số chuyển đổi số (DIT), Thông kê nền tảng số...	Thông kê số đơn đăng ký, Tỷ lệ xử lý hồ sơ, Phân tích xu hướng đăng ký SHTT	Thông kê TC, QC, Báo cáo hoạt động kiểm định, hiệu chuẩn...	Thông kê sử dụng băng tần, Thông kê tài nguyên Internet...	Báo cáo ATBX, Thông kê thành tích kiểm tra, Phân tích mức độ tuân thủ...	Thông kê DN bưu chính, Báo cáo mạng lưới điểm phục vụ...	Văn bản QPPL, Văn bản chỉ đạo điều hành, Danh mục chương trình, đề án...	Chỉ tiêu công khai, Báo cáo công khai, Thông kê kết quả thực hiện nhiệm vụ...	DỮ LIỆU TỔNG HỢP (biểu đồ, bảng hợp, phân tích)
DL phục vụ tiếp nhận và xử lý TTHC, DL điều phối, quản trị dữ liệu	Phiếu tiếp nhận/hiến tại kết quả, Ngày tiếp nhận, Trang thái xử lý hồ sơ...	Ngày tiếp nhận, Trang thái nhiệm vụ, Trạng thái thực hiện...	Trạng thái hoạt động, Số hiệu Giấy chứng nhận đăng ký doanh nghiệp...	Trạng thái văn bản, Ngày cấp (giấy phép), Số hiệu...	Số đơn, Ngày nộp đơn, Số bằng, Ngày cấp bằng...	Tình trạng, Văn bản công bố bản hành, Năm công bố...	Ngày cấp phép, Ngày lập Báo cáo vi phạm, Tần số vi phạm...	Số giấy phép, Ngày, tháng, năm cấp giấy phép, Thời hạn của Giấy đăng ký...	Tình hình hoạt động, Ngày cấp phép, Thời hạn giấy phép...	Kiểm góp ý, Kết quả chấm điểm, Phân bản dự thảo, Quyết định ban hành...	Phiếu giao nhiệm vụ, Cập nhật tiến độ, Báo cáo định kỳ/đột xuất...	DỮ LIỆU GIAO DỊCH (nghiệp vụ chuyên ngành)
DLOG về tổ chức, DLOG về con người dân cư DN, DLOG khác (TTHC, địa danh...)	Mã số doanh nghiệp, Tên doanh nghiệp, số chức, Văn đặc lệ...	Mã số nhiệm vụ khoa học, Tên nhiệm vụ khoa học...	Mã số doanh nghiệp, Tên doanh nghiệp, Tỉnh...	Mã số doanh nghiệp, Tên CA	Số đơn, Ngày nộp đơn, Số bằng, Ngày cấp bằng, Tên SHCN	Mã GTN, Số hiệu, Tên QC VN, Số hiệu tiêu chuẩn...	Tỷ lệ kết nối, Tỷ lệ kết nối khác cơ quan Nhà nước...	ID giấy phép, Mã nhân viên học xã, Mã cơ sở dịch vụ...	Mã số doanh nghiệp bưu chính, Tên doanh nghiệp bưu chính...	Số văn bản, Loại văn bản, Cơ quan ban hành, Trạng thái hiệu lực...	Mã định danh cơ quan/đơn vị, Loại văn bản chỉ đạo...	DỮ LIỆU CHỦ
Dữ liệu danh mục đang chung quốc gia	Tên TTHC, Mã TTHC, Cơ quan gửi quyết, Họ và tên Chủ hồ sơ...	Mã nhiệm vụ, Tên nhiệm vụ, Mã hồ sơ, Tên hồ sơ, Cấp quản lý...	Mã doanh nghiệp, Tên doanh nghiệp, Loại hình startup...	Tên hệ thống, Loại nền tảng, Cấp độ hệ thống, Đơn vị quản lý...	Loại đối tượng SHTT, Loại đơn, Tên SHCN, Chủ sở hữu...	Mã GTN, Tên sản phẩm, Nhóm sản phẩm, Mã thị trường mục tiêu...	Dữ liệu số, Loại giấy phép, Năm báo cáo, Tháng báo cáo...	Mã hồ sơ, Tên hồ sơ, Tên tổ chức cấp giấy đăng ký...	Tên doanh nghiệp, Mã số doanh nghiệp, Tình hình hoạt động...	Văn bản quy phạm pháp luật, Văn bản chỉ đạo điều hành...	Chỉ tiêu thống kê, Báo cáo thống kê, Biểu mẫu báo cáo...	DỮ LIỆU CỐC
Từ dữ liệu quốc gia, SDL, kiểm toán dữ liệu...	Tên TTHC, Mã TTHC, Cơ quan gửi quyết, Trạng thái hồ sơ...	Mã nhiệm vụ KH&CN, Tên nhiệm vụ, Mã số nhiệm vụ KH&CN	Mã doanh nghiệp, Tên doanh nghiệp, Loại hình startup...	Tên hệ thống, Loại nền tảng, Cấp độ hệ thống, Đơn vị quản lý...	Số đơn, Loại đối tượng SHTT, Ngày nộp đơn, Số bằng...	Mã GTN, Tên sản phẩm, Mã số, Nhóm sản phẩm, Tên doanh nghiệp...	Dữ liệu số, Loại giấy phép, Ngày cấp phép, Tên miền...	Mã hồ sơ, Tên hồ sơ, Số giấy phép, Ngày, tháng, năm cấp giấy phép...	Tên doanh nghiệp, Mã số doanh nghiệp, Tình hình hoạt động...	Số văn bản, Loại văn bản, Cơ quan ban hành, Ngày hiệu lực...	Mã nhiệm vụ, Đơn vị thực hiện, Thời hạn, Trạng thái, KPI liên quan...	SIÊU DỮ LIỆU (dữ liệu đặc tả)

Hình 2. Mô hình tham chiếu dữ liệu

Mô hình dữ liệu tham chiếu thể hiện các miền dữ liệu được khái quát hóa từ nghiệp vụ, các hoạt động và các lớp dữ liệu. Dữ liệu của Bộ KH&CN được phân theo 11 nhóm nghiệp vụ tương ứng với 11 miền dữ liệu như trên mô hình:

STT	Nhóm nghiệp vụ	Mô tả về dữ liệu
1	Dữ liệu phục vụ người dân, doanh nghiệp	Trọng tâm là dịch vụ công (DVC) và thủ tục hành chính (TTHC). Bao gồm: Danh sách TTHC, DVC quốc gia; dữ liệu mở về hướng dẫn dịch vụ; dữ liệu tổng hợp về hồ sơ; dữ liệu giao dịch xử lý hồ sơ TTHC; định danh người dân/doanh nghiệp; danh mục liên quan DVC/TTHC; Siêu dữ liệu phục vụ việc kiểm toán dữ liệu này.
2	Dữ liệu nghiệp vụ Khoa học công nghệ	Dữ liệu về hoạt động Khoa học, Công nghệ. Bao gồm: Nhiệm vụ KH&CN, kết quả nghiên cứu (mở); tổng hợp về nhân lực, tổ chức KH&CN; dữ liệu giao dịch về hoạt động nghiệp vụ, kiểm tra, đánh giá; dữ liệu chủ về nhà khoa học, tổ chức, nhiệm vụ; danh mục loại nhiệm vụ; và siêu dữ liệu tương ứng.
3	Dữ liệu nghiệp vụ Đổi mới sáng tạo	Là tập hợp các dữ liệu được tạo ra, thu thập, quản lý và khai thác trong quá trình tổ chức, vận hành, quản lý và hỗ trợ hoạt động đổi mới sáng tạo, phục vụ cho việc ra quyết định, điều hành, đánh giá và phát triển hệ sinh thái ĐMST.
4	Dữ liệu nghiệp vụ Chuyển đổi số và công nghiệp công nghệ số	Dữ liệu về Chuyển đổi số và Công nghiệp Công nghệ số. Bao gồm: Nền tảng số, DN công nghệ số (mở); tổng hợp về CĐS, công nghiệp số; dữ liệu giao dịch quản lý nền tảng, hệ thống số; dữ liệu chủ nền tảng, DN công nghệ số; danh mục nền tảng, hệ thống số; và siêu dữ liệu.

5	Dữ liệu nghiệp vụ Sở hữu trí tuệ	Quản lý bằng sáng chế, nhãn hiệu. Bao gồm: Sáng chế, nhãn hiệu, kiểu dáng (mở); tổng hợp đơn đăng ký, văn bằng; dữ liệu giao dịch tiếp nhận, thẩm định, xử lý đơn; dữ liệu chủ về đơn, bằng, chủ thể quyền; danh mục đối tượng, trạng thái đơn; và siêu dữ liệu.
6	Dữ liệu nghiệp vụ Tiêu chuẩn đo lường chất lượng	Quản lý quy chuẩn, tiêu chuẩn kỹ thuật. Bao gồm: TCVN, QCVN (mở); tổng hợp tiêu chuẩn, đo lường, chất lượng (CL); dữ liệu giao dịch đánh giá hợp chuẩn, hợp quy, kiểm định, hiệu chuẩn; dữ liệu chủ TCVN, QCVN, chuẩn đo lường, PT (phương tiện) đo; danh mục loại tiêu chuẩn, đo lường, đánh giá CL; và siêu dữ liệu.
7	Dữ liệu nghiệp vụ Viễn thông, tần số, Internet	Quản lý hạ tầng thông tin và truyền thông. Bao gồm: Tên miền, tài nguyên Internet (mở); tổng hợp về VT (viễn thông), TS (tần số), Internet, BC (bưu chính); dữ liệu giao dịch cấp phép, phân bổ, thanh tra, giám sát; dữ liệu chủ tên miền, IP, ASN, mạng lưới, DN; danh mục tài nguyên số, loại giấy phép; và siêu dữ liệu.
8	Dữ liệu nghiệp vụ Năng lượng nguyên tử và ATBX	Là tập hợp các thông tin, số liệu được tạo lập, thu thập, cập nhật và quản lý trong quá trình thực hiện chức năng quản lý nhà nước, hoạt động chuyên môn, nghiệp vụ về năng lượng nguyên tử, bao gồm: thông tin về cấp phép, nguồn phóng xạ, cơ sở bức xạ – hạt nhân, an toàn bức xạ, an ninh hạt nhân, thanh tra, giám sát và ứng phó sự cố...

9	Dữ liệu nghiệp vụ Bưu chính	Là tập hợp thông tin, dữ liệu phát sinh trong quá trình cung ứng và sử dụng dịch vụ bưu chính, phục vụ cho hoạt động quản lý, khai thác, vận hành, giám sát và đảm bảo an toàn trong lĩnh vực bưu chính...
10	Dữ liệu thể chế, chính sách, nguồn lực	Là tập hợp dữ liệu phản ánh hệ thống văn bản pháp luật, các chủ trương – chính sách của Nhà nước và các nguồn lực (tài chính, nhân lực, hạ tầng, tổ chức) phục vụ quản lý, điều hành và phát triển kinh tế - xã hội.
11	Dữ liệu phục vụ quản lý, chỉ đạo điều hành	Dữ liệu cho lãnh đạo Bộ. Bao gồm: Chỉ tiêu, báo cáo, thống kê (mở); tổng hợp Dashboard, KPI, cảnh báo, dự báo; dữ liệu giao dịch theo dõi nhiệm vụ chỉ đạo, tiến độ, cảnh báo; dữ liệu chủ nhiệm vụ, chỉ tiêu, chủ đề, khung chỉ số, KPI; danh mục chỉ tiêu báo cáo, thống kê; và siêu dữ liệu phục vụ chỉ đạo, điều hành.

3. Danh mục dữ liệu

3.1. Danh mục cơ sở dữ liệu chuyên ngành

Danh mục các cơ sở dữ liệu chuyên ngành được quy định và ban hành theo Quyết định về Danh mục cơ sở dữ liệu chuyên ngành thuộc phạm vi quản lý của Bộ Khoa học và Công nghệ, để bảo đảm tính linh hoạt trong việc bổ sung, cập nhật và chuẩn hóa dữ liệu theo đặc thù của từng lĩnh vực quản lý.

3.2. Dữ liệu dùng chung

STT	Loại dữ liệu	Tên Cơ sở dữ liệu	Mô tả
1	Siêu dữ liệu	CSDL Siêu dữ liệu	Dữ liệu dùng để mô tả, định nghĩa và quản lý các

			lớp dữ liệu khác. Cung cấp thông tin về nguồn gốc, cấu trúc, quy tắc chất lượng và điều kiện khai thác, giúp thống nhất quản trị và chia sẻ dữ liệu trong toàn Bộ. Chi tiết được thể hiện tại mục Siêu dữ liệu tại Phụ lục kèm theo.
2	Danh mục dùng chung	CSDL Danh mục dùng chung	Là dữ liệu về các danh mục, bảng mã phân loại do Bộ thống nhất ban hành. Dữ liệu danh mục dùng chung được sử dụng thống nhất, đồng bộ và tương thích giữa các hệ thống thông tin từ Bộ đến Trung ương và đến các địa phương. Chi tiết xem tại mục Danh mục dùng chung tại Phụ lục kèm theo.
3	Dữ liệu chủ	CSDL chủ	Thông tin gốc, ổn định về các đối tượng quản lý trọng tâm của Bộ KH&CN. Dùng làm chuẩn tham chiếu chung để đảm bảo tính đồng nhất dữ liệu giữa các hệ thống chuyên ngành và phục vụ phân tích, điều hành. Chi tiết xem tại

			mục Dữ liệu chủ tại Phụ lục kèm theo.
4	Dữ liệu giao dịch (Dữ liệu chuyên ngành)	CSDL giao dịch.	Dữ liệu phát sinh liên tục từ các hoạt động nghiệp vụ, xử lý hồ sơ và cung cấp dịch vụ công hàng ngày của Bộ KH&CN. Chi tiết xem tại mục Dữ liệu giao dịch tại Phụ lục kèm theo.
5	Dữ liệu gốc	CSDL gốc	Dữ liệu gốc là dữ liệu được tạo lập trong quá trình hoạt động của cơ quan, tổ chức, cá nhân hoặc thu thập, tạo lập từ số hóa bản chính giấy tờ, tài liệu, các dạng vật chất khác. Chi tiết xem tại mục Dữ liệu gốc tại Phụ lục kèm theo.
6	Dữ liệu phân tích, tổng hợp	CSDL phân tích, tổng hợp.	Dữ liệu t được hình thành từ quá trình tổng hợp, thống kê và mô hình hóa dữ liệu gốc (dữ liệu chủ và giao dịch). Chi tiết xem tại mục Dữ liệu

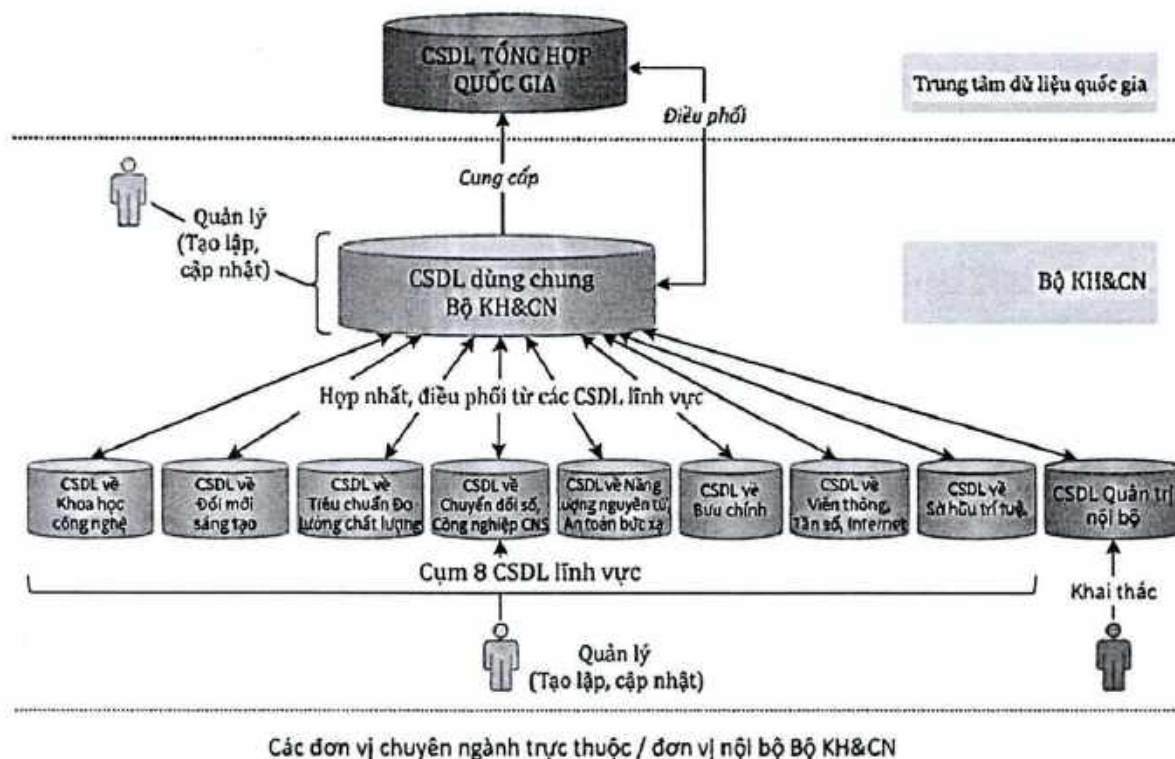
			phân tích tại Phụ lục kèm theo
7	Dữ liệu mở	CSDL dữ liệu mở.	Dữ liệu được chọn lọc từ kho dữ liệu của Bộ KH&CN, công bố công khai dưới định dạng máy có thể đọc được để tổ chức, cá nhân có thể tự do tiếp cận, tái sử dụng và chia sẻ. Chi tiết xem tại mục Dữ liệu mở tại Phụ lục kèm theo.

3.3. Dữ liệu ngoài ngành

Dữ liệu nhận từ bên ngoài về để phục vụ dữ liệu chuyên ngành của Bộ Khoa học và Công nghệ.

STT	Loại dữ liệu	Nguồn dữ liệu	Mô tả
1	Dữ liệu người dân	CSDL Dân cư	Một số thông tin chủ của người dân đã được xác thực CSDL Dân cư
2	Dữ liệu cán bộ viên chức	CSDL Cán bộ công nhân viên chức	Thông tin cán bộ viên chức thuộc Bộ KH&CN nguồn từ CSDL dùng chung Bộ Nội vụ
3	Dữ liệu tổ chức, doanh nghiệp	CSDL Quốc gia về doanh nghiệp	Thông tin dữ liệu chủ của doanh nghiệp có liên quan tới các nghiệp vụ. Nguồn từ CSDL dùng chung Bộ Tài chính

4. Mô hình tích hợp và điều phối dữ liệu



Hình 3. Mô hình tích hợp và điều phối dữ liệu Bộ KH&CN

- Mô hình quản lý dữ liệu dùng chung của Bộ KH&CN được thiết kế nhằm hình thành một nguồn dữ liệu thống nhất cấp Bộ cho các nhóm dữ liệu dùng chung, đặc biệt là dữ liệu chủ, dữ liệu danh mục dùng chung và dữ liệu tổng hợp phục vụ điều hành, đồng thời vẫn bảo đảm tính phân quyền, chủ động và bám nghiệp vụ đối với các đơn vị chuyên ngành trong quá trình tạo lập, cập nhật và duy trì dữ liệu.

- Khác với mô hình quản lý tập trung hoàn toàn, mô hình của Bộ KH&CN không yêu cầu mọi dữ liệu đều phải được tạo lập tại một đầu mối duy nhất. Ngược lại, dữ liệu vẫn được hình thành chủ yếu tại các hệ thống nghiệp vụ chuyên ngành, nơi gắn trực tiếp với quy trình nghiệp vụ và với các đơn vị có thẩm quyền quản lý dữ liệu. Tuy nhiên, để tránh phân tán, trùng lặp và không thống nhất giữa các hệ thống, Bộ xây dựng một nền tảng dữ liệu dùng chung làm lõi hợp nhất, chuẩn hóa và điều phối dữ liệu trên phạm vi toàn Bộ.

- Về bản chất, đây là mô hình phân quyền một phần: một phần dữ liệu được quản lý tại nguồn chuyên ngành, một phần dữ liệu được quản lý tập trung tại nền tảng dữ liệu dùng chung của Bộ, và giữa hai bên tồn tại cơ chế hợp nhất, đồng bộ và điều phối hai chiều. Cách tiếp cận này phù hợp với đặc thù của Bộ KH&CN là

có nhiều lĩnh vực quản lý khác nhau, mỗi lĩnh vực có hệ thống nghiệp vụ và cơ sở dữ liệu riêng, nhưng vẫn cần một bức tranh dữ liệu thống nhất để phục vụ điều hành, phân tích và kết nối Trung tâm dữ liệu quốc gia và các Bộ ngành khác.

Dữ liệu chuyên ngành và các nguồn dữ liệu của Bộ (nguồn tạo lập)

Đây là lớp hình thành dữ liệu ban đầu, cũng là nơi dữ liệu có tính nghiệp vụ sâu nhất được tạo lập và cập nhật. Theo mô hình Khung kiến trúc dữ liệu Bộ KH&CN, phần này không chỉ gồm các cơ sở dữ liệu chuyên ngành, mà là một khối nguồn dữ liệu đa dạng, chia thành ba nhóm lớn.

- Thứ nhất là dữ liệu chuyên ngành, bao gồm dữ liệu phát sinh từ các lĩnh vực quản lý nhà nước cốt lõi của Bộ như: khoa học và công nghệ; đổi mới sáng tạo; sở hữu trí tuệ; tiêu chuẩn, đo lường, chất lượng; chuyển đổi số và các lĩnh vực chuyên ngành khác. Đây là những dữ liệu phản ánh trực tiếp hoạt động nghiệp vụ của các đơn vị chuyên môn, ví dụ hồ sơ nhiệm vụ KH&CN, dữ liệu về tổ chức KH&CN, dữ liệu sáng chế/nhãn hiệu/kiểu dáng công nghiệp, dữ liệu tiêu chuẩn - quy chuẩn - đo lường, hay dữ liệu liên quan đến các chương trình, đề án, nhiệm vụ chuyển đổi số...

- Thứ hai là dữ liệu nội bộ, bao gồm tài chính, nhân sự, văn bản điều hành, kế hoạch, nhiệm vụ, tài sản và quản trị nội bộ... Mặc dù không phải dữ liệu chuyên ngành theo nghĩa hẹp, nhưng đây là nhóm dữ liệu rất quan trọng để hình thành bức tranh điều hành tổng thể của Bộ, đặc biệt khi xây dựng dashboard lãnh đạo, báo cáo tổng hợp, giám sát KPI hay trợ lý AI nội bộ. Các dữ liệu này không phục vụ chia sẻ ra ngoài mà hỗ trợ các hoạt động chỉ đạo, điều hành và nâng cao năng lực nội bộ của Bộ KH&CN.

- Thứ ba là nguồn dữ liệu khác, bao gồm dữ liệu từ các bộ, ngành, địa phương; dữ liệu từ viện nghiên cứu, trường đại học, doanh nghiệp và đối tác; cùng với dữ liệu công khai, dữ liệu mở và các nguồn Internet. Với Bộ KH&CN, nhóm này đặc biệt quan trọng vì nhiều bài toán phân tích và dự báo không thể chỉ dựa trên dữ liệu nội bộ của Bộ, mà cần kết hợp với dữ liệu ngoài Bộ để làm giàu dữ liệu, đối soát và nâng cao giá trị khai thác.

Theo nguyên tắc “dữ liệu được tạo lập một lần tại nguồn”, các đơn vị chuyên ngành chịu trách nhiệm tạo lập, cập nhật và bảo đảm chất lượng dữ liệu thuộc phạm vi mình quản lý. Nói cách khác, chất lượng dữ liệu trước hết phải được bảo đảm tại nơi phát sinh dữ liệu. Đây là điểm rất quan trọng, vì nếu dữ liệu nguồn

sai, thiếu hoặc không đồng nhất, thì dù nền tảng dữ liệu dùng chung có tốt đến đâu cũng chỉ hợp nhất được một tập dữ liệu có chất lượng hạn chế.

Dữ liệu dùng chung của Bộ (lỗi hợp nhất, chuẩn hóa và điều phối)

Đây là phần trung tâm của mô hình, tương ứng với thành phần Dữ liệu trong kiến trúc dữ liệu Bộ KH&CN theo thiết kế: khối dữ liệu chủ, dữ liệu danh mục dùng chung, khối dữ liệu tổng hợp và phân tích, cùng với Data Lakehouse và AI/ML Platform.

Về mặt quản lý dữ liệu, lớp này là nơi Bộ KH&CN hình thành “một nguồn dữ liệu thống nhất cấp Bộ”, không phải bằng cách thay thế hoàn toàn các hệ thống nguồn, mà bằng cách thu nhận, làm sạch, chuẩn hóa, hợp nhất, loại bỏ trùng lặp và tái cấu trúc dữ liệu từ nhiều nguồn khác nhau. Thông qua các cơ chế tích hợp như ETL/ELT, CDC/Data Streaming và các thành phần trung gian, dữ liệu được đổ vào nền tảng dữ liệu tập trung của Bộ và tổ chức thành các lớp Bronze, Silver, Gold để phục vụ các mục tiêu khai thác khác nhau.

Trong lớp này, Bộ KH&CN sẽ quản lý thống nhất các nhóm dữ liệu cốt lõi như dữ liệu chủ (Master data), dữ liệu danh mục dùng chung (Reference data). Đây chính là những nhóm dữ liệu mà Khung kiến trúc dữ liệu quốc gia đã khuyến nghị nên thiên về mô hình tập trung hoặc mô hình lai, thay vì để phân tán hoàn toàn ở các hệ thống chuyên ngành. Với Bộ KH&CN, điều đó có nghĩa là các khái niệm và đối tượng dùng chung như tổ chức KH&CN, danh mục lĩnh vực, nhiệm vụ KH&CN, các danh mục phân loại chuyên ngành, hay các tập dữ liệu thống kê dùng chung không nên để mỗi hệ thống hiểu một kiểu, mà cần có một nguồn chuẩn duy nhất cấp Bộ.

Điểm rất quan trọng của lớp này là nó không chỉ “lưu lại dữ liệu sau khi hợp nhất”, mà còn đóng vai trò điều phối dữ liệu ngược trở lại cho các hệ thống khai thác trong Bộ. Nghĩa là dữ liệu không chỉ đi từ dưới lên trên, mà còn đi từ trung tâm ra các hệ thống tác nghiệp, hệ thống báo cáo, dashboard, chatbot, trợ lý AI hay các cổng thông tin. Đây chính là điểm phân biệt của mô hình 03: một số dữ liệu vẫn được quản lý tại nguồn chuyên ngành, nhưng một số dữ liệu khác lại được quản lý tập trung tại hệ thống dùng chung và cung cấp ngược trở lại cho toàn hệ thống sử dụng.

Ngoài ra, Khung dữ liệu của Bộ đã tích hợp AI/ML Platform ngay trong lớp dữ liệu dùng chung của Bộ KH&CN không chỉ là kho lưu trữ, mà còn là nền tảng

để sinh ra giá trị gia tăng từ dữ liệu. Tại đây, dữ liệu được tổ chức để phục vụ phân tích nâng cao, dự báo, mô hình AI, RAG, NL2SQL, Knowledge Graph, VectorDB... Tức là lớp này vừa là “nơi thống nhất dữ liệu”, vừa là “nơi chuẩn bị dữ liệu” cho các dịch vụ dữ liệu và AI.

Kết nối, chia sẻ và liên thông quốc gia

Đây là lớp bảo đảm dữ liệu của Bộ KH&CN không tồn tại biệt lập, mà có thể tham gia vào hệ sinh thái dữ liệu quốc gia. Trong kiến trúc Bộ KH&CN của bạn, lớp này tương ứng với phần Tích hợp, chia sẻ và điều phối dữ liệu, nơi có LGSP/LDOP, API Gateway, Service Mesh, Information Mediator, Agent Node, và kết nối với NDXP/NDOP.

Từ nền tảng dữ liệu dùng chung của Bộ KH&CN, các tập dữ liệu cần thiết sẽ được lựa chọn, kiểm soát, đóng gói dưới dạng dịch vụ hoặc luồng dữ liệu và cung cấp ra bên ngoài thông qua cơ chế tích hợp, chia sẻ dữ liệu của Bộ. Việc kết nối với bên ngoài không thực hiện trực tiếp từ từng hệ thống chuyên ngành riêng lẻ, mà được điều phối qua lớp tích hợp của Bộ để bảo đảm thống nhất về tiêu chuẩn kỹ thuật, bảo mật, kiểm soát truy cập, ghi log và giám sát.

Thông qua LGSP/LDOP và Agent Node, Bộ KH&CN kết nối với NDOP và các nền tảng quốc gia, từ đó thực hiện chia sẻ, đồng bộ hoặc khai thác dữ liệu với Trung tâm dữ liệu quốc gia và các cơ quan khác trong hệ thống chính trị. Điều này làm cho nền tảng dữ liệu dùng chung của Bộ trở thành đầu mối liên thông giữa dữ liệu nội bộ Bộ với không gian dữ liệu quốc gia. Nói cách khác, Bộ KH&CN vừa giữ được tính chủ động trong quản trị dữ liệu của mình, vừa tuân thủ Khung kiến trúc dữ liệu quốc gia và cơ chế chia sẻ dữ liệu bắt buộc.

Dòng chảy dữ liệu trong mô hình

(1) Luồng hợp nhất từ dưới lên (Bottom-up)

Dữ liệu được tạo lập tại các hệ thống chuyên ngành, dữ liệu nội bộ và các nguồn dữ liệu khác của Bộ KH&CN, sau đó được tích hợp thông qua Tầng tích hợp (LGSP/ETL/CDC) để đưa vào nền tảng dữ liệu dùng chung (Data Lakehouse).

Tại đây, dữ liệu được xử lý qua các bước làm sạch, chuẩn hóa, hợp nhất và loại bỏ trùng lặp để hình thành các lớp dữ liệu có thể dùng chung, bao gồm dữ liệu chủ, dữ liệu danh mục và dữ liệu tổng hợp phục vụ phân tích. Trong quá trình

này, Bộ KH&CN áp dụng hai phương thức chia sẻ dữ liệu cơ bản nhằm bảo đảm vừa tính đồng bộ, vừa tính linh hoạt:

a) Chia sẻ dữ liệu chủ động (Push-based)

Cơ chế:

- Đơn vị chủ quản dữ liệu chủ động đóng gói và gửi dữ liệu về nền tảng dữ liệu dùng chung của Bộ ngay khi có thay đổi tại hệ thống nguồn (thêm mới, cập nhật, xóa), thông qua các cơ chế ETL, CDC hoặc streaming.

Đặc điểm:

- Dữ liệu được đồng bộ gần thời gian thực.
- Giảm phụ thuộc vào truy vấn trực tiếp hệ thống nguồn.
- Phù hợp với dữ liệu cần thống nhất trên toàn Bộ KH&CN.

Áp dụng cho:

- Dữ liệu chủ (Master data) như tổ chức KH&CN, nhiệm vụ KH&CN
- Dữ liệu danh mục dùng chung (lĩnh vực, phân loại, tiêu chuẩn...)
- Các dữ liệu nghiệp vụ cốt lõi cần tính đồng bộ cao

b) Chia sẻ dữ liệu theo yêu cầu (Pull-based / On-demand)

Cơ chế:

Hệ thống khai thác gửi yêu cầu truy vấn thông qua API hoặc dịch vụ dữ liệu; hệ thống nguồn hoặc dịch vụ trung gian trả về kết quả tương ứng.

Đặc điểm:

- Truy xuất dữ liệu theo nhu cầu.
- Hệ thống nguồn giữ quyền kiểm soát dữ liệu.
- Có kiểm soát truy cập và ghi nhật ký.

Áp dụng cho:

- Tra cứu dữ liệu chuyên ngành
- Xác thực thông tin
- Khai thác dữ liệu không thường xuyên
- Các báo cáo hoặc truy vấn ad-hoc

Hai phương thức này được sử dụng kết hợp để bảo đảm dữ liệu vừa được hợp nhất kịp thời tại nền tảng dùng chung, vừa bảo đảm tính linh hoạt và kiểm soát khi khai thác dữ liệu chuyên ngành.

(2) Luồng điều phối từ trung tâm ra các hệ thống sử dụng (Top-down)

Sau khi dữ liệu chủ, dữ liệu danh mục và dữ liệu chuyên ngành dùng chung được hình thành tại nền tảng dữ liệu của Bộ KH&CN, các dữ liệu này được cung cấp lại cho các hệ thống khai thác thông qua tầng dịch vụ dưới dạng API, dịch vụ dữ liệu hoặc dịch vụ phân tích.

Luồng này bảo đảm các hệ thống trong Bộ KH&CN sử dụng cùng một nguồn dữ liệu chuẩn, tránh tình trạng mỗi hệ thống xây dựng và duy trì một phiên bản dữ liệu riêng, từ đó nâng cao tính nhất quán và giảm trùng lặp dữ liệu.

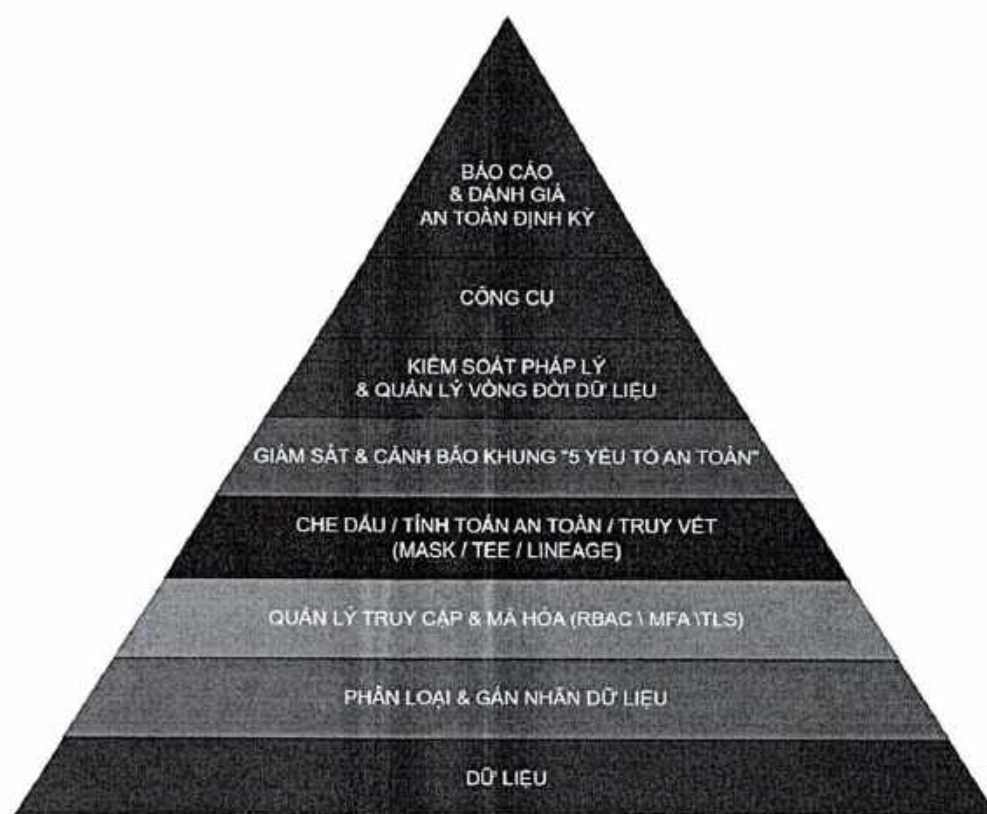
(3) Luồng liên thông và chia sẻ quốc gia (Cross-system / External)

Từ nền tảng dữ liệu dùng chung của Bộ KH&CN, các dữ liệu cần chia sẻ được điều phối ra bên ngoài thông qua phân tích hợp, sử dụng LGSP/LDOP, API Gateway và Agent Node để kết nối với các nền tảng quốc gia như NDXP và NDOP.

Thông qua đó, dữ liệu của Bộ KH&CN được kết nối với Trung tâm dữ liệu quốc gia và các bộ, ngành, địa phương khác, bảo đảm liên thông theo Khung Kiến trúc dữ liệu quốc gia.

Ở chiều ngược lại, dữ liệu từ các hệ thống bên ngoài cũng có thể được tiếp nhận, tích hợp và làm giàu cho nền tảng dữ liệu của Bộ KH&CN, phục vụ các bài toán phân tích, dự báo và điều hành.

5. Kiến trúc an toàn dữ liệu



Hình 4. Kiến trúc an toàn dữ liệu

Quản lý An toàn bảo mật dữ liệu: Hệ thống dữ liệu của Bộ phải tuân thủ mô hình Khung an toàn dữ liệu 8 lớp sau đây:

- Lớp 1 - Dữ liệu gốc: Là tầng nền tảng của toàn bộ hệ thống thông tin. Mọi tập dữ liệu phải được gắn mã định danh duy nhất và siêu dữ liệu chuẩn hóa.

- Lớp 2 - Phân loại và gán nhãn: Bắt buộc phân loại dữ liệu theo quy định về Bảo vệ bí mật nhà nước (Không mật, Mật, Tối mật, Tuyệt mật), theo quy định về Bảo vệ dữ liệu cá nhân (Dữ liệu cá nhân cơ bản, Dữ liệu cá nhân nhạy cảm, Dữ liệu cá nhân khác) và gán nhãn tự động. Nhãn này là cơ sở để hệ thống tự động áp dụng các chính sách bảo mật tương ứng.

- Lớp 3 - Quản lý truy cập và mã hóa dữ liệu:

+ Phân quyền: Cấp quyền truy cập dựa trên vai trò, vị trí việc làm, bảo đảm người dùng chỉ được truy cập đúng phạm vi chức năng, nhiệm vụ và phạm vi dữ liệu được giao.

+ Xác thực đa yếu tố: Bắt buộc áp dụng đối với: (1) Tài khoản quản trị hệ thống (Admin); (2) Người dùng truy cập dữ liệu nhạy cảm cấp độ 4 trở lên (dữ

liệu dân cư, hồ sơ sức khỏe, tài chính, dữ liệu điều hành chiến lược) hoặc dữ liệu có ảnh hưởng trực tiếp đến an toàn xã hội và vận hành quốc gia.

+ Mã hóa: Thực hiện mã hóa dữ liệu tại 3 trạng thái: khi truyền tải (đường truyền), khi lưu trữ (trong ổ cứng) và khi sao lưu. Công nghệ mã hóa sử dụng các tiêu chuẩn và thuật toán hiện đại như TLS 1.3 cho mã hóa đường truyền và AES-256 cho mã hóa lưu trữ, bảo đảm an toàn dữ liệu trong suốt vòng đời khai thác và sử dụng.

+ Quản lý khóa: Tổ chức quản lý, lưu trữ và phân quyền sử dụng khóa mã hóa độc lập; Phân vùng quản lý khóa mã hóa riêng biệt cho từng đơn vị để tránh rủi ro tập trung.

- Lớp 4 - Che giấu dữ liệu, tính toán an toàn và truy vết:

+ Che giấu dữ liệu: Chỉ hiển thị thông tin tối thiểu cần thiết cho cán bộ xử lý (Ví dụ: Che bớt số điện thoại, số CCCD). Bắt buộc áp dụng với dữ liệu cá nhân cấp độ 3 trở lên.

+ Token hóa: Khi chia sẻ dữ liệu để huấn luyện AI hoặc phân tích Big Data, phải tách rời định danh cá nhân khỏi nội dung gốc.

+ Truy vết: Hệ thống phải ghi lại được lịch sử hình thành và thay đổi của dữ liệu.

- Lớp 5 - Giám sát và đánh giá theo mô hình “5 Yếu tố an toàn”:

+ Người dùng an toàn: Định danh, xác thực rõ ràng.

+ Mục đích an toàn: Chỉ khai thác khi có văn bản/thẩm quyền phê duyệt.

+ Dữ liệu phù hợp: Chỉ cung cấp dữ liệu đúng phạm vi cho phép.

+ Môi trường an toàn: Hạ tầng vật lý và mạng được kiểm soát, cách ly môi trường thử nghiệm (Test) và sản xuất (Production).

+ Đầu ra an toàn: Kết quả báo cáo không được tiết lộ thông tin cá nhân/bí mật.

- Lớp 6 - Kiểm soát pháp lý và quản lý vòng đời dữ liệu:

+ Ban hành quy định về vòng đời dữ liệu, bảo đảm dữ liệu được kiểm soát chặt chẽ từ khâu tạo lập, lưu trữ, khai thác/chia sẻ, đến lưu trữ lịch sử hoặc hủy bỏ.

- + Rà soát và xóa định kỳ các dữ liệu rác, dữ liệu hết hạn lưu trữ.
- + Việc chia sẻ dữ liệu giữa các đơn vị phải thông qua “Biên bản thỏa thuận chia sẻ dữ liệu” quy định rõ trách nhiệm pháp lý.
- Lớp 7 - Công cụ kỹ thuật bắt buộc: Nghiên cứu hoàn thiện các nền tảng, công cụ kỹ thuật, đảm bảo các yêu cầu như sau:
 - + Metadata and Lineage Tools: Quản lý siêu dữ liệu và truy vết nguồn gốc.
 - + SIEM (Security Information and Event Management): Giám sát sự kiện bảo mật, tích hợp với SOC của Bộ.
 - + DLP (Data Loss Prevention): Ngăn chặn copy dữ liệu ra USB, gửi email trái phép (bắt buộc với dữ liệu nhạy cảm/bí mật nhà nước).
 - + DAM (Database Activity Monitoring): Giám sát mọi câu lệnh truy vấn vào CSDL quan trọng.
 - + Compliance Dashboard: Bảng điều khiển hiển thị mức độ tuân thủ an toàn thông tin của từng đơn vị.
- Lớp 8 - Báo cáo và đánh giá an toàn dữ liệu định kỳ: Các đơn vị thực hiện chế độ báo cáo định kỳ về Trung tâm Công nghệ thông tin và báo cáo ngay lập tức khi có sự cố cấp độ III trở lên. Các chỉ số báo cáo bao gồm:
 - + Tỷ lệ dữ liệu được mã hóa;
 - + Số lượng truy cập trái phép bị ngăn chặn;
 - + Tốc độ phản ứng sự cố trung bình;
 - + Tính toàn vẹn của nhật ký truy cập (Logs).

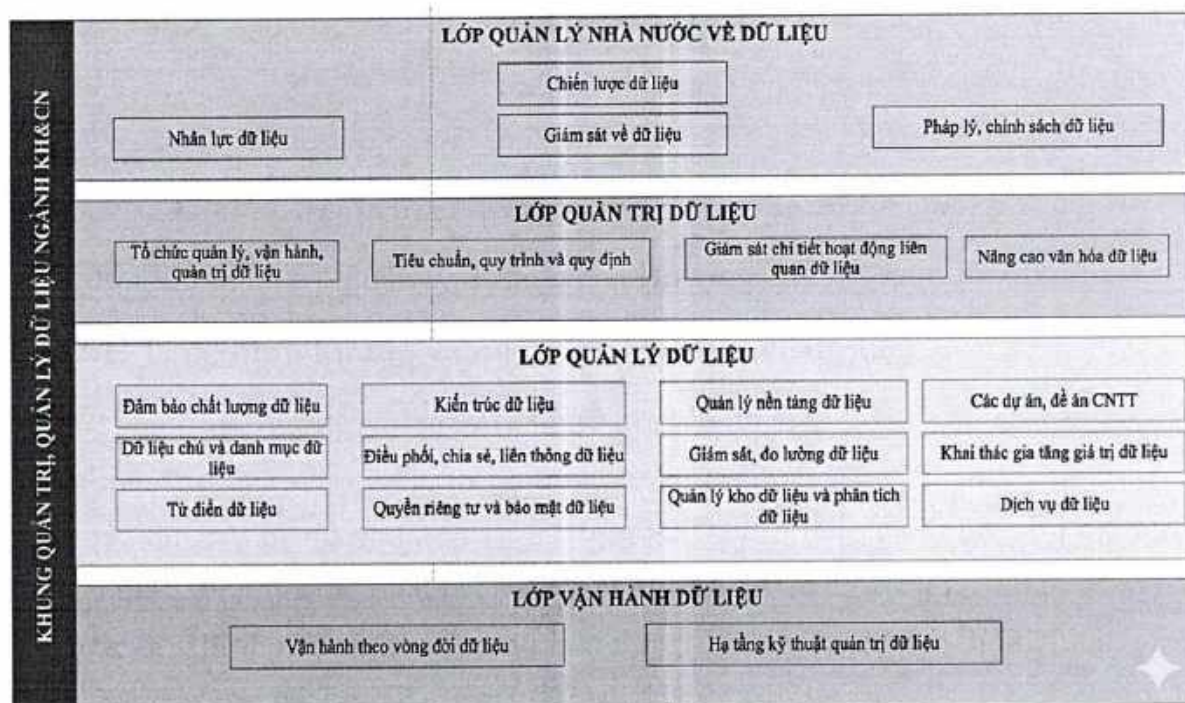
IV. KHUNG QUẢN TRỊ VÀ QUẢN LÝ DỮ LIỆU

1. Mục tiêu

- Thiết lập hệ thống dữ liệu thống nhất: xây dựng cơ chế quản trị và hạ tầng kỹ thuật đồng bộ, kết nối xuyên suốt từ Bộ đến các đơn vị thuộc Bộ; giải quyết triệt để tình trạng phân mảnh, thiếu đồng bộ giữa các đơn vị để đảm bảo dữ liệu luôn đúng, đủ, sạch, sống, thống nhất, dùng chung, sẵn sàng chia sẻ và liên thông.
- Chuyển đổi tư duy quản lý dữ liệu: từ quản lý dữ liệu đơn thuần sang quản trị và khai thác dữ liệu như một tài sản số quan trọng, phục vụ hiệu quả công tác

chỉ đạo, điều hành của Lãnh đạo Bộ; đồng thời cung cấp dữ liệu mở phục vụ người dân, doanh nghiệp, cộng đồng nghiên cứu và hệ sinh thái đổi mới sáng tạo.

2. Mô hình tổ chức quản trị, quản lý dữ liệu của Bộ



Hình 5. Mô hình tổ chức quản trị, quản lý dữ liệu của Bộ

Mô hình quản trị, quản lý dữ liệu của Bộ gồm các lớp sau:

2.1. Tầng định hướng và quản lý nhà nước về dữ liệu

Tầng định hướng và quản lý nhà nước về dữ liệu đóng vai trò xác lập nền tảng chiến lược và khung chính sách cho toàn bộ hoạt động quản trị, quản lý và khai thác dữ liệu trong phạm vi Bộ KH&CN. Đây là tầng cao nhất của mô hình, nơi xác định định hướng phát triển dữ liệu, thiết lập các nguyên tắc quản lý và ban hành các quy định nhằm đảm bảo dữ liệu được quản lý một cách thống nhất, minh bạch và hiệu quả.

a) Định hướng chiến lược dữ liệu Bộ KH&CN:

- Định hướng tổng thể của Bộ KH&CN trong việc phát triển và khai thác dữ liệu. Nội dung bao gồm việc xác định dữ liệu là một nguồn tài nguyên quan trọng phục vụ quản lý nhà nước, nghiên cứu khoa học và phát triển công nghệ.

- Bộ xác định các mục tiêu dài hạn như xây dựng hệ sinh thái dữ liệu khoa học và công nghệ, phát triển kho dữ liệu tập trung, thúc đẩy chia sẻ dữ liệu giữa các đơn vị trong Bộ và Bộ với các cơ quan khác. Đồng thời, chiến lược dữ liệu

cũng hướng tới việc tăng cường ứng dụng phân tích dữ liệu trong hoạt động hoạch định chính sách và ra quyết định.

b) Hành lang pháp lý và chính sách dữ liệu:

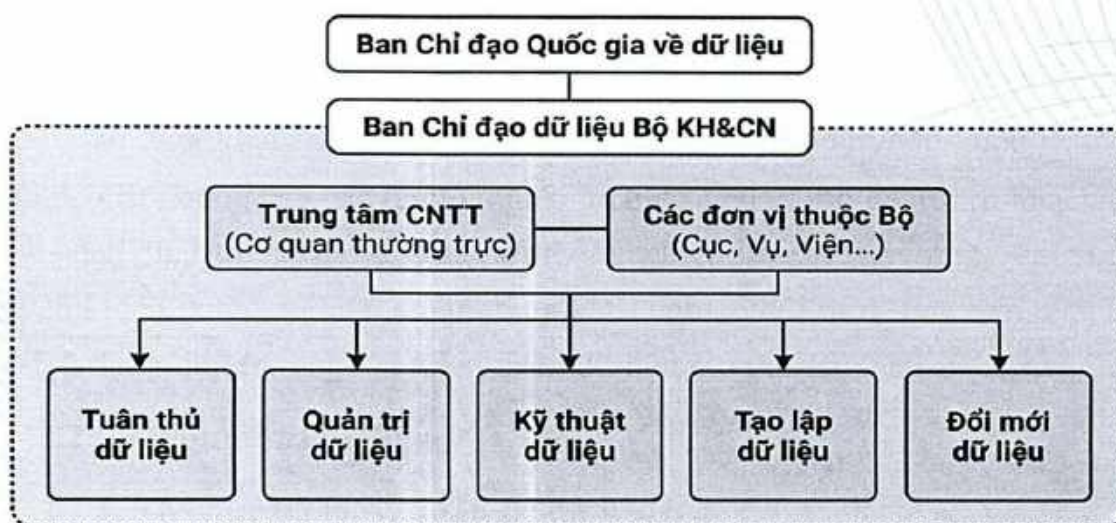
- Bao gồm hệ thống văn bản quy phạm pháp luật, quy chế, quy định và chính sách liên quan đến việc quản lý, khai thác và chia sẻ dữ liệu. Các quy định này xác định rõ trách nhiệm của các đơn vị trong việc thu thập, quản lý và sử dụng dữ liệu.

- Hành lang pháp lý cũng quy định các nguyên tắc về bảo mật, bảo vệ dữ liệu cá nhân, quyền truy cập dữ liệu, cũng như các cơ chế chia sẻ dữ liệu giữa các cơ quan nhà nước. Việc xây dựng hành lang pháp lý đầy đủ giúp đảm bảo dữ liệu được quản lý theo đúng quy định của pháp luật và hạn chế các rủi ro liên quan đến an toàn thông tin.

- Bảo đảm khai thác dữ liệu để phục vụ hoạt động chỉ đạo, điều hành của Lãnh đạo Bộ; cung cấp các dịch vụ phân tích dữ liệu vĩ mô để hỗ trợ hoạch định chính sách phát triển kinh tế - xã hội một cách chính xác; xây dựng chính sách khuyến khích doanh nghiệp, cộng đồng khởi nghiệp đổi mới sáng tạo, phát triển các ứng dụng dựa trên nguồn tài nguyên dữ liệu mở của Bộ.

2.2. Lớp quản trị dữ liệu

2.2.1. Tổ chức bộ máy, quản trị dữ liệu



Hình 6. Mô hình tổ chức bộ máy quản trị dữ liệu

Mô hình tổ chức bộ máy quản trị dữ liệu của Bộ bao gồm:

- Ban Chỉ đạo quốc gia về dữ liệu chỉ đạo chiến lược ở phạm vi quốc gia, định hướng chính sách và chiến lược phát triển dữ liệu quốc gia. Ban Chỉ đạo có nhiệm vụ điều phối hoạt động phát triển dữ liệu giữa các bộ, ngành và địa phương nhằm bảo đảm việc xây dựng và khai thác dữ liệu được thực hiện thống nhất trên phạm vi toàn quốc.

- Ban Chỉ đạo dữ liệu Bộ KH&CN chỉ đạo trong phạm vi Bộ, có nhiệm vụ tổ chức triển khai các chiến lược và chương trình phát triển dữ liệu quốc gia trong lĩnh vực khoa học và công nghệ. Ban Chỉ đạo chịu trách nhiệm điều phối các hoạt động quản trị dữ liệu trong các đơn vị thuộc Bộ, đồng thời giám sát việc thực hiện các quy định và tiêu chuẩn liên quan đến dữ liệu.

- Trung tâm Công nghệ thông tin (Cơ quan thường trực Ban Chỉ đạo) chủ trì triển khai, điều phối, giám sát việc tuân thủ Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu của Bộ; có nhiệm vụ tham mưu và hỗ trợ Ban Chỉ đạo trong việc tổ chức triển khai các chương trình và kế hoạch phát triển dữ liệu, đồng thời điều phối các hoạt động liên quan đến quản trị dữ liệu trong toàn Bộ.

- Các đơn vị thuộc Bộ là các đơn vị chuyên môn trực tiếp tạo lập và quản lý các nguồn dữ liệu chuyên ngành. Các đơn vị này chịu trách nhiệm thu thập, cập nhật và khai thác dữ liệu phục vụ cho công tác quản lý nhà nước, nghiên cứu khoa học và hoạt động đổi mới sáng tạo trong lĩnh vực khoa học và công nghệ.

- Các bộ phận chức năng: Được tổ chức thành 05 bộ phận chuyên trách, bao gồm:

(1) Bộ phận Tuân thủ dữ liệu:

- Nhiệm vụ: Bảo đảm các hoạt động quản lý và khai thác dữ liệu trong Bộ được thực hiện theo đúng quy định của pháp luật, các tiêu chuẩn dữ liệu và các quy chế quản lý nội bộ.

- Đơn vị chủ trì: Trung tâm Công nghệ thông tin

- Đơn vị phối hợp: Các đơn vị thuộc Bộ.

(2) Bộ phận Quản trị dữ liệu:

- Nhiệm vụ: Xây dựng các chính sách, quy trình và tiêu chuẩn quản lý dữ liệu nhằm bảo đảm dữ liệu được quản lý thống nhất, có chất lượng và có khả năng chia sẻ giữa các hệ thống thông tin.

- Đơn vị chủ trì: Trung tâm Công nghệ thông tin

- Đơn vị phối hợp: Các đơn vị thuộc Bộ.

(3) Bộ phận Kỹ thuật dữ liệu:

- Nhiệm vụ: Triển khai các giải pháp công nghệ phục vụ quản lý và khai thác dữ liệu, bao gồm các hệ thống lưu trữ dữ liệu, các nền tảng tích hợp dữ liệu và các công cụ phân tích dữ liệu.

- Đơn vị chủ trì: Trung tâm Công nghệ thông tin

- Đơn vị phối hợp: Các đơn vị thuộc Bộ.

(4) Bộ phận Tạo lập dữ liệu:

- Nhiệm vụ: Tạo lập dữ liệu bao gồm việc thu thập, tổng hợp và cập nhật dữ liệu từ các nguồn khác nhau trong Bộ nhằm bảo đảm dữ liệu được duy trì đầy đủ và chính xác.

- Đơn vị chủ trì: Các đơn vị thuộc Bộ

- Đơn vị phối hợp: Trung tâm Công nghệ thông tin.

(5) Bộ phận Đổi mới dữ liệu:

- Nhiệm vụ: Tập trung vào việc nghiên cứu và ứng dụng các công nghệ và phương pháp khai thác dữ liệu mới nhằm tạo ra giá trị gia tăng từ dữ liệu và hỗ trợ quá trình ra quyết định trong quản lý nhà nước.

- Đơn vị chủ trì: Trung tâm Công nghệ thông tin

- Đơn vị phối hợp: Các đơn vị thuộc Bộ

2.2.2. Xây dựng các quy định, quy trình và tiêu chuẩn quản trị dữ liệu

Xây dựng và hoàn thiện hành lang pháp lý - kỹ thuật đồng bộ, làm cơ sở để chuẩn hóa hoạt động quản lý, kết nối và chia sẻ dữ liệu số của Bộ:

- Xây dựng quy định về quản trị dữ liệu

+ Quy định về quản lý, tạo lập, cập nhật, khai thác và sử dụng dữ liệu.

+ Quy định về phân loại dữ liệu, xác định mức độ mật, phạm vi chia sẻ và công bố dữ liệu.

+ Quy định về trách nhiệm của các đơn vị, tổ chức, cá nhân trong việc quản lý và sử dụng dữ liệu.

+ Quy định về chia sẻ, kết nối và khai thác dữ liệu giữa các đơn vị trong Bộ và với các bộ, ngành, địa phương.

+ Quy định về bảo đảm an toàn thông tin, bảo vệ dữ liệu trong suốt vòng đời dữ liệu.

- Xây dựng quy trình quản trị dữ liệu theo vòng đời

+ Quy trình thu thập, tạo lập dữ liệu từ các hệ thống thông tin, cơ sở dữ liệu chuyên ngành.

+ Quy trình kiểm tra, làm sạch và chuẩn hóa dữ liệu theo từ điển dữ liệu ngành.

+ Quy trình tích hợp, đồng bộ và lưu trữ dữ liệu vào các kho dữ liệu tập trung.

+ Quy trình chia sẻ, cung cấp và khai thác dữ liệu thông qua các nền tảng tích hợp, chia sẻ dữ liệu.

+ Quy trình cập nhật, chỉnh sửa và kiểm soát thay đổi dữ liệu.

+ Quy trình lưu trữ, bảo quản và kết thúc sử dụng dữ liệu theo quy định.

2.2.3. Giám sát hoạt động liên quan đến dữ liệu

Việc giám sát các hoạt động liên quan đến dữ liệu nhằm bảo đảm việc quản lý, khai thác, sử dụng và chia sẻ dữ liệu trong Bộ KH&CN được thực hiện đúng quy định, bảo đảm chất lượng, an toàn và hiệu quả, phục vụ công tác quản lý nhà nước và phát triển khoa học và công nghệ.

a) Nguyên tắc giám sát

- Liên tục và thời gian thực: Hoạt động giám sát phải được thực hiện liên tục.

- Minh bạch và truy vết: Mọi hành vi tác động vào dữ liệu (xem, thêm, sửa, xóa, chia sẻ) bắt buộc phải được theo dõi và ghi vào nhật ký hệ thống.

- Cảnh báo chủ động: Hệ thống phải có khả năng tự phát hiện rủi ro và gửi cảnh báo trước khi sự cố xảy ra.

- Tuân thủ pháp lý: Đảm bảo sự phù hợp với Khung kiến trúc dữ liệu Bộ và các quy định về bảo vệ dữ liệu cá nhân.

b) Nội dung giám sát:

- Giám sát việc thu thập, tạo lập và cập nhật dữ liệu tại các hệ thống thông tin, cơ sở dữ liệu chuyên ngành.

- Giám sát việc tuân thủ các quy định, quy trình và tiêu chuẩn về quản trị dữ liệu.

- Giám sát chất lượng dữ liệu, bao gồm các tiêu chí về tính đầy đủ, chính xác, nhất quán và kịp thời.

- Giám sát hoạt động tích hợp, chia sẻ và kết nối dữ liệu giữa các hệ thống trong và ngoài Bộ.

- Giám sát việc khai thác, sử dụng dữ liệu, bảo đảm đúng mục đích, đúng thẩm quyền và phù hợp với quy định pháp luật.

- Giám sát việc công bố, cung cấp dữ liệu mở và dịch vụ dữ liệu.

- Giám sát việc bảo đảm an toàn thông tin, bảo mật dữ liệu trong quá trình xử lý và sử dụng.

c) Cơ chế giám sát

- Thiết lập hệ thống giám sát dữ liệu tập trung của Bộ, bảo đảm khả năng thu thập log truy cập, xử lý và phát hiện các hành vi bất thường.

- Áp dụng công nghệ AI/Big Data để phân tích hành vi người dùng.

d) Trách nhiệm giám sát

- Trung tâm Công nghệ thông tin là đầu mối tổ chức, giám sát, kiểm tra định kỳ và đột xuất các hoạt động dữ liệu của Bộ.

- Các đơn vị thuộc Bộ có trách nhiệm thực hiện giám sát nội bộ đối với dữ liệu thuộc phạm vi quản lý, bảo đảm tính toàn vẹn và sẵn sàng kiểm tra.

2.2.4. Xây dựng văn hóa dữ liệu

Việc xây dựng văn hóa dữ liệu trong Bộ KH&CN được thực hiện theo các nguyên tắc sau:

- Dữ liệu được tạo lập trong quá trình hoạt động của Bộ là tài sản chung, được quản lý tập trung, chia sẻ và khai thác theo quy định; hạn chế tình trạng cục bộ, phân tán dữ liệu tại từng đơn vị.

- Mọi hoạt động chỉ đạo, điều hành, xây dựng chính sách, kế hoạch, chương trình khoa học và công nghệ phải dựa trên dữ liệu và các bằng chứng số, bảo đảm tính khách quan và khoa học.

- Dữ liệu được chia sẻ giữa các đơn vị trong Bộ theo nguyên tắc chủ động, kịp thời, đúng thẩm quyền và tuân thủ quy định của pháp luật.

- Dữ liệu được công khai, minh bạch theo quy định; các đơn vị, cá nhân chịu trách nhiệm về tính chính xác, đầy đủ và kịp thời của dữ liệu do mình quản lý.

- Việc tạo lập, quản lý, chia sẻ và khai thác dữ liệu phải tuân thủ quy định của pháp luật về bảo vệ dữ liệu, an toàn thông tin và bí mật nhà nước.

- Triển khai đồng bộ khung quản trị, quản lý dữ liệu của Bộ; xác lập rõ vai trò, trách nhiệm của các đơn vị trong tạo lập, cập nhật, chia sẻ và khai thác dữ liệu; tổ chức đào tạo, bồi dưỡng kỹ năng sử dụng, phân tích dữ liệu phục vụ công tác pháp lý, quản lý nhà nước; khuyến khích sáng kiến đổi mới, ứng dụng công nghệ số và AI dựa trên dữ liệu trong các lĩnh vực nghiệp vụ của Bộ.

2.3. Lớp quản lý dữ liệu

2.3.1. Quản lý chất lượng dữ liệu

- Dữ liệu trong Bộ phải có độ tin cậy cao, phục vụ chính xác cho công tác chỉ đạo điều hành và giải quyết thủ tục hành chính. Dữ liệu phải đảm bảo “Đúng - Đủ - Sạch - Sống - Thống nhất - Dùng chung”.

- Trong Bộ KH&CN được thực hiện thống nhất theo quy định của pháp luật và Khung quản trị, quản lý dữ liệu của Bộ; coi dữ liệu là tài sản dùng chung của ngành, phục vụ xây dựng, tổ chức thi hành pháp luật và quản lý nhà nước.

2.3.2. Quản lý dữ liệu chủ, dữ liệu danh mục dùng chung

Bộ KH&CN chủ trì xác định, công bố và quản lý danh mục dữ liệu chủ và danh mục dữ liệu dùng chung trong toàn ngành (như danh mục tiêu chuẩn đo lường chất lượng, đổi mới sáng tạo, KH&CN...), đảm bảo thống nhất với danh mục dữ liệu chủ và danh mục dùng chung quốc gia. Phân công rõ vai trò chủ sở hữu đối với từng loại dữ liệu; xác định quy trình tạo, cập nhật, lưu trữ, quản lý

phiên bản và thực hiện truy vết thay đổi để đảm bảo tính nhất quán; định kỳ tổ chức rà soát, cập nhật danh sách, công bố thông tin điều chỉnh.

2.3.3. Từ điển dữ liệu

Xây dựng Từ điển dữ liệu dùng chung của Bộ KH&CN làm công cụ chuẩn hóa, chứa định nghĩa, kiểu dữ liệu, mã tham chiếu và metadata; bảo đảm áp dụng nhất quán trong các hệ thống, tuân thủ theo cấu trúc của từ điển dữ liệu quốc gia; xác định quy trình đăng ký, phê duyệt, công bố và quản lý thay đổi từ điển dữ liệu dùng chung của Bộ KH&CN.

2.3.4. Kiến trúc dữ liệu

- Kiến trúc dữ liệu của Bộ phải mô tả đầy đủ các nguồn dữ liệu, quy trình nghiệp vụ KH&CN có sử dụng dữ liệu, các hệ thống thông tin/cơ sở dữ liệu phục vụ lưu trữ, chia sẻ dữ liệu và luồng xử lý, cung cấp dữ liệu trong toàn ngành; bảo đảm phù hợp Khung Kiến trúc số quốc gia, Kiến trúc số Bộ KH&CN và Khung kiến trúc dữ liệu quốc gia, từ điển dùng chung.

- Kiến trúc dữ liệu được thẩm định, phê duyệt theo thẩm quyền; được rà soát, cập nhật khi có thay đổi về cấu trúc, nguồn, hệ thống hoặc phương thức tích hợp, chia sẻ dữ liệu; mọi thay đổi phải được quản lý phiên bản, bảo đảm truy vết, phục hồi và phục vụ kiểm tra theo quy định.

2.3.5. Quản lý quyền riêng tư và bảo mật dữ liệu

Thực hiện kiểm soát dữ liệu trong toàn bộ vòng đời; bảo vệ dữ liệu cá nhân chuyên gia, nhà khoa học và dữ liệu nhạy cảm trong hệ thống, bao gồm cả dữ liệu tiếp nhận từ đơn vị khác và dữ liệu đã chia sẻ. Áp dụng đồng bộ các lớp kiểm soát an toàn dữ liệu, gồm: định danh, phân loại và gán nhãn dữ liệu; quản lý truy cập, xác thực và mã hóa; che giấu dữ liệu hoặc mã hóa khi khai thác; truy vết dữ liệu (data lineage); giám sát an toàn; triển khai các công cụ tự động và thực hiện chế độ giám sát, báo cáo định kỳ theo quy định.

2.3.6. Quản lý nền tảng dữ liệu

Nền tảng dữ liệu của Bộ phải bảo đảm tính sẵn sàng, an toàn, khả năng mở rộng, hiệu năng và chất lượng dữ liệu phục vụ hệ thống nghiệp vụ, công tác chỉ đạo, điều hành, phân tích và chia sẻ dữ liệu; dữ liệu được tích hợp từ các hệ thống chuyên ngành và nguồn bên ngoài qua nền tảng tích hợp, chia sẻ dữ liệu và API an toàn; đồng thời áp dụng cơ chế tự động kiểm soát chất lượng, phát hiện – cảnh

báo sai lệch, truy vết nguyên nhân và báo cáo phục vụ công tác quản lý, giám sát của Bộ.

2.3.7. Quản lý kho dữ liệu và phân tích

Các kho dữ liệu của Bộ là nền tảng chuẩn hóa và lưu trữ dữ liệu tập trung từ các hệ thống, cơ sở dữ liệu chuyên ngành (như tiêu chuẩn đo lường chất lượng, đổi mới sáng tạo, KH&CN...), phục vụ tổng hợp, báo cáo, phân tích và hỗ trợ lãnh đạo Bộ và các đơn vị trực thuộc trong công tác chỉ đạo, điều hành. Các chỉ tiêu, cấu trúc báo cáo và logic nghiệp vụ được xây dựng thống nhất, gắn với mô hình dữ liệu kho, bảo đảm tính nhất quán, toàn vẹn và khả năng truy vết nguồn dữ liệu.

2.3.8. Điều phối, chia sẻ liên thông dữ liệu

Bộ Khoa học và Công nghệ tổ chức điều phối, kết nối, chia sẻ dữ liệu trong phạm vi ngành và liên thông với các bộ, ngành, địa phương theo quy định; bảo đảm kết nối kỹ thuật thống nhất, kiểm soát an ninh, an toàn thông tin, giám sát sử dụng và truy vết dữ liệu. Việc truy cập, khai thác, chia sẻ dữ liệu phải đúng mục đích, đúng thẩm quyền, tuân thủ quy định về bảo vệ dữ liệu cá nhân, bí mật nhà nước và được ghi nhận đầy đủ nhật ký.

2.3.9. Giám sát đo lường dữ liệu

Xây dựng bộ chỉ số đo lường quản trị và chất lượng dữ liệu (KPI, KQI) phục vụ đánh giá mức độ trưởng thành trong quản trị dữ liệu của Bộ và các đơn vị trực thuộc, làm cơ sở cải tiến năng lực quản trị dữ liệu. Áp dụng phương pháp đánh giá có thang điểm, báo cáo định kỳ và công bố kết quả xếp hạng năng lực quản trị dữ liệu của các cơ quan, đơn vị.

2.3.10. Các dự án, đề án CNTT

Quản lý các dự án, đề án công nghệ thông tin của Bộ KH&CN là quá trình lập kế hoạch, tổ chức thực hiện, giám sát và đánh giá các chương trình, đề án, dự án liên quan đến dữ liệu (thu thập, tích hợp, chuẩn hóa, xây dựng cơ sở dữ liệu chuyên ngành về tiêu chuẩn đo lường chất lượng, sở hữu trí tuệ, KH&CN phân tích, chia sẻ liên thông và dữ liệu mở). Đồng thời thiết lập cơ chế quản lý thay đổi, nhận diện và kiểm soát rủi ro về dữ liệu, bảo đảm tuân thủ pháp luật, an toàn thông tin và hiệu quả triển khai, tránh lãng phí, bảo đảm tính thống nhất, đồng bộ trong triển khai thực hiện các hệ thống thông tin phục vụ quản lý nhà nước.

2.3.11. Khai thác giá trị gia tăng dựa trên dữ liệu

- Các dữ liệu của Bộ (bao gồm cả dữ liệu mở) được công bố theo danh mục được phê duyệt, bảo đảm tuân thủ pháp luật về bảo mật, bảo vệ dữ liệu cá nhân và bí mật nhà nước, đặc biệt là các thông tin nhạy cảm về công nghệ và năng lượng nguyên tử. Áp dụng định dạng mở kèm siêu dữ liệu đầy đủ, cập nhật định kỳ và cung cấp kênh truy cập phù hợp như Cổng dữ liệu quốc gia, Cổng dữ liệu mở của Bộ.

- Việc khai thác dữ liệu có thể thông qua chia sẻ có điều kiện, hợp tác công – tư hoặc các dịch vụ phân tích xu hướng công nghệ theo yêu cầu, khai thác dữ liệu theo quy định của pháp luật. Việc cung cấp dịch vụ dữ liệu có thu phí (nếu có) thực hiện theo cơ chế tài chính, phí – lệ phí do cơ quan có thẩm quyền quy định, phải được đánh giá và quản lý chặt chẽ, bảo đảm cạnh tranh lành mạnh, an ninh quốc gia và đúng quy định pháp luật.

2.3.12. Khai thác dịch vụ dữ liệu

Xây dựng các công cụ và nền tảng kỹ thuật cho phép cung cấp, khai thác, sử dụng dữ liệu khoa học, công nghệ và đổi mới sáng tạo một cách hiệu quả. Bảo đảm khả năng truy cập, tiêu thụ, phân tích và khai thác các nguồn dữ liệu chuyên ngành theo nhiều hình thức khác nhau, từ tra cứu văn bản bảo hộ đến khai thác kết quả nghiên cứu sử dụng ngân sách nhà nước, nhằm phục vụ yêu cầu quản trị, vận hành và ra quyết định dựa trên dữ liệu.

2.4. Lốp vận hành dữ liệu

Thiết lập cơ chế vận hành kỹ thuật và quản lý vòng đời dữ liệu chuyên nghiệp, bảo đảm hạ tầng số của Bộ luôn sẵn sàng, an toàn và đáp ứng nhu cầu khai thác dữ liệu phục vụ nghiên cứu, đổi mới sáng tạo và quản lý nhà nước.

2.4.1. Vận hành vòng đời dữ liệu

- Thực hiện quản lý tập trung và xuyên suốt toàn bộ vòng đời của dữ liệu khoa học, công nghệ và đổi mới sáng tạo, từ khâu thu nhận kết quả nghiên cứu, xác lập quyền sở hữu trí tuệ đến lưu trữ lịch sử và hủy bỏ dữ liệu theo quy định. Quá trình vận hành phải tuân thủ nghiêm ngặt các quy định của pháp luật về dữ liệu, các chính sách bảo mật thông tin và bảo vệ bí mật nhà nước trong lĩnh vực khoa học công nghệ. Thiết lập cơ chế ghi nhật ký hệ thống (log) đầy đủ, phân

quyền truy cập chặt chẽ, xác định rõ trách nhiệm giải trình của đội ngũ vận hành và xây dựng quy trình tiêu hủy dữ liệu an toàn, không để rò rỉ thông tin nhạy cảm.

- Nguyên tắc quản lý:

- + Vòng đời dữ liệu không được quản lý cắt khúc mà phải được kiểm soát liên mạch trên một nền tảng thống nhất.

- + Toàn bộ các khâu trong vòng đời dữ liệu bắt buộc phải tuân thủ nghiêm ngặt các quy định về an toàn an ninh mạng và bảo vệ dữ liệu cá nhân. Tối ưu hóa giá trị: dữ liệu không chỉ được cất giữ mà phải được làm giàu. Trong vòng đời hoạt động, dữ liệu phải liên tục được chuẩn hóa, làm sạch, phân tích và chia sẻ để phục vụ chỉ đạo điều hành và phát triển dịch vụ số, tránh tình trạng dữ liệu không được sử dụng.

2.4.2. Hạ tầng kỹ thuật quản trị dữ liệu

- Bảo đảm hệ thống hạ tầng kỹ thuật (bao gồm trung tâm dữ liệu, nền tảng điện toán đám mây và mạng chuyên dùng) đáp ứng yêu cầu về tính sẵn sàng cao, khả năng mở rộng linh hoạt và an toàn thông tin tối đa cho các cơ sở dữ liệu quốc gia và chuyên Bộ KH&CN.

- Xây dựng và áp dụng thống nhất các quy trình cấp phát tài nguyên số, quy chuẩn về sao lưu và phục hồi dữ liệu sau sự cố nhằm bảo đảm tính liên tục của các dịch vụ công trực tuyến và hệ thống nghiệp vụ. Triển khai các công cụ vận hành tự động và hệ thống giám sát tập trung hoạt động 24/7 đối với các hệ thống thông tin quan trọng của Bộ để kịp thời phát hiện, xử lý các nguy cơ mất an toàn dữ liệu

- Đầu tư và áp dụng các công nghệ tiên tiến để tự động hóa công tác quản trị, tự động quét và thu thập thông tin về cấu trúc dữ liệu từ các hệ thống nguồn, giúp duy trì bản đồ dữ liệu luôn cập nhật mà không cần con người nhập liệu thủ công.

3. Từ điển dữ liệu dùng chung

3.1. Mục đích

- Từ điển dữ liệu dùng chung là tập hợp các định nghĩa, mô tả và quy định chuẩn hóa đối với các thành phần dữ liệu được sử dụng trong các hệ thống thông tin của Bộ. Từ điển dữ liệu dùng chung đóng vai trò làm cơ sở thống nhất về khái niệm, cấu trúc và ý nghĩa của dữ liệu, nhằm bảo đảm các dữ liệu được tạo lập,

quản lý và khai thác trong các hệ thống thông tin của Bộ có sự đồng bộ và nhất quán.

- Việc xây dựng từ điển dữ liệu dùng chung giúp các đơn vị thuộc Bộ hiểu rõ nội dung và cách thức sử dụng dữ liệu, đồng thời tạo điều kiện thuận lợi cho việc tích hợp, chia sẻ và khai thác dữ liệu giữa các hệ thống thông tin trong nội bộ Bộ cũng như giữa Bộ với các cơ quan, tổ chức khác.

- Là nền tảng quan trọng cho công tác quản trị, quản lý dữ liệu Bộ KH&CN, làm cơ sở để quản lý siêu dữ liệu (Metadata), quản lý chất lượng dữ liệu, xác định dữ liệu chủ và phân loại mức độ chia sẻ dữ liệu theo quy định

- Góp phần bảo đảm tuân thủ pháp luật về bảo vệ dữ liệu cá nhân chuyên gia, nhà khoa học, an toàn thông tin và bí mật nhà nước về công nghệ thông qua việc gắn từng trường dữ liệu với căn cứ pháp lý và chuẩn kỹ thuật.

- Tạo điều kiện thuận lợi cho việc tích hợp dữ liệu với Nền tảng chia sẻ, điều phối dữ liệu quốc gia, các cơ sở dữ liệu quốc gia và hệ thống của các bộ, ngành, địa phương, bảo đảm tuân thủ Khung kiến trúc dữ liệu quốc gia.

3.2. Nguyên tắc

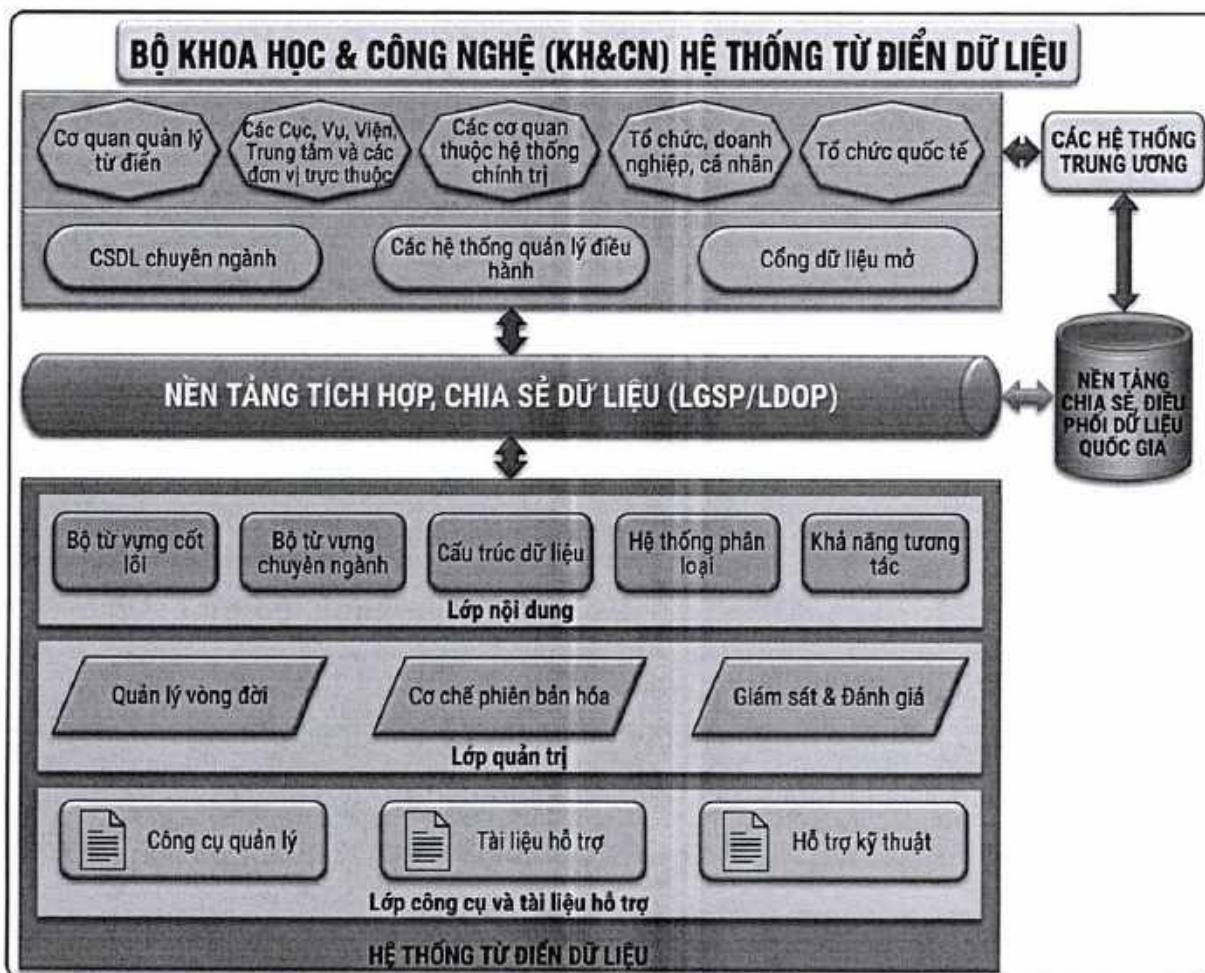
- Nguyên tắc kế thừa: Từ điển dữ liệu phải bảo đảm kế thừa và tuân thủ các tiêu chuẩn, quy chuẩn và danh mục dữ liệu đã được ban hành ở cấp quốc gia. Điều này nhằm bảo đảm sự thống nhất giữa hệ thống dữ liệu của Bộ với hệ thống dữ liệu quốc gia, đồng thời tạo điều kiện thuận lợi cho việc kết nối và chia sẻ dữ liệu với các bộ, ngành và địa phương.

- Nguyên tắc mở rộng: Từ điển dữ liệu cần được thiết kế theo hướng linh hoạt, cho phép bổ sung các thành phần dữ liệu mới mà không làm ảnh hưởng đến cấu trúc dữ liệu đã được xây dựng trước đó. Các thành phần dữ liệu mới khi được bổ sung cần tuân thủ các tiêu chuẩn chung về đặt tên, định nghĩa và cấu trúc dữ liệu để bảo đảm sự tương thích với các dữ liệu hiện có. Trong trường hợp phát sinh các thành phần dữ liệu mới, các đơn vị cần thực hiện rà soát, đối chiếu với các dữ liệu đã có để bảo đảm không tạo ra các định nghĩa trùng lặp hoặc mâu thuẫn.

- Nguyên tắc nhất quán: Xây dựng từ điển dữ liệu, các trường dữ liệu phải được định nghĩa rõ ràng về tên gọi, ý nghĩa, kiểu dữ liệu, định dạng và phạm vi giá trị. Các thành phần dữ liệu giống nhau chỉ được định nghĩa một lần trong từ

điển dữ liệu dùng chung và được sử dụng thống nhất trong toàn bộ các hệ thống thông tin của Bộ. Điều này giúp hạn chế tình trạng trùng lặp dữ liệu hoặc sử dụng các khái niệm dữ liệu khác nhau cho cùng một nội dung thông tin.

3.3. Hệ thống từ điển dữ liệu



Hình 7. Sơ đồ tổng quát Hệ thống Từ điển dữ liệu của Bộ

Hệ thống Từ điển dữ liệu của Bộ bao gồm các lớp sau:

3.3.1. Lớp nội dung

- Bộ từ vựng cốt lõi: tập hợp thuật ngữ cơ bản dùng chung và xuyên suốt trong nhiều ngành, hệ thống và cấp quản lý khác nhau (dữ liệu chủ, dữ liệu danh mục dùng chung và các dữ liệu dùng chung liên ngành khác), bảo đảm tính phổ cập (tính chung), khả năng tái sử dụng.

- Bộ từ vựng theo chuyên ngành: bao gồm các thuật ngữ đặc thù phục vụ quản lý nhà nước trong các lĩnh vực trọng tâm của Bộ như: Tiêu chuẩn đo lường chất lượng, sở hữu trí tuệ, KH&CN, Năng lượng nguyên tử...

- Cấu trúc dữ liệu: tổ chức theo thực thể (entity) và mối quan hệ (relationship) giữa các thực thể, hỗ trợ mô tả rõ ràng kèm định dạng kỹ thuật để dễ liên kết.

- Hệ thống phân loại từ vựng (taxonomy): xây dựng danh mục phân loại theo ngành, chủ đề, loại thực thể, mức độ nhạy cảm... để hỗ trợ tìm kiếm, khám phá dữ liệu (data discovery) và quản lý dữ liệu.

- Khả năng tương tác quốc tế: bảo đảm trao đổi và liên thông dữ liệu xuyên biên giới thông qua cơ chế ánh xạ (data mapping) tới các chuẩn quốc tế phù hợp.

3.3.2. Lớp quản trị

Quy trình quản lý vòng đời từ vựng theo các bước sau:

- Khởi tạo/Đề xuất: các đơn vị thuộc Bộ có thể đề xuất từ vựng mới hoặc thay đổi thuật ngữ đã có.

- Xây dựng/Rà soát sơ bộ: Trung tâm Công nghệ thông tin phối hợp với các đơn vị có liên quan rà soát hồ sơ đề xuất, rà soát trùng lặp, ảnh hưởng.

- Tham vấn/Lấy ý kiến: đối với từ vựng có phạm vi sử dụng liên ngành hoặc tác động lớn, thực hiện tham vấn với các bên liên quan để đánh giá ý nghĩa, tính liên ngành.

- Phê duyệt/Đăng ký: từ vựng được phê duyệt chính thức và được đăng ký công khai trên Hệ thống Từ điển dữ liệu dùng chung.

Cơ chế phiên bản hóa:

- Quản lý các thay đổi, bảo đảm tính nhất quán và tránh gây gián đoạn cho các hệ thống hiện hành khi cập nhật:

- Phiên bản Chính (Major Release): được phát hành theo chu kỳ 3 năm, bao gồm cập nhật toàn diện cả Bộ Từ vựng cốt lõi và kiến trúc kỹ thuật của Hệ thống từ điển;

- Phiên bản Phụ (Minor Release): có thể được phát hành 12 tháng (không cùng năm với phiên bản chính), các phiên bản này chỉ cập nhật nội dung liên quan đến các lĩnh vực chuyên ngành (Domains) và các bảng mã;

- Quản lý thay đổi: các thay đổi được ghi lại chi tiết trong nhật ký phiên bản, các phiên bản cũ được duy trì trong một khoảng thời gian xác định để hỗ trợ các hệ thống chưa kịp cập nhật.

Giám sát và đánh giá:

- Giám sát liên tục: Trung tâm Công nghệ thông tin giám sát việc sử dụng Từ điển dữ liệu dùng chung thông qua phản hồi từ các đơn vị thuộc Bộ. Các vấn đề (như thuật ngữ không phù hợp, lỗi dữ liệu) được ghi nhận và xử lý theo quy trình phê duyệt.

- Đánh giá định kỳ: Trung tâm Công nghệ thông tin đánh giá hàng năm hoặc khi có nhu cầu để xem xét tính phù hợp của các thuật ngữ và định nghĩa hiện đánh giá nhu cầu thêm thuật ngữ mới hoặc sửa đổi thuật ngữ hiện có.

3.3.3. Lớp công cụ và hỗ trợ

Cung cấp các phương tiện kỹ thuật để các đơn vị khai thác từ điển dữ liệu:

- Công cụ quản lý: Các phần mềm, nền tảng số hỗ trợ tra cứu trực tuyến, biên tập siêu dữ liệu và quản lý các bảng mã dùng chung.

- Tài liệu hỗ trợ: Các văn bản hướng dẫn kỹ thuật, quy định về chuẩn hóa đặt tên trường dữ liệu và cấu trúc gói tin trao đổi dữ liệu.

- Hỗ trợ kỹ thuật: Duy trì đội ngũ và kênh tiếp nhận yêu cầu để xử lý các vấn đề phát sinh trong quá trình tích hợp dữ liệu giữa các đơn vị thông qua nền tảng tích hợp, chia sẻ dữ liệu của Bộ (LGSP/LDOP).

3.4. Cấu trúc và tiêu chuẩn kỹ thuật

Mỗi thuật ngữ trong Từ điển dữ liệu dùng chung được mô tả theo hai nhóm chính: định nghĩa nghiệp vụ (để hiểu và dùng đúng) và định nghĩa kỹ thuật (để hệ thống tương tác). Các thành phần chính bao gồm:

3.4.1. Định nghĩa nghiệp vụ (Business Metadata)

Nhóm này phục vụ mục đích quản lý hành chính, giúp cán bộ và các tổ chức, cá nhân tham gia vào hệ sinh thái KH&CN hiểu và sử dụng đúng ý nghĩa bản chất của dữ liệu.

- *Mã định danh duy nhất (ID)*: Là mã số nhận diện duy nhất, không thay đổi trong suốt vòng đời của đối tượng dữ liệu để phục vụ việc liên thông, tích hợp.

- *Tên nghiệp vụ*: Tên gọi chính thức bằng tiếng Việt, phản ánh đúng khái niệm chuyên ngành được quy định trong các văn bản quy phạm pháp luật của Bộ KH&CN.

- *Định nghĩa nghiệp vụ*: Mô tả chi tiết, đầy đủ về nội dung, ý nghĩa và phạm vi của thuật ngữ, đảm bảo không gây nhầm lẫn giữa các lĩnh vực (ví dụ: phân biệt rõ "Sáng chế" và "Giải pháp hữu ích").

- *Ngữ cảnh sử dụng*: Xác định cụ thể quy trình nghiệp vụ hoặc hệ thống thông tin chuyên ngành sử dụng dữ liệu này (ví dụ: dùng trong đăng ký bảo hộ quyền sở hữu công nghiệp hoặc nghiệm thu nhiệm vụ KH&CN).

- *Nhóm dữ liệu/Phân loại*: Thuộc lĩnh vực quản lý nhà nước cụ thể của Bộ (như: Tiêu chuẩn Đo lường Chất lượng, Sở hữu trí tuệ, Năng lượng nguyên tử...).

- *Mối quan hệ*: Mô tả các mối liên kết nghiệp vụ với các đối tượng dữ liệu khác (ví dụ: "Số văn bằng bảo hộ" có quan hệ thuộc tính với đối tượng "Chủ bằng").

- *Đơn vị tạo lập và đơn vị sở hữu*: Xác định rõ cơ quan chịu trách nhiệm khởi tạo dữ liệu gốc và cơ quan có thẩm quyền pháp lý quản lý dữ liệu đó (ví dụ: Cục Sở hữu trí tuệ là đơn vị sở hữu dữ liệu về Sáng chế).

- *Tài liệu tham chiếu*: Trích dẫn các điều khoản tại Luật Khoa học, Công nghệ và Đổi mới sáng tạo, Luật Sở hữu trí tuệ và các Nghị định, Thông tư hướng dẫn làm căn cứ định nghĩa thuật ngữ.

3.4.2. Định nghĩa kỹ thuật

Nhóm này quy định các thông số kỹ thuật để các hệ thống thông tin có thể tự động hiểu, xử lý và trao đổi gói tin dữ liệu chuẩn hóa.

- *Tên trường kỹ thuật*: Tên định danh dùng trong mã nguồn và cơ sở dữ liệu, tuân thủ quy tắc đặt tên viết liền không dấu, đảm bảo tương thích với các chuẩn kết nối quốc tế.

- *Kiểu dữ liệu và định dạng*: Quy định định dạng lưu trữ vật lý (chuỗi, số, ngày tháng...) và cấu trúc hiển thị bắt buộc (ví dụ: định dạng ngày tháng theo chuẩn YYYY-MM-DD).

- *Kiểu dữ liệu mô hình*: Sử dụng các chuẩn mô tả ngữ nghĩa để hỗ trợ dữ liệu liên kết và truy vấn thông minh, đảm bảo khả năng tương tác ở mức độ cao.

- *Quan hệ kỹ thuật*: Định nghĩa các ràng buộc và liên kết kỹ thuật giữa các trường thông tin (như: quan hệ khóa chính, khóa ngoại) để đảm bảo tính toàn vẹn dữ liệu xuyên hệ thống.

- *Danh sách giá trị*: Đối với các thuật ngữ dạng danh mục, phải liệt kê đầy đủ các mã giá trị cho phép và ý nghĩa tương ứng (ví dụ: 01 - Đang thực hiện, 02 - Đã nghiệm thu).

- *Quy tắc chuẩn hóa và Kiểm tra*: Thiết lập các điều kiện bắt buộc để dữ liệu được chấp nhận vào hệ thống (ví dụ: quy tắc về độ dài, tính bắt buộc phải có thông tin, quy tắc kiểm tra định dạng mã số).

- *API truy cập*: Chỉ dẫn địa chỉ dịch vụ dữ liệu trên nền tảng LGSP/LDOP của Bộ để các đơn vị có thẩm quyền thực hiện kết nối và khai thác trực tiếp.

- *Tương thích chuẩn quốc tế*: Mô tả khả năng ánh xạ (mapping) với các chuẩn dữ liệu quốc tế chuyên ngành mà Việt Nam tham gia (ví dụ: chuẩn ST.96 của WIPO về dữ liệu sở hữu trí tuệ).

3.5. Quy trình xây dựng và cập nhật

Bước 1. Phân tích và rà soát

Các đơn vị chuyên ngành thuộc Bộ thực hiện:

- Đánh giá nhu cầu dữ liệu dùng chung trong phạm vi quản lý nhà nước của đơn vị.

- Rà soát toàn bộ các văn bản quy phạm pháp luật, biểu mẫu nghiệp vụ và các hệ thống thông tin hiện có.

- Tổng hợp danh mục các thuật ngữ, trường dữ liệu từ các cơ sở dữ liệu chuyên ngành.

- Xác định các nội dung cần chuẩn hóa hoặc mở rộng so với danh mục dữ liệu dùng chung quốc gia.

Kết quả đầu ra: danh sách từ vựng thô (bao gồm thuật ngữ, mô tả, ngữ cảnh sử dụng).

Bước 2. Chuẩn hóa

Các đơn vị chuyên ngành thuộc Bộ thực hiện:

- Chuẩn hóa tên gọi và định nghĩa nghiệp vụ theo quy định pháp luật chuyên ngành; loại bỏ các thuật ngữ đồng nghĩa hoặc gây nhầm lẫn.

- Phân loại từ vựng thành 02 nhóm: từ vựng cốt lõi (dùng chung liên ngành) và từ vựng chuyên ngành (đặc thù KH&CN).

- Xác định rõ cơ quan chịu trách nhiệm tạo lập và sở hữu đối với từng thuật ngữ.

- Thực hiện ánh xạ (mapping) với các tiêu chuẩn quốc tế chuyên ngành như chuẩn của WIPO (Sở hữu trí tuệ) hoặc ISO (Tiêu chuẩn).

Kết quả đầu ra: danh mục từ vựng chuẩn hóa, có định nghĩa và phân loại rõ ràng.

Bước 3. Mô hình hóa

Các đơn vị chuyên ngành phối hợp với Trung tâm Công nghệ thông tin thực hiện:

- Thiết lập cấu trúc kỹ thuật cho từng thuật ngữ: tên trường kỹ thuật (không dấu), kiểu dữ liệu, định dạng và quy tắc kiểm tra.

- Xây dựng mối quan hệ phân cấp và liên kết giữa các thực thể dữ liệu.

- Mô hình hóa dữ liệu theo các chuẩn kỹ thuật quốc tế (RDF/OWL) để đảm bảo khả năng truy vấn ngữ nghĩa.

Kết quả đầu ra: mô hình từ vựng có cấu trúc kỹ thuật thống nhất, sẵn sàng cho việc tích hợp hệ thống.

Bước 4. Xác thực

- Trung tâm Công nghệ thông tin chủ trì, phối hợp với các đơn vị liên quan rà soát tính thống nhất của mô hình.

- Tổ chức lấy ý kiến tham vấn của các bên liên quan đối với các thuật ngữ có tính liên ngành cao.

- Tổng hợp và trình Lãnh đạo Bộ phê duyệt danh mục Từ điển dữ liệu dùng chung của Bộ.

Bước 5. Công bố

- Sau khi có Quyết định phê duyệt của Lãnh đạo Bộ, Trung tâm Công nghệ thông tin thực hiện đăng ký và công bố công khai trên Hệ thống Từ điển dữ liệu dùng chung của Bộ.

- Công bố các dịch vụ dữ liệu (API) tương ứng trên nền tảng LGSP/LDOP của Bộ để các đơn vị thực hiện kết nối.

Bước 6. Quản trị

- Định kỳ rà soát (quý hoặc năm) để cập nhật, bổ sung thuật ngữ khi có sự thay đổi về quy định pháp luật hoặc phát sinh nghiệp vụ mới.

- Thực hiện cơ chế phiên bản hóa để đảm bảo các hệ thống cũ vẫn vận hành ổn định trong khi cập nhật các tiêu chuẩn mới.

V.DANH MỤC NHIỆM VỤ VÀ LỘ TRÌNH TRIỂN KHAI

Stt	Nhiệm vụ	Sản phẩm	Đơn vị chủ trì	Đơn vị phối hợp	Thời gian thực hiện
I	Thiết lập Thể chế và Tổ chức				
1	Cập nhật Chiến lược dữ liệu của Bộ KH&CN đến năm 2030	Quyết định phê duyệt Chiến lược dữ liệu của Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Thường xuyên
2	Thành lập Ban Chỉ đạo về Dữ liệu của Bộ KH&CN	Quyết định thành lập Ban Chỉ đạo về Dữ liệu của Bộ KH&CN	Cục Thông tin, thống kê	Vụ Tổ chức cán bộ và các đơn vị có liên quan	Quý I/2026
3	Ban hành Quy chế hoạt động Ban Chỉ đạo về dữ liệu của Bộ KH&CN	- Quy chế hoạt động Ban Chỉ đạo về dữ liệu của Bộ KH&CN	Trung tâm Công nghệ thông tin	Vụ Tổ chức cán bộ và các đơn vị có liên quan	4/2026
4	Ban hành quy định về quản trị, khai thác và chia sẻ dữ liệu trong Bộ KH&CN	Quyết định phê duyệt Quy chế	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026

5	Rà soát, cập nhật danh mục dữ liệu dùng chung, danh mục dữ liệu mở của Bộ KH&CN	Danh mục dữ liệu dùng chung, danh mục dữ liệu mở	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Nhiệm vụ thường xuyên
6	Kế hoạch triển khai nhiệm vụ Kho dữ liệu tập trung của Bộ KH&CN	Kế hoạch triển khai nhiệm vụ Kho dữ liệu tập trung của Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026 - 2027
7	Ban hành quy định an toàn dữ liệu và bảo vệ dữ liệu đáp ứng Khung kiến trúc dữ liệu của Bộ KH&CN	Quyết định/Quy chế về an toàn dữ liệu và bảo vệ dữ liệu của Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026
II Quy hoạch dữ liệu					
1	Xây dựng Quy hoạch dữ liệu của Bộ KH&CN	Quy hoạch dữ liệu của Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026 - 2027
2	Xây dựng, cập nhật Từ điển dữ liệu dùng chung của Bộ KH&CN	Danh mục từ điển dữ liệu	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Nhiệm vụ thường xuyên

III	Triển khai các công cụ kỹ thuật				
1	Nâng cấp, hoàn thiện Nền tảng tích hợp, chia sẻ dữ liệu của Bộ KH&CN (LGSP)	Nền tảng tích hợp, chia sẻ dữ liệu của Bộ	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Nhiệm vụ thường xuyên
2	Xây dựng Nền tảng phân tích, xử lý dữ liệu tổng hợp Bộ KH&CN	Nền tảng phân tích, xử lý dữ liệu tổng hợp Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026 - 2030
3	Xây dựng Kho dữ liệu tập trung của Bộ KH&CN	Kho dữ liệu tập trung Bộ KH&CN	Trung tâm Công nghệ thông tin	Các đơn vị có liên quan	Năm 2026 - 2027
IV	Vận hành				
1	Triển khai các chiến dịch làm sạch và chuyển đổi dữ liệu	Xây dựng Kế hoạch và tổ chức triển khai	Các đơn vị quản lý CSDL	Trung tâm Công nghệ thông tin	Nhiệm vụ thường xuyên
2	Thực hiện tích hợp dữ liệu, làm sạch, chuẩn hóa và các bài toán phân tích tại Nền tảng phân tích, xử lý dữ liệu tổng hợp	Data Mart chuyên biệt phục vụ báo cáo, phân tích	Các đơn vị quản lý CSDL	Trung tâm Công nghệ thông tin	Nhiệm vụ thường xuyên

V	Xây dựng Văn hóa dữ liệu và Đào tạo				
1	Đào tạo tư duy dữ liệu cho lãnh đạo và kỹ năng phân tích cho cán bộ, công chức, viên chức, người lao động của Bộ	Kế hoạch đào tạo	Các đơn vị thuộc Bộ	Trung tâm Công nghệ thông tin	Nhiệm vụ thường xuyên

VI. GIẢI PHÁP QUẢN TRỊ

1. Nguyên tắc triển khai Khung kiến trúc và quản trị dữ liệu

1.1. Tính thống nhất và tuân thủ kiến trúc

- Khung kiến trúc dữ liệu và Khung quản trị dữ liệu là căn cứ pháp lý, kỹ thuật bắt buộc để các cơ quan, đơn vị thuộc Bộ lập thiết kế cơ sở, thẩm định và triển khai các chương trình, dự án, nhiệm vụ có nội dung xây dựng, vận hành hệ thống thông tin.

- Mọi hoạt động tạo lập, quản lý và khai thác dữ liệu phải bảo đảm tính đồng bộ, liên thông theo mô hình kiến trúc tổng thể của Bộ nhằm tối ưu hóa nguồn lực và khả năng chia sẻ dữ liệu giữa các đơn vị nghiệp vụ chuyên ngành.

- Các đề xuất đầu tư không chứng minh được sự phù hợp với Khung kiến trúc hoặc không đáp ứng các tiêu chuẩn về kết nối, chia sẻ dữ liệu dùng chung sẽ không được xem xét phê duyệt.

- Trường hợp phát sinh các yêu cầu nghiệp vụ đặc thù nằm ngoài phạm vi của Khung, đơn vị chủ trì có trách nhiệm báo cáo Bộ trưởng (thông qua Trung tâm Công nghệ thông tin) để xem xét điều chỉnh, cập nhật Khung trước khi thực hiện.

1.2. Quản trị dữ liệu tập trung và thúc đẩy chia sẻ

- Hoạt động quản lý dữ liệu được thực hiện theo nguyên tắc quản trị dữ liệu thống nhất, xuyên suốt vòng đời dữ liệu, bảo đảm tính xác thực, toàn vẹn và được quản lý bởi đơn vị chủ quản nghiệp vụ chịu trách nhiệm về chất lượng dữ liệu.

- Cấu trúc dữ liệu của Bộ được phân lớp khoa học thành: dữ liệu chủ, dữ liệu danh mục chuyên ngành (Tiêu chuẩn đo lường chất lượng, sở hữu trí tuệ, KH&CN...), dữ liệu tổng hợp và dữ liệu mở.

- Nền tảng tích hợp, chia sẻ dữ liệu của Bộ (LGSP/LDOP) là trung tâm điều phối duy nhất cho việc kết nối giữa các đơn vị; thực hiện nguyên tắc "chia sẻ dữ liệu là mặc định", hạn chế tối đa việc cung cấp dữ liệu thủ công hoặc phân mảnh.

- Khuyến khích việc tái sử dụng dữ liệu từ Kho dữ liệu dùng chung để nâng cao năng lực phân tích, dự báo và hỗ trợ ra quyết định trong công tác quản lý nhà nước của Bộ.

1.3. Bảo đảm an toàn, an ninh dữ liệu và bảo vệ dữ liệu cá nhân

- An toàn, an ninh dữ liệu là điều kiện tiên quyết và xuyên suốt trong quá trình hình thành, vận hành Khung kiến trúc, tuân thủ nghiêm ngặt các quy định về an toàn hệ thống thông tin theo cấp độ và bảo vệ bí mật nhà nước.

- Thực hiện phân loại và gán nhãn dữ liệu theo độ mật và tính chất nhạy cảm (như bí mật công nghệ, hồ sơ chuyên gia) để áp dụng các chính sách kiểm soát truy cập, mã hóa và giám sát phù hợp.

- Thiết lập cơ chế giám sát tự động và kiểm tra định kỳ để kịp thời phát hiện, xử lý các hành vi truy cập trái phép hoặc làm lộ lọt dữ liệu ngành.

1.4. Duy trì, cập nhật Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu của Bộ

- Khung kiến trúc dữ liệu và Khung quản trị, quản lý dữ liệu của Bộ phải được duy trì, rà soát và cập nhật định kỳ nhằm bảo đảm sự tương thích với tầm nhìn, mục tiêu chiến lược phát triển khoa học, công nghệ và đổi mới sáng tạo; đồng thời bảo đảm tính thống nhất với Khung kiến trúc dữ liệu quốc gia, Khung quản trị, quản lý dữ liệu quốc gia và các quy định pháp luật hiện hành.

- Quá trình rà soát, cập nhật được thực hiện dựa trên việc đánh giá định kỳ hiện trạng tổ chức, mô hình dữ liệu, cơ chế kết nối và hiệu quả khai thác dữ liệu thực tế tại các đơn vị trực thuộc Bộ. Đồng thời, Bộ thực hiện đánh giá tác động kịp thời khi các văn bản quy phạm pháp luật hoặc các phiên bản Khung kiến trúc dữ liệu quốc gia được điều chỉnh, bổ sung hoặc thay thế.

- Công tác cập nhật phải bảo đảm tính khả thi, phù hợp với định hướng phát triển Bộ KH&CN trong từng giai đoạn. Lộ trình thực hiện cần được thiết kế tương ứng với nguồn lực tài chính, năng lực hạ tầng kỹ thuật và điều kiện triển khai thực tế của Bộ.

- Các nội dung thay đổi trong Khung kiến trúc và quản trị dữ liệu phải được xác định và phân loại theo mức độ tác động để lựa chọn phương thức điều chỉnh phù hợp. Quá trình này ưu tiên tính kế thừa, bảo đảm việc triển khai đồng bộ và tuyệt đối không gây gián đoạn các hoạt động quản lý, khai thác dữ liệu hiện hành của các đơn vị thuộc Bộ.

- Kết quả cập nhật Khung kiến trúc và quản trị dữ liệu là căn cứ pháp lý cao nhất để điều chỉnh, phê duyệt các chương trình, đề án và dự án đầu tư về dữ liệu

của Bộ. Điều này đảm bảo sự thống nhất toàn diện trong công tác tổ chức, quản trị và khai thác nguồn tài nguyên dữ liệu chuyên ngành trong hệ thống quản lý nhà nước của Bộ.

2. Giải pháp thực hiện

2.1. Giải pháp về nguồn lực

- Kiến toàn bộ máy quản trị dữ liệu tại các đơn vị thuộc Bộ; xác định rõ vai trò, trách nhiệm của bộ phận chuyên trách về kiến trúc và kỹ thuật dữ liệu tại từng cơ quan.

- Đào tạo nâng cao năng lực số cho lãnh đạo và bồi dưỡng kỹ năng phân tích chuyên sâu cho cán bộ chuyên môn nhằm hình thành văn hóa quản trị dựa trên dữ liệu.

- Xây dựng cơ chế thu hút chuyên gia đầu ngành về dữ liệu và an toàn thông tin tham gia hỗ trợ triển khai các nền tảng dữ liệu chuyên ngành của Bộ.

2.2. Giải pháp về cơ chế, chính sách

- Ban hành và hoàn thiện các quy chế về kết nối, tích hợp và chia sẻ dữ liệu nội bộ giữa các đơn vị thuộc Bộ và giữa Bộ với các cơ quan Trung ương, địa phương.

- Chuẩn hóa danh mục điện tử, từ điển dữ liệu và mã định danh duy nhất cho các thực thể quản lý trong toàn hệ thống.

- Thiết lập bộ chỉ số đánh giá mức độ trưởng thành về dữ liệu (KPI/KQI) và gắn trách nhiệm của người đứng đầu đơn vị với kết quả thực hiện quản trị dữ liệu.

2.3. Giải pháp về tài chính

- Đảm bảo ưu tiên nguồn kinh phí từ ngân sách KH&CN hàng năm cho việc phát triển hạ tầng số, duy trì Kho dữ liệu dùng chung và các hoạt động làm sạch dữ liệu chuyên ngành.

- Lồng ghép tối ưu các nhiệm vụ dữ liệu trong kế hoạch chuyển đổi số và cải cách hành chính, đảm bảo đầu tư có trọng tâm, tránh lãng phí.

3. Tổ chức thực hiện

Ban Chỉ đạo về dữ liệu của Bộ: Ban Chỉ đạo về dữ liệu Bộ chỉ đạo chung việc triển khai Khung kiến trúc dữ liệu, Khung quản trị, quản lý dữ liệu Bộ.

Trung tâm Công nghệ thông tin: Là đơn vị đầu mối, thường trực; chủ trì hướng dẫn, giám sát việc tuân thủ Khung kiến trúc và quản lý vận hành hạ tầng tích hợp, chia sẻ dữ liệu của Bộ.

Các đơn vị trực thuộc Bộ: Thực hiện rà soát, số hóa và chuẩn hóa dữ liệu chuyên ngành thuộc thẩm quyền quản lý, chịu trách nhiệm đảm bảo dữ liệu “Đúng - Đủ - Sạch - Sóng - Thống nhất - Dùng chung”.

Vụ Kế hoạch - Tài chính: Chủ trì, phối hợp với các đơn vị trực thuộc trong công tác tổng hợp, tham mưu Lãnh đạo Bộ trình cấp có thẩm quyền bố trí kinh phí từ nguồn ngân sách nhà nước để triển khai các nhiệm vụ trong Kế hoạch này.

PHỤ LỤC

DANH MỤC DỮ LIỆU CHI TIẾT

(Kèm theo Quyết định số /QĐ-BKHCN ngày tháng năm 2026
của Bộ Khoa học và Công nghệ)

1. Danh mục siêu dữ liệu

STT	Nhóm siêu dữ liệu	Mô tả một số danh mục siêu dữ liệu
1	Phục vụ người dân, doanh nghiệp	Siêu dữ liệu định danh, mô tả, cấu trúc, nguồn gốc, quy trình tạo lập và sử dụng, đơn vị chủ quản, chia sẻ và truy vết dữ liệu đối với các tập dữ liệu về thủ tục hành chính, dịch vụ công và hồ sơ giải quyết thủ tục hành chính...
2	Khoa học và Công nghệ	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chất lượng và khai thác dữ liệu đối với các tập dữ liệu về nhiệm vụ KH&CN, tổ chức KH&CN, nhà khoa học và kết quả nghiên cứu.
3	Đổi mới sáng tạo	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chất lượng và khai thác dữ liệu đối với các tập dữ liệu về doanh nghiệp khởi nghiệp sáng tạo, tổ chức hỗ trợ, chương trình hỗ trợ, quỹ đầu tư và hệ sinh thái đổi mới sáng tạo.
4	Chuyển đổi số, công nghiệp công nghệ số	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, dòng chảy, thời gian và vòng đời, cập nhật, quản trị, chất lượng và khai thác dữ liệu đối với các tập dữ liệu về hệ thống thông tin, nền tảng số, doanh nghiệp công nghệ số, tổ chức cung cấp dịch

		vụ số và giấy phép, chứng nhận trong lĩnh vực...
5	Sở hữu trí tuệ	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chất lượng, truy vết và khai thác dữ liệu đối với các tập dữ liệu về đơn đăng ký, văn bằng bảo hộ, đối tượng sở hữu công nghiệp, chủ thể quyền và quá trình xử lý hồ sơ sở hữu trí tuệ...
6	Tiêu chuẩn, đo lường, chất lượng	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chuẩn hóa, chất lượng và khai thác dữ liệu đối với các tập dữ liệu về tiêu chuẩn, quy chuẩn kỹ thuật, đo lường, đánh giá sự phù hợp, sản phẩm, hàng hóa và tổ chức liên quan...
7	Viễn thông, tần số, Internet	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, dòng chảy, thời gian và vòng đời, cập nhật, quản trị, chất lượng, truy vết và khai thác dữ liệu đối với các tập dữ liệu về giấy phép viễn thông, tài nguyên tần số, tài nguyên Internet, vi phạm và xử lý can nhiễu...
8	Năng lượng nguyên tử, ATBX	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, mức độ nhạy cảm, chất lượng, truy vết và khai thác dữ liệu đối với các tập dữ liệu về giấy phép, cơ sở bức xạ, nguồn phóng xạ, nhân sự bức xạ và hoạt động bảo đảm an toàn bức xạ, hạt nhân...

9	Bưu chính	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chất lượng và khai thác dữ liệu đối với các tập dữ liệu về doanh nghiệp bưu chính, giấy phép hoạt động bưu chính, bộ tem, mẫu tem và các sản phẩm, ấn phẩm bưu chính...
10	Thẻ chế, chính sách	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, truy vết và khai thác dữ liệu đối với các tập dữ liệu về văn bản quy phạm pháp luật, văn bản quản lý, chính sách, chương trình, đề án và tài liệu điều hành...
11	Quản lý, chỉ đạo điều hành	Siêu dữ liệu định danh, mô tả, cấu trúc, quan hệ, chủ thể liên quan, nguồn gốc, thời gian và vòng đời, cập nhật, quản trị, chất lượng, truy vết và khai thác dữ liệu đối với các tập dữ liệu về nhiệm vụ, chỉ tiêu, báo cáo, cảnh báo và hệ thống phục vụ công tác quản lý, chỉ đạo, điều hành...

2. Danh mục dùng chung

STT	Nhóm dữ liệu danh mục dùng chung	Mô tả một số dữ liệu danh mục dùng chung
1	Phục vụ người dân, doanh nghiệp	- Danh mục thủ tục hành chính; - Danh mục loại hồ sơ; - Danh mục trạng thái xử lý hồ sơ; - Danh mục cơ quan giải quyết; - Danh mục loại đối tượng; - Danh mục mức độ dịch vụ công.

2	Khoa học và Công nghệ	- Danh mục lĩnh vực nghiên cứu; - Danh mục loại nhiệm vụ; - Danh mục chương trình, đề án, nhiệm vụ; - Danh mục tổ chức KH&CN; - Danh mục trạng thái nhiệm vụ; - Danh mục loại sản phẩm nghiên cứu.
3	Đổi mới sáng tạo	- Danh mục loại hình startup; - Danh mục lĩnh vực công nghệ; - Danh mục loại hình tổ chức hỗ trợ khởi nghiệp sáng tạo; - Danh mục chương trình hỗ trợ đổi mới sáng tạo; - Danh mục giai đoạn phát triển startup; - Danh mục loại hình hỗ trợ - Danh mục loại quỹ đầu tư. - Danh mục loại hoạt động thương mại hóa công nghệ.
4	Chuyển đổi số, công nghiệp công nghệ số	- Danh mục loại nền tảng / hệ thống thông tin; - Danh mục loại cơ sở dữ liệu; - Danh mục loại doanh nghiệp công nghệ số; - Danh mục loại sản phẩm, giải pháp số; - Danh mục chỉ số, mức độ chuyển đổi số. - Danh mục loại giấy phép, chứng nhận.
5	Sở hữu trí tuệ	- Danh mục loại đối tượng sở hữu trí tuệ; - Danh mục loại đơn; - Danh mục trạng thái đơn;

		- Danh mục trạng thái văn bằng; - Danh mục loại hình bảo hộ; - Danh mục nhóm sản phẩm, dịch vụ; - Danh mục loại yêu cầu xử lý.
6	Tiêu chuẩn, đo lường, chất lượng	- Danh mục loại tiêu chuẩn; - Danh mục loại quy chuẩn kỹ thuật; - Danh mục lĩnh vực áp dụng; - Danh mục loại hình đánh giá sự phù hợp; - Danh mục nhóm sản phẩm, hàng hóa; - Danh mục trạng thái hiệu lực.
7	Viễn thông, tần số, Internet	- Danh mục loại giấy phép; - Danh mục loại tài nguyên tần số; - Danh mục loại mạng viễn thông; - Danh mục đơn vị tần số; - Danh mục loại tài nguyên Internet; - Danh mục loại vi phạm; - Danh mục loại nhiễu.
8	Năng lượng nguyên tử, ATBX	- Danh mục loại giấy phép; - Danh mục loại nguồn phóng xạ; - Danh mục loại cơ sở bức xạ; - Danh mục loại chứng chỉ; - Danh mục mức độ an toàn; - Danh mục mức độ rủi ro; - Danh mục trạng thái cấp phép.

9	Bưu chính	- Danh mục loại hình doanh nghiệp bưu chính; - Danh mục loại dịch vụ bưu chính; - Danh mục phạm vi hoạt động; - Danh mục loại giấy phép; - Danh mục trạng thái hoạt động; - Danh mục loại sản phẩm bưu chính; - Danh mục loại tem.
10	Thể chế, chính sách	- Danh mục loại văn bản; - Danh mục lĩnh vực quản lý; - Danh mục cơ quan ban hành; - Danh mục trạng thái hiệu lực; - Danh mục loại chương trình, đề án;
11	Quản lý, chỉ đạo điều hành	- Danh mục loại nhiệm vụ; - Danh mục trạng thái nhiệm vụ; - Danh mục loại chỉ tiêu; - Danh mục loại báo cáo; - Danh mục mức độ hoàn thành;

3. Danh mục Dữ liệu chủ

STT	Nhóm dữ liệu chủ	Mô tả một số danh mục dữ liệu chủ
1	Phục vụ người dân, doanh nghiệp	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi phục vụ cung cấp dịch vụ công, giải quyết thủ tục hành chính và tương tác với người dân, doanh nghiệp.

		Một số dữ liệu chủ cơ bản: Thủ tục hành chính; Dịch vụ công; Hồ sơ thủ tục hành chính; Thành phần hồ sơ; Kết quả giải quyết thủ tục hành chính; Cá nhân; Tổ chức; Doanh nghiệp; Cơ quan, đơn vị giải quyết thủ tục hành chính;...
2	Khoa học và Công nghệ	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong hoạt động khoa học và công nghệ. Một số dữ liệu chủ cơ bản: Nhiệm vụ khoa học và công nghệ; Hồ sơ nhiệm vụ khoa học và công nghệ; Chương trình, đề án khoa học và công nghệ; Tổ chức khoa học và công nghệ; Doanh nghiệp khoa học và công nghệ; Nhà khoa học, chuyên gia; Kết quả nghiên cứu; Sản phẩm khoa học và công nghệ;...
3	Đổi mới sáng tạo	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong hệ sinh thái đổi mới sáng tạo. Một số dữ liệu chủ cơ bản: Doanh nghiệp khởi nghiệp sáng tạo; Tổ chức hỗ trợ đổi mới sáng tạo; Quỹ đầu tư; Chương trình hỗ trợ đổi mới sáng tạo; Dự án, sáng kiến đổi mới sáng tạo; Sản phẩm, giải pháp đổi mới sáng tạo;...
4	Chuyển đổi số, công nghiệp công nghệ số	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong hệ sinh thái số và công nghiệp công nghệ số. Một số dữ liệu chủ cơ bản: Hệ thống thông tin; Cơ sở dữ liệu; Nền tảng số; Dịch vụ số; Doanh nghiệp công nghệ số; Tổ chức cung cấp dịch vụ số; Sản phẩm, giải pháp công nghệ số; Giấy phép, chứng nhận trong lĩnh vực công nghệ số;...

5	Sở hữu trí tuệ	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong lĩnh vực sở hữu trí tuệ và sở hữu công nghiệp. Một số dữ liệu chủ cơ bản: Đơn đăng ký sở hữu công nghiệp; Văn bằng bảo hộ; Đối tượng sở hữu công nghiệp; Chủ thể quyền sở hữu trí tuệ; Người nộp đơn; Đại diện sở hữu công nghiệp;...
6	Tiêu chuẩn, đo lường, chất lượng	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong lĩnh vực tiêu chuẩn, đo lường, chất lượng và đánh giá sự phù hợp. Một số dữ liệu chủ cơ bản: Tiêu chuẩn; Quy chuẩn kỹ thuật; Sản phẩm, hàng hóa; Tổ chức đánh giá sự phù hợp; Phòng thử nghiệm, hiệu chuẩn; Mã số mã vạch/GTIN; Doanh nghiệp, tổ chức liên quan;...
7	Viễn thông, tần số, Internet	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong quản lý tài nguyên, hạ tầng và hoạt động viễn thông, tần số, Internet. Một số dữ liệu chủ cơ bản: Giấy phép viễn thông, tần số, Internet; Tài nguyên tần số; Tài nguyên Internet; Tên miền; Địa chỉ IP; Doanh nghiệp viễn thông, công nghệ thông tin;...
8	Năng lượng nguyên tử, ATBX	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong quản lý năng lượng nguyên tử, nguồn phóng xạ và an toàn bức xạ. Một số dữ liệu chủ cơ bản: Giấy phép trong lĩnh vực năng lượng nguyên tử, an toàn bức xạ; Cơ sở bức xạ; Nguồn phóng xạ; Thiết bị bức xạ; Chứng chỉ chuyên môn;...

9	Bưu chính	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong lĩnh vực bưu chính và sản phẩm bưu chính. Một số dữ liệu chủ cơ bản: Doanh nghiệp bưu chính; Dịch vụ bưu chính; Giấy phép hoạt động bưu chính; Mạng lưới bưu chính; Bộ tem bưu chính; Mẫu tem; Ấn phẩm, sản phẩm bưu chính;...
10	Thể chế, chính sách	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi trong xây dựng, quản lý và theo dõi thể chế, chính sách. Một số dữ liệu chủ cơ bản: Văn bản quy phạm pháp luật; Văn bản quản lý; Chính sách; Chương trình, đề án; Kế hoạch thực hiện;...
11	Quản lý, chỉ đạo điều hành	Bao gồm danh mục dữ liệu chủ về các đối tượng cốt lõi phục vụ công tác quản lý, giám sát, theo dõi và chỉ đạo điều hành. Một số dữ liệu chủ cơ bản: Nhiệm vụ; Chỉ tiêu; Báo cáo; Cảnh báo; Đơn vị thực hiện; Hệ thống điều hành, giám sát;...

4. Danh mục Dữ liệu giao dịch (Chuyên ngành)

STT	Nhóm dữ liệu giao dịch/chuyên ngành	Mô tả một số danh mục dữ liệu giao dịch/chuyên ngành
1	Phục vụ người dân, doanh nghiệp	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình tiếp nhận, xử lý, giải quyết hồ sơ và cung cấp dịch vụ công cho người dân, doanh nghiệp.

		Một số dữ liệu nghiệp vụ (giao dịch) cơ bản: Hồ sơ tiếp nhận; Phiếu xử lý hồ sơ; Kết quả giải quyết hồ sơ; Lịch sử xử lý hồ sơ; Giao dịch nộp hồ sơ; Giao dịch thanh toán phí, lệ phí; Phản ánh, kiến nghị; Đánh giá mức độ hài lòng;..
2	Khoa học và Công nghệ	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình đề xuất, tuyển chọn, phê duyệt, triển khai, theo dõi, đánh giá và nghiệm thu nhiệm vụ khoa học và công nghệ. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ đề xuất nhiệm vụ; Hồ sơ tuyển chọn, giao trực tiếp; Quyết định phê duyệt; Hợp đồng thực hiện nhiệm vụ; Báo cáo tiến độ; Báo cáo kết quả; Hồ sơ nghiệm thu; Hồ sơ thanh quyết toán; Hồ sơ đánh giá, xếp loại;..
3	Đổi mới sáng tạo	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình tiếp nhận, xét chọn, hỗ trợ, theo dõi, đánh giá các doanh nghiệp, tổ chức và hoạt động đổi mới sáng tạo. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ đăng ký tham gia chương trình hỗ trợ; Hồ sơ xét chọn startup; Hồ sơ công nhận doanh nghiệp khởi nghiệp sáng tạo; Báo cáo hoạt động; Hồ sơ hỗ trợ tài chính; Hồ sơ kết nối đầu tư; Hồ sơ đánh giá kết quả hỗ trợ.
4	Chuyên đổi số, công nghiệp công nghệ số	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình đăng ký, cấp phép, quản lý, giám sát, vận hành và đánh giá các hệ thống, nền tảng, doanh nghiệp và dịch vụ số. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ đăng ký hệ thống; Hồ sơ cấp phép; Hồ sơ chứng nhận;

		Nhật ký vận hành; Dữ liệu giám sát hệ thống; Dữ liệu sử dụng nền tảng số; Báo cáo an toàn thông tin; Báo cáo đánh giá, kiểm tra;..
5	Sở hữu trí tuệ	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình nộp đơn, thẩm định, công bố, cấp văn bằng, sửa đổi, gia hạn, chuyển nhượng và xử lý các hồ sơ sở hữu trí tuệ. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ nộp đơn; Hồ sơ thẩm định hình thức; Hồ sơ thẩm định nội dung; Dữ liệu công bố đơn; Dữ liệu cấp văn bằng; Hồ sơ sửa đổi, gia hạn; Hồ sơ chuyển nhượng; Lịch sử xử lý đơn; Kết quả xử lý hồ sơ;..
6	Tiêu chuẩn, đo lường, chất lượng	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình xây dựng, ban hành, áp dụng, kiểm tra, đánh giá và quản lý các hoạt động tiêu chuẩn, đo lường, chất lượng và đánh giá sự phù hợp. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ xây dựng tiêu chuẩn, quy chuẩn; Hồ sơ công bố; Hồ sơ đánh giá sự phù hợp; Kết quả kiểm định, thử nghiệm, hiệu chuẩn; Hồ sơ chứng nhận; Hồ sơ thanh tra, kiểm tra; Báo cáo đánh giá chất lượng;..
7	Viễn thông, tần số, Internet	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình cấp phép, phân bổ, giám sát, kiểm tra, xử lý vi phạm và xử lý can nhiễu trong lĩnh vực viễn thông, tần số, Internet. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ cấp phép; Hồ sơ gia hạn, sửa đổi giấy phép; Dữ liệu

		phân bổ tài nguyên; Báo cáo giám sát; Báo cáo vi phạm; Hồ sơ xử lý can nhiễu; Biên bản kiểm tra; Biên bản xử lý vi phạm;..
8	Năng lượng nguyên tử, ATBX	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình cấp phép, đăng ký, kiểm tra, giám sát, đánh giá và quản lý an toàn bức xạ, hạt nhân. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ cấp phép; Hồ sơ đăng ký; Hồ sơ cấp chứng chỉ; Báo cáo an toàn bức xạ; Kết quả kiểm tra, giám sát; Hồ sơ thanh tra, kiểm tra; Hồ sơ gia hạn, sửa đổi giấy phép;..
9	Bưu chính	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình cấp phép, xác nhận hoạt động, quản lý dịch vụ bưu chính và phát hành các sản phẩm bưu chính. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ cấp phép; Hồ sơ xác nhận thông báo hoạt động bưu chính; Hồ sơ đăng ký dịch vụ; Dữ liệu phát hành tem; Hồ sơ thiết kế, phê duyệt mẫu tem; Báo cáo hoạt động doanh nghiệp bưu chính; Hồ sơ kiểm tra, giám sát;..
10	Thể chế, chính sách	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình xây dựng, lấy ý kiến, thẩm định, ban hành, theo dõi thi hành và đánh giá văn bản, chính sách, chương trình, đề án. Một số dữ liệu nghiệp vụ cơ bản: Hồ sơ xây dựng văn bản; Hồ sơ lấy ý kiến; Hồ sơ thẩm định; Hồ sơ trình ban hành; Dữ liệu theo dõi thi

		hành; Báo cáo rà soát, đánh giá; Hồ sơ sửa đổi, bổ sung chính sách;..
11	Quản lý, chỉ đạo điều hành	Bao gồm danh mục dữ liệu nghiệp vụ phát sinh trong quá trình giao nhiệm vụ, theo dõi tiến độ, tổng hợp báo cáo, giám sát, cảnh báo và điều hành công việc. Một số dữ liệu nghiệp vụ cơ bản: Giao việc; Cập nhật tiến độ; Báo cáo định kỳ, đột xuất; Cảnh báo quá hạn; Nhật ký xử lý công việc; Kết quả thực hiện nhiệm vụ; Dữ liệu giám sát, điều hành;..

5. Danh mục Dữ liệu gốc

STT	Nhóm dữ liệu gốc	Mô tả một số danh mục dữ liệu gốc
1	Phục vụ người dân, doanh nghiệp	Là tập hợp dữ liệu được tạo lập lần đầu, có giá trị pháp lý, phản ánh trực tiếp thông tin về cá nhân, tổ chức/doanh nghiệp trong quá trình cung cấp dịch vụ công, giải quyết thủ tục hành chính, và các giao dịch với cơ quan nhà nước; được quản lý bởi cơ quan có thẩm quyền và là nguồn dữ liệu chuẩn để chia sẻ, khai thác.
2	Khoa học và Công nghệ	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình thực hiện hoạt động khoa học, công nghệ và đổi mới sáng tạo; phản ánh trực tiếp kết quả nghiên cứu, thông tin nhiệm vụ KH&CN, tổ chức, cá nhân, tài sản trí tuệ... và có giá trị pháp lý, chuyên môn, được cơ quan có thẩm

		quyền quản lý làm nguồn dữ liệu chuẩn để lưu trữ, khai thác và chia sẻ.
3	Đổi mới sáng tạo	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình hình thành, vận hành và phát triển hệ sinh thái đổi mới sáng tạo; phản ánh trực tiếp thông tin về doanh nghiệp khởi nghiệp sáng tạo, tổ chức hỗ trợ, quỹ đầu tư, hoạt động ươm tạo – tăng tốc, kết quả đổi mới và thương mại hóa công nghệ; có giá trị pháp lý và thực tiễn, được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn để khai thác, chia sẻ và phục vụ hoạch định chính sách.
4	Chuyển đổi số, công nghiệp công nghệ số	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình quản lý, triển khai và vận hành các hoạt động chuyển đổi số, phát triển công nghiệp công nghệ số; phản ánh trực tiếp thông tin về hạ tầng số, nền tảng số, hệ thống thông tin, doanh nghiệp công nghệ số, sản phẩm – dịch vụ số, và mức độ chuyển đổi số của các cơ quan, tổ chức; có giá trị pháp lý và quản lý, được cơ quan có thẩm quyền xác lập, duy trì làm nguồn dữ liệu chuẩn để chia sẻ và khai thác.
5	Sở hữu trí tuệ	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình xác lập, đăng ký, thẩm định, bảo hộ và khai thác quyền sở hữu trí tuệ; phản ánh trực tiếp thông tin về đối tượng SHTT (sáng chế, kiểu dáng công nghiệp, nhãn hiệu, quyền tác giả...), chủ thể quyền và tình trạng pháp lý; có giá trị pháp lý, được cơ quan có thẩm quyền

		quản lý làm nguồn dữ liệu chuẩn phục vụ tra cứu, bảo hộ, thực thi và chia sẻ.
6	Tiêu chuẩn, đo lường, chất lượng	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình xây dựng, ban hành, áp dụng tiêu chuẩn – quy chuẩn kỹ thuật; hoạt động đo lường; đánh giá sự phù hợp và quản lý chất lượng sản phẩm, hàng hóa; phản ánh trực tiếp thông tin về tiêu chuẩn, phương tiện đo, kết quả thử nghiệm – kiểm định – hiệu chuẩn, chứng nhận, công bố hợp chuẩn/hợp quy; có giá trị pháp lý, kỹ thuật và được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn phục vụ quản lý nhà nước, tra cứu và chia sẻ.
7	Viễn thông, tần số, Internet	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình quản lý, cấp phép, khai thác và vận hành hạ tầng viễn thông, tài nguyên tần số vô tuyến điện và dịch vụ Internet; phản ánh trực tiếp thông tin về mạng viễn thông, thuê bao, tài nguyên tần số, tên miền, địa chỉ Internet, doanh nghiệp cung cấp dịch vụ và lưu lượng khai thác; có giá trị pháp lý, kỹ thuật và được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn phục vụ quản lý nhà nước, điều hành, giám sát và chia sẻ.
8	Năng lượng nguyên tử, ATBX	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình quản lý, cấp phép, kiểm soát và vận hành các hoạt động liên quan đến năng lượng nguyên tử và bức xạ hạt nhân; phản ánh trực tiếp thông tin về nguồn phóng xạ, thiết bị bức xạ, cơ

		sở hạt nhân, liều chiếu xạ, quan trắc phóng xạ môi trường, sự cố bức xạ – hạt nhân và nhân sự bức xạ; có giá trị pháp lý, kỹ thuật cao và được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn phục vụ quản lý an toàn, an ninh và ứng phó sự cố.
9	Bưu chính	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình cung ứng, quản lý và giám sát dịch vụ bưu chính; phản ánh trực tiếp thông tin về doanh nghiệp bưu chính, mạng bưu chính, điểm phục vụ, bưu gửi, hành trình vận chuyển, giá cước, chất lượng dịch vụ và các hoạt động nghiệp vụ liên quan; có giá trị pháp lý và nghiệp vụ, được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn phục vụ quản lý nhà nước, điều hành và chia sẻ.
10	Thể chế, chính sách	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình xây dựng, ban hành, tổ chức thực hiện và theo dõi, đánh giá các văn bản quy phạm pháp luật, cơ chế, chính sách; phản ánh trực tiếp thông tin về văn bản pháp luật, dự thảo, hồ sơ xây dựng chính sách, quá trình lấy ý kiến, ban hành, hiệu lực và kết quả thực thi; có giá trị pháp lý chính thức và được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn để tra cứu, áp dụng, chia sẻ.
11	Quản lý, chỉ đạo điều hành	Là tập hợp dữ liệu được tạo lập lần đầu trong quá trình lãnh đạo, quản lý, điều hành của cơ quan nhà nước; phản ánh trực tiếp thông tin về chỉ tiêu, nhiệm vụ, chương trình công tác, văn bản chỉ đạo, kết quả thực hiện, tình hình hoạt động và các báo cáo điều hành; có giá trị pháp

		lý, quản trị và được cơ quan có thẩm quyền quản lý làm nguồn dữ liệu chuẩn phục vụ theo dõi, giám sát, phân tích và ra quyết định.
--	--	--

6. Danh mục Dữ liệu tổng hợp, phân tích

STT	Nhóm dữ liệu tổng hợp, phân tích	Mô tả một số danh mục dữ liệu tổng hợp, phân tích
1	Phục vụ người dân, doanh nghiệp	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hiệu quả cung cấp dịch vụ công và giải quyết thủ tục hành chính. Một số dữ liệu tổng hợp cơ bản: Số lượng hồ sơ theo thời gian; Tỷ lệ giải quyết đúng hạn, quá hạn; Thời gian xử lý trung bình; Số lượng hồ sơ theo lĩnh vực; Mức độ hài lòng của người dân, doanh nghiệp; Tỷ lệ sử dụng dịch vụ công trực tuyến;..
2	Khoa học và Công nghệ	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động khoa học và công nghệ. Một số dữ liệu tổng hợp cơ bản: Số lượng nhiệm vụ theo lĩnh vực; Tỷ lệ hoàn thành nhiệm vụ; Tỷ lệ nghiệm thu đạt yêu cầu; Tổng kinh phí theo chương trình; Hiệu quả đầu ra nghiên cứu; Phân bổ nhiệm vụ theo địa phương, tổ chức;..
3	Đổi mới sáng tạo	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hệ sinh thái đổi mới sáng tạo. Một số dữ liệu tổng hợp cơ bản: Số lượng startup theo lĩnh vực; Số lượng doanh nghiệp

		được hỗ trợ; Tổng vốn đầu tư; Tỷ lệ tăng trưởng startup; Hiệu quả chương trình hỗ trợ; Phân bố hệ sinh thái theo địa phương;..
4	Chuyển đổi số, công nghiệp công nghệ số	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá mức độ phát triển và vận hành hệ sinh thái số. Một số dữ liệu tổng hợp cơ bản: Số lượng hệ thống thông tin; Tỷ lệ hệ thống đạt cấp độ; Mức độ sử dụng nền tảng số; Số lượng doanh nghiệp công nghệ số; Tỷ lệ dịch vụ số hóa; Mức độ an toàn thông tin;..
5	Sở hữu trí tuệ	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động đăng ký và bảo hộ sở hữu trí tuệ. Một số dữ liệu tổng hợp cơ bản: Số lượng đơn đăng ký theo thời gian; Tỷ lệ cấp văn bằng; Thời gian xử lý trung bình; Phân bố theo loại đối tượng; Số lượng đơn theo quốc gia, tổ chức; Xu hướng đăng ký sáng chế;..
6	Tiêu chuẩn, đo lường, chất lượng	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động tiêu chuẩn, đo lường, chất lượng. Một số dữ liệu tổng hợp cơ bản: Số lượng tiêu chuẩn, quy chuẩn ban hành; Số lượng sản phẩm được chứng nhận; Tỷ lệ kiểm định đạt yêu cầu; Số lượng tổ chức đánh giá sự phù hợp; Kết quả kiểm tra chất lượng;..

7	Viễn thông, tần số, Internet	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động viễn thông, tần số, Internet. Một số dữ liệu tổng hợp cơ bản: Số lượng giấy phép cấp; Tỷ lệ sử dụng tài nguyên tần số; Số lượng tên miền, địa chỉ IP; Tình hình vi phạm; Số lượng sự cố, can nhiễu; Hiệu quả khai thác hạ tầng;..
8	Năng lượng nguyên tử, ATBX	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động an toàn bức xạ và năng lượng nguyên tử. Một số dữ liệu tổng hợp cơ bản: Số lượng giấy phép cấp; Số lượng cơ sở bức xạ; Tình trạng an toàn; Số vụ vi phạm; Kết quả kiểm tra, giám sát; Mức độ tuân thủ quy định;..
9	Bưu chính	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá hoạt động bưu chính. Một số dữ liệu tổng hợp cơ bản: Số lượng doanh nghiệp bưu chính; Sản lượng dịch vụ; Phạm vi cung cấp dịch vụ; Hiệu quả hoạt động; Số lượng tem phát hành; Tình hình phát triển mạng lưới;..
10	Thế chế, chính sách	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ theo dõi, đánh giá quá trình xây dựng và thực thi chính sách. Một số dữ liệu tổng hợp cơ bản: Số lượng văn bản ban hành; Tiến độ xây dựng văn bản; Tỷ lệ văn bản đúng hạn; Kết quả rà soát, đánh giá; Hiệu quả thực thi chính sách;..

11	Quản lý, chỉ đạo điều hành	Bao gồm danh mục dữ liệu tổng hợp, phân tích phục vụ trực tiếp công tác quản lý, giám sát và chỉ đạo điều hành. Một số dữ liệu tổng hợp cơ bản: Tổng số nhiệm vụ; Tỷ lệ hoàn thành; Tỷ lệ chậm tiến độ; Phân bổ nhiệm vụ theo đơn vị; KPI theo lĩnh vực; Cảnh báo rủi ro; Xu hướng thực hiện nhiệm vụ;..
----	----------------------------	--

7. Danh mục Dữ liệu mở

STT	Nhóm dữ liệu mở	Mô tả một số danh mục tập dữ liệu mở (Dự kiến công bố)
1	Phục vụ người dân, doanh nghiệp	Danh mục thủ tục hành chính; Danh mục dịch vụ công; Biểu mẫu thủ tục hành chính; Thống kê kết quả giải quyết hồ sơ; Dữ liệu mức độ hài lòng của người dân, doanh nghiệp;...
2	Khoa học và Công nghệ	Danh mục tổ chức KH&CN; Danh mục nhiệm vụ KH&CN; Danh mục nhà khoa học; Danh mục kết quả nghiên cứu công bố; Thống kê hoạt động khoa học và công nghệ;...
3	Đổi mới sáng tạo	Danh mục doanh nghiệp khởi nghiệp sáng tạo; Danh mục tổ chức hỗ trợ đổi mới sáng tạo; Danh mục chương trình hỗ trợ; Danh mục quỹ đầu tư; Thống kê hệ sinh thái đổi mới sáng tạo;...

4	Chuyển đổi số, công nghiệp công nghệ số	Danh mục nền tảng số; Danh mục hệ thống thông tin được công bố; Danh mục doanh nghiệp công nghệ số; Danh mục sản phẩm, giải pháp số; Thống kê chuyển đổi số và công nghiệp công nghệ số;...
5	Sở hữu trí tuệ	Dữ liệu tra cứu đơn đăng ký; Dữ liệu tra cứu văn bằng bảo hộ; Danh mục đối tượng sở hữu công nghiệp; Thống kê sở hữu trí tuệ; Thông tin công bố sở hữu công nghiệp;...
6	Tiêu chuẩn, đo lường, chất lượng	Danh mục tiêu chuẩn; Danh mục quy chuẩn kỹ thuật; Danh mục tổ chức đánh giá sự phù hợp; Danh mục mã số mã vạch; Thống kê hoạt động tiêu chuẩn, đo lường, chất lượng;...
7	Viễn thông, tần số, Internet	Danh mục tài nguyên Internet; Danh mục tên miền; Danh mục giấy phép công khai; Thống kê viễn thông; Thống kê tài nguyên tần số, Internet;...
8	Năng lượng nguyên tử, ATBX	Danh mục cơ sở được công khai; Danh mục giấy phép được công khai; Thống kê hoạt động an toàn bức xạ; Báo cáo tổng hợp về an toàn bức xạ, hạt nhân;...
9	Bưu chính	Danh mục doanh nghiệp bưu chính; Danh mục điểm phục vụ bưu chính; Danh mục dịch vụ bưu chính; Danh mục tem bưu chính; Thống kê hoạt động bưu chính,...

10	Thể chế, chính sách	Văn bản quy phạm pháp luật; Văn bản quản lý được công khai; Danh mục chương trình, đề án; Chính sách công bố; Thống kê văn bản;...
11	Quản lý, chỉ đạo điều hành	Chỉ tiêu công khai; Báo cáo công khai; Thống kê kết quả thực hiện nhiệm vụ; Dữ liệu dashboard công khai;...