

Collaboratory Digital Environment

Supplement to the
Collaboratory 2030 Strategy



World Health
Organization

HUB
Pandemic and Epidemic
Intelligence

Collaboratory Digital Environment

Supplement to the
Collaboratory 2030 Strategy



Contents

Introduction	6
Reference Architecture	8
Overview	8
Experience Layer	10
Capability Domains	10
Community and Identity	10
Digital Public Goods and Knowledge Ecosystem	11
Analytics and Execution	11
Data Harmonization	12
Cross Cutting Capabilities	12
Architecture Principles	13
Conclusion	14

Introduction

Collaboratory is a flagship initiative of the World Health Organization Hub for Pandemic and Epidemic Intelligence (WHO HUB). Collaboratory convenes and connects a global public health analytics community to collectively and rapidly understand and characterize emerging health threats to support public health decision making. Since launching in 2024, Collaboratory has grown into a trusted digital environment where technical experts can exchange knowledge, share new developments and innovations, codevelop analytical tools, and rapidly mobilize analytical support during public health emergencies.

The Collaboratory 2025-2030 Strategy outlines how Collaboratory will continue maturing into a global, connected, and resilient analytics ecosystem. This supplement focuses on the opportunities and key growth areas for the Collaboratory Digital Environment and is designed to complement the work laid out in the 2030 Strategy. Some of these areas will take longer than five years, but all will need key investment within the next five years to achieve Collaboratory's vision of an interconnected global pandemic and epidemic intelligence network engaged in collaborative analytical response.

The Community Digital Infrastructure Opportunity

Alongside the community facing work, Collaboratory serves as a shared digital infrastructure for public health intelligence, enabling a global community of analysts, researchers, and public health partners to work together in support of evidence-informed decision-making. As public health threats grow in scale and complexity, there is an increasing need for a digital commons that connects data, computational resources, and analytical outputs across institutions and regions—moving beyond fragmented tools toward a cohesive, cross-cutting digital ecosystem. Within this ecosystem, access to interoperable data, trusted and scalable compute, and reusable analytical products is critical to generating timely insights that can inform preparedness, response, and recovery. The World Health Organization, through the Hub for Pandemic and Epidemic Intelligence, provides stewardship of this shared environment—setting direction, governance, and safeguards—while the global analytics community drives innovation and impact through contribution, collaboration, and collective ownership of Collaboratory's outcomes.

The Digital Infrastructure Challenge

Building a digital infrastructure capable of supporting collaborative pandemic and epidemic intelligence is challenged by a set of persistent, system-level barriers. Global analytics collaboration is still shaped by fragmented ecosystems in which data products, tools, and analytical efforts are developed in isolation, limiting interoperability and collective impact. The duplication of models and tools, alongside inconsistent data structures and practices, increases inefficiency and slows the translation of analysis into decision-relevant insights. Participation in collaborative analytics is further constrained by inequitable access to computational resources, including limited availability of secure and trusted compute environments across regions and institutions. At the same time, expertise is distributed across disconnected networks, making it difficult to identify, connect, and mobilize subject-matter experts at the pace required during emergencies. Left unaddressed, these challenges reinforce fragmentation and inequity; addressing them requires a deliberate shift toward shared digital infrastructure, common frameworks, and a coordinated ecosystem that enables collaboration at a global scale while preserving and enabling stakeholder sovereignty over data and analyses.

Alignment with Hub Data and Digital Enabler Strategy

In addition to complementing the Collaboratory 2030 Strategy, this Digital Environment Supplement aligns with and operationalizes the Data and Digital strategy of the WHO Hub for Pandemic and Epidemic Intelligence (the Hub), particularly its emphasis on federated approaches, ethical governance, and advanced analytical capability. While the Hub-wide strategy defines foundational principles and priorities, this supplement articulates how those principles are implemented within the Collaboratory as a digital public infrastructure for collaborative analytics.

Reference Architecture

Overview

The Collaboratory architecture defines a federated and community-driven digital environment that brings together people, tools, data, and analytical workflows to support collaborative analytics.

The architecture is structured around three key elements:

- **A unified experience layer** that provides access to the ecosystem.
- **A set of capability domains** that enable collaboration, discovery and analysis, including cross-cutting capability domains that capture AI and trust frameworks.
- **Federated data and infrastructure** that support distributed data access and computation.

At the center of the architecture is the Collaboratory Digital Core, which provides shared platform services for discovery, curation, and promotion of analytical resources, as well as enabling key applications across the ecosystem.

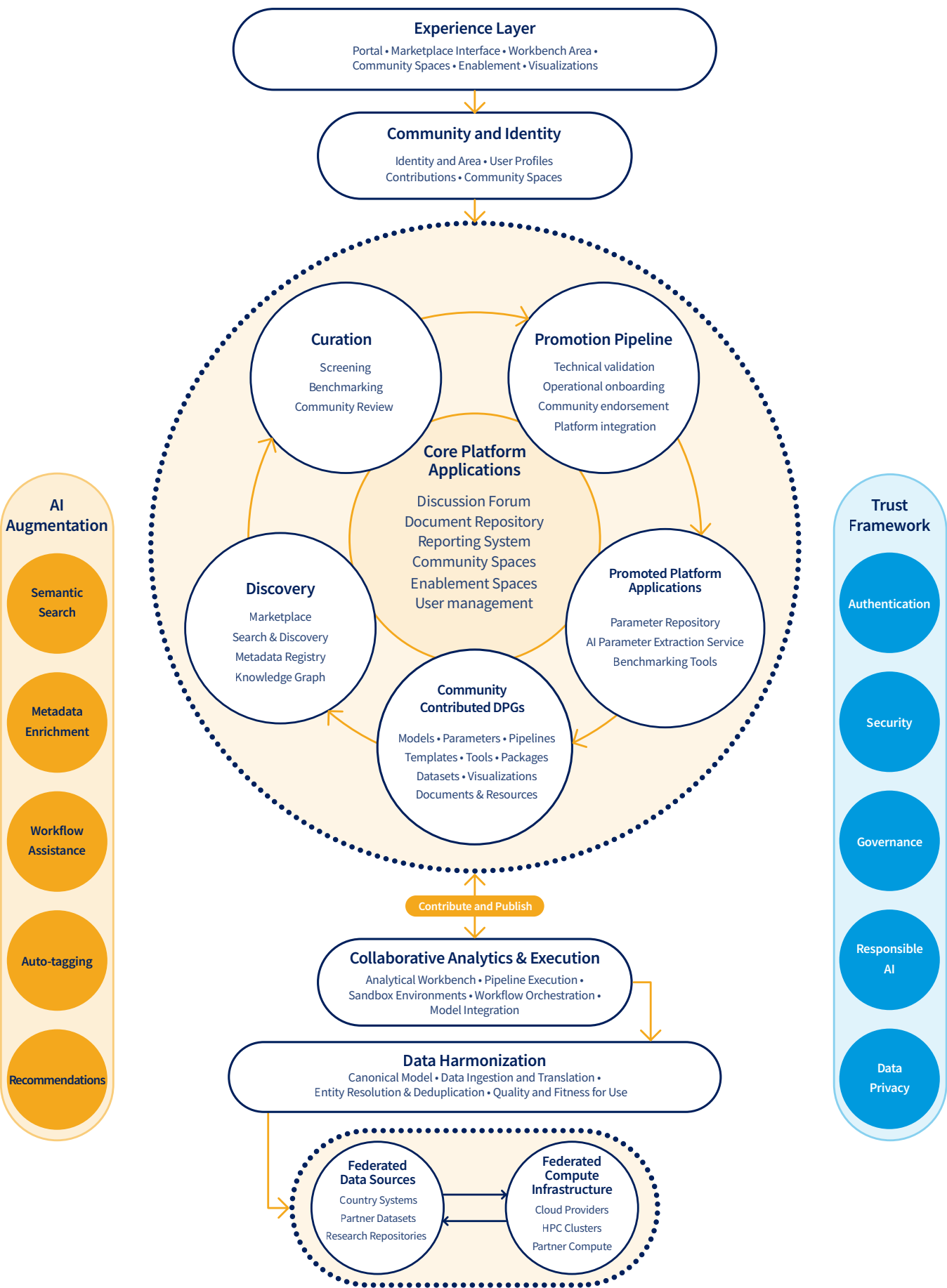
Capability:

Throughout this document, a capability refers to the ability to achieve a specific outcome, enabled through a combination of digital tools, data, processes, and community contributions. The Collaboratory Digital Environment is designed to bring these capabilities together to support its mission:

To connect a global public health analytics community that conducts collaborative analytics – pooling knowledge and experience to develop or curate analytic tools, models and resources for a more collective, rapid and community driven approach to analytics for public health decision making and emergency response.

Figure 1: Collaboratory Goal Architecture with Reference Architecture realized.
Note that this does not represent the current state but rather the goal.

Collaboratory Digital Environment



Experience Layer

The Collaboratory Experience Layer is composed of the external facing components of the Collaboratory and, to most members, will be what defines their experience with the Collaboratory Digital Environment. Its design and functionality enable (or hinder) collaboration and technical value, membership and participation.

The Experience Layer provides the unified interface through which members access and interact with the ecosystem. It brings together capabilities across all domains into a coherent and user-centered environment that supports discovery, collaboration, and analytical work.

The key components of this layer are:

- **Portal:** A central entry point providing access to the Collaboratory, including communities of practice, tools, resources, and updates.
- **Marketplace Interface and Search:** User-facing interfaces that enable intuitive discovery and navigation of analytical tools, models, and resources.
- **Analytical Workbench Access:** Integrated access to analytical environments and workflows, supporting seamless transition from discovery to execution.
- **Community and Learning /Training Spaces:** Interfaces supporting communities of practice, discussions, and access to training and learning resources to enable onboarding and effective use of the Collaboratory and its capabilities by the community.

Capability Domains

The Collaboratory Capability Domains define the capabilities required to operate and evolve the Collaboratory ecosystem. They describe what the architecture must be able to support, independent of specific technologies, platforms, or implementations.

These domains provide a structured view of the digital environment's functional scope, guiding the design and integration of components across the ecosystem.

Community and Identity

The Collaboratory is a community-driven ecosystem connecting researchers, modeling networks, country analysts, and partners. This domain provides the capabilities needed to enable visible, and scalable collaboration across this diverse network.

Key Capabilities

- **Identity and Access:** Secure user authentication and access control, supporting participation across institutions, ensuring a trusted and secure environment, and enabling participation across a variety of components.
- **User Profiles and Affiliations:** Structured profiles capturing roles, expertise, and institutional connections within the global analytics community, ensuring consistency and clarity of members.
- **Connection, Contribution and Attribution:** Connecting members to each other and enabling insights into and connections across membership, linking members to models, pipelines, datasets, and analytical outputs to ensure visibility, recognition, and reuse of contributions.
- **Community Spaces and Knowledge Exchange:** Communities of practice, discussion forums, and collaboration spaces that enable coordination, peer learning, and joint analytical work.

Digital Public Goods and Knowledge Ecosystem

The Collaboratory enables the discovery, sharing, and reuse of analytical resources. This domain provides a structured environment for accessing models, pipelines, tools, datasets, and analytical outputs developed by the global analytics community.

It supports the creation of technical value by making resources discoverable, comparable, and reusable across contexts, reducing duplication and accelerating analytical impact.

At the center of this domain is the Collaboratory Digital Core, which provides the services required to organize, evaluate, and scale community-contributed resources, as well as the operation of key ecosystem applications.

Key Capabilities

- **Discovery (Marketplace and Search):** Centralized marketplace and search capability that enables members to discover analytical tools, models, pipelines, and resources across the ecosystem.
- **Community-Driven Curation and Review:** Screening, benchmarking and feedback mechanisms that support the identification of trusted and high-quality resources via community participation.
- **Promotion Pipeline:** A pathway for external and community-developed tools that show great value to the work and mission to be integrated into Collaboratory architecture and core services, enabling wider adoption
- **Platform Applications and Services:** The Collaboratory Digital Core hosts and enables platform applications and services, including implementation services for user management, community spaces and discussion forum.

Analytics and Execution

The Collaboratory enables collaborative analytics by connecting members to the environments, tools, and workflows used to develop, execute, and refine analytical approaches. This domain supports the joint creation and use of models, pipelines, and analytical outputs across the global analytics community. Consistent with the federated model, execution takes place across distributed, often partner-hosted environments rather than on a single centrally operated platform.

Key Capabilities

- **Analytical Workbench:** Access to shared environments for collaborative modeling and analysis – hosted locally, by partners or within the Collaboratory where appropriate, enabling members to configure, run, and adapt analytical workflows.
- **Pipelines and Workflow Execution:** Shared patterns and interfaces for developing and executing analytical pipelines that link data, models, and outputs in a reproducible and interoperable way.
- **Integration of Tools and Models:** Mechanisms to integrate community-developed tools, models, and methods into analytical workflows, enabling members to build on and reuse work.
- **Reusable Outputs and Templates:** Generation, visualization and sharing of analytical outputs, templates, and workflows that can be reused and adapted across contexts.

Data Harmonization

Collaboratory supports the harmonization of disparate public health data across a federated ecosystem of diverse and distributed systems, serving as the interoperability backbone that connects existing human health, animal health, environmental monitoring, laboratory, and research systems rather than replacing them. Because data originates in systems built on different structures, standards, and reporting logic, consistent analysis depends on shared harmonization capabilities. This domain establishes a canonical model and a translation layer which together provide the ingestion, translation, deduplication, and quality capabilities required to make distributed data analysis-ready without assuming central custodianship of source datasets.

Key Capabilities

- **Canonical Model:** A shared, standards-aligned representation into which case-based and aggregate epidemiological data from diverse systems is translated, enabling consistent analysis across the surveillance ecosystem.
- **Data Ingestion and Translation Layer:** Mechanisms to access and integrate data from diverse sources across and beyond the One Health continuum — including human health, animal health, and environmental monitoring systems, alongside research datasets and partner platforms — rendering it into a consistent, analysis-ready form across the ecosystem.
- **Entity Resolution and Deduplication:** Capabilities to recognize when data from different systems refers to the same case, person, or entity, enabling “collect once, use many times” thus avoiding duplication and facilitating enrichment across the different sources.
- **Quality and Fitness for Use:** Automated and community-contributed validation, standardization, and quality checks that assess the reliability, completeness, and consistency of public health data as it moves through the translation layer, supporting informed, context-aware use of data across human, animal, and environmental health sources.

Cross Cutting Capabilities

While all the above capabilities interact in many important ways to create the full Collaboratory Digital Environment, there are a few additional capabilities that apply across all and will be integrated throughout all the areas of work and development.

Key Capabilities

- **AI Augmentation:** AI capabilities are integrated across Collaboratory to enhance discovery, analysis, and usability. These include support for semantic search, metadata enrichment, knowledge graph inference, human in the loop by design and assistance in developing and executing analytical workflows. AI is applied as an augmenting capability, supporting members and workflows rather than replacing domain expertise, and is integrated in a modular and responsible manner.
- **Trust Framework:** From a digital environment perspective, the Trust Framework provides the foundations required for secure, responsible, and transparent use of the Collaboratory. It includes identity and access control, data governance, security, and responsible use of AI and analytical outputs. These capabilities ensure that collaboration across institutions and regions is supported by clear rules, safeguards, and accountability mechanisms.

Architecture Principles

The Collaboratory architecture is defined by a set of key design patterns that guide how capabilities are structured, deployed, and evolved across the ecosystem.

Federated Architecture Model

Collaboratory is designed as a federated system, where data, compute, and analytical capabilities are distributed across participating institutions, countries, and partners.

Rather than centralizing data and infrastructure, the architecture enables collaboration across decentralized resources, allowing data to remain within local systems while supporting shared analysis and insight generation.

Federation applies across multiple layers of the architecture:

- **Data remains** within country systems, research institutions, and partner platforms
- **Compute and analytical environments** may be deployed locally, within partner infrastructure or within the Collaboratory ecosystem
- **Tools and pipelines** are contributed by the community and integrated through shared interfaces

This approach supports:

- **Data sovereignty** and governance requirements
- **Flexibility** in deployment across diverse contexts
- **Scalability** through distributed infrastructure and contributions

At the same time, the Collaboratory Digital Core provides mechanisms to connect these distributed components, enabling discovery, interoperability, and reuse across the ecosystem.

Layered Architecture

The architecture follows a layered model, spanning from infrastructure and data through to analytical execution and shared digital public goods, and ultimately to community and identity. Each layer builds on the capabilities of the layer below, enabling a clear separation of concerns while supporting interoperability across the system.

Modular Deployment Patterns

The architecture supports modular deployment, allowing components to be developed, deployed, and scaled independently. This enables flexibility in implementation, including the use of open-source tools, partner-developed solutions, and locally deployed components in different contexts.

Integration of Inputs and Outputs

A key feature of the architecture is the ability to integrate diverse inputs and outputs across the ecosystem. Data, models, and analytical outputs can be connected through shared interfaces and workflows, enabling interoperability between community-contributed resources and analytical pipelines.

Promotion of Digital Public Goods

The architecture enables the selective promotion of digital public goods into shared execution capabilities within Collaboratory. Through processes of curation, validation, and reuse, widely adopted tools and approaches can be integrated into the analytical environment, supporting standardization and scalability across the ecosystem.

Conclusion

This Digital Environment Supplement outlines a vision for Collaboratory as an interconnected, federated digital ecosystem—one in which people, data, tools, and analytical workflows can be progressively connected to support collaborative pandemic and epidemic intelligence.

The architecture is designed as a layered and modular system, where capabilities can be developed, adopted, and integrated over time. Individual components and investments can deliver value independently, while their combination strengthens interoperability, reuse, and collective impact across the ecosystem.

Realizing this vision does not depend on a single implementation path. Instead, Collaboratory will evolve incrementally, guided by community needs, partner contributions, and targeted investments across specific capability areas. This approach enables flexibility in how the ecosystem grows, allowing different actors to contribute to and benefit from Collaboratory at different levels of engagement and maturity.

Achieving this next phase of Collaboratory's evolution will depend on active participation from the global analytics community and clear governance to guide growth, trust, and accountability. Community members are not only users of the digital environment, but co-creators of its value—contributing analytical assets, shaping shared practices, and advancing reusable and scalable approaches to analytics. At the same time, governance structures must evolve alongside the ecosystem to ensure responsible use, equitable participation, and alignment with public health priorities.

Together, this positions Collaboratory not as a monolithic platform, but as a scalable and adaptable digital public infrastructure—one that can grow over time, respond to emerging needs, and support collective action for global public health decision-making.

Find out more and join the Collaboratory at: www.who.int/initiatives/collaboratory

Contact us at: collaboratory@who.int



A Laboratory for Pandemic and
Epidemic Intelligence

www.who.int/initiatives/collaboratory

collaboratory@who.int