

IOWN Data Hub to Become a Reality

*Kunihiko Minamihata, Tomohiro Inoue,
and Kazuyuki Takaya*

Abstract

The IOWN Global Forum is working on the social implementation of the Innovative Optical and Wireless Network (IOWN) technology. This article introduces the progress of the storage service study in the IOWN Global Forum, focusing on the newly published IOWN Data Hub Functional Architecture version 2.0 and the Proof of Concept Reference and NTT laboratories' activities related to the IOWN Data Hub.

Keywords: storage, database, virtual data lake

1. Background of the need for IOWN Data Hub

Integrating the real and virtual worlds and feeding the results of simulations in the virtual world back to the real world in real time can provide many benefits, such as the ability to detect and quickly respond to hazards and maintain a comfortable environment with minimal changes. However, to achieve this integration, it is necessary to synchronize in real time the data and behavior of a vast number of objects sensed in the real world with those in cyberspace. Due to their high latency, current networks cannot synchronize vast amounts of data over a wide area in real time, making it difficult to construct a high-definition digital twin that covers a vast space. To solve this problem, we need the IOWN Data Hub (IDH), a new data-sharing platform built on the Open All-Photonic Network (APN) and data-centric infrastructure (DCI), which are both infrastructures of the Innovative Optical and Wireless Network (IOWN). This article focuses on the update of the IDH architecture document and proof of concept (PoC) activities undertaken to enable the IDH.

2. IDH activities at the IOWN Global Forum

In January 2022, IOWN Global Forum members, including NTT, Oracle, and NEC, collaborated to develop and release IOWN Data Hub Functional Architecture version 1.0. For the basic mechanism

and concept of the IDH, refer to “Study of Storage Services at IOWN Global Forum” [1] that can be found in the NTT Technical Review. After the release of the above architecture document, the IOWN Global Forum held a series of discussions toward social implementation, and in July 2023, updated the document to release the IOWN Data Hub Functional Architecture version 2.0 [2]. The main updates are as follows:

- (1) Specific description of use-case requirements: Identification of data types, bandwidth requirements, etc. for each use case presented in the Cyber-Physical System (CPS) and AI (artificial intelligence) Integrated Communications (AIC) use-case documents [3, 4].
- (2) Identification of inherent gaps: Gap analysis based on more specific implementation models.

Some of the requirements for the use cases described in update (1) are listed in **Table 1**. These future use cases will need to meet diverse and demanding data processing and sharing requirements. For example, to meet real-time requirements, it is necessary to process data as-is without any conversion, close to the original data. To handle a large number of data flows, it is necessary to significantly improve distributed data management engines so that they can operate over ultra-high-speed, high-quality networks. Robust data security and data-usage control mechanisms will be required to enable fast and

Table 1. IOWN Global Forum use cases and high-level data requirements.

Use case	Expected requirements			
	Data type	Data producers and consumers	Data collection bandwidth	Producer-consumer latency
Area management security	<ul style="list-style-type: none"> • Video cameras • LiDAR sensors 	Data Producers <ul style="list-style-type: none"> • Video cameras: 1,000 per monitored area (e.g., building) x 100–10,000 monitored areas Continuously sending data • LiDAR sensors: 1,000 per monitored area (e.g., building) x 100–10,000 monitored areas Continuously sending data Data Consumers <ul style="list-style-type: none"> • Security officers: Up to 1,000 per monitored area Receiving alerts when detected • Security monitors: Up to 50 per monitored area Displaying the selected cameras 	<ul style="list-style-type: none"> • 45–60 Mbit/s per camera (assuming Full HD motion JPEG) • 600–800 cameras i.e., 27–48 Gbit/s per building • Thousands or millions of buildings (depending upon the required control) 	From data collection through analysis to alerting: 1 sec, ideally 100 ms
Industry management Remote factory control	<ul style="list-style-type: none"> • Robot/drone's high-definition (8K) cameras • Robot/drone's sound/odor sensors 	Data Producers <ul style="list-style-type: none"> • Machines and/or robots Up to 10,000 per factory Continuously sending data when they are in operation Data Consumers <ul style="list-style-type: none"> • Machines and/or robots Up to 10,000 per factory Receiving control signals when required • Factory Engineers Up to thousands Receiving alert signals and other information when required 	<ul style="list-style-type: none"> • 2.35 Gbit/s per robot/drone 	From data collection through analysis to feedback controls: 10 ms, sub-second, or a second (depending upon the use case scenario assumed)
Smart-grid management Renewable-energy optimization	<ul style="list-style-type: none"> • Sensor data from prosumer • External data sources (weather data, human flow data, etc.) 	Data Producers <ul style="list-style-type: none"> • Prosumer devices (e.g., smart meters, electric vehicles, etc.) and power grid system equipment (e.g., switchgear, transformer, etc.) More than 100,000 Continuously sending data Data Consumers <ul style="list-style-type: none"> • Prosumer devices and power-grid-system equipment More than 100,000 Receiving control signals when required 	<ul style="list-style-type: none"> • 24 Tbit/s (received from data producer) • 128 Pbit/s (data transmission from electric vehicles in 2030) • 3.2 Gbit/s (other data sources) 	From data collection to display: 50 ms From data collection to analysis and feedback control: 200 ms

HD: high definition

JPEG: Image compression formats created by the Joint Photographic Experts Group.

LiDAR: light detection and ranging

reliable data exchange between multiple parties.

Update (2) identified gaps in the computing environment as well as problems in datacenters (DCs) and models for data sharing and processing models that make up the computing environment.

2.1 Fundamental problems in DCs

The current networks and DC technology have several fundamental problems in the following areas. These must be solved if we are to build IDH services that meet the use-case requirements (**Fig. 1**).

(1) Inter-DC network quality

The network quality between DCs used in today's cloud is focused on current use cases. Since packet reordering and loss occur in networks, the Transmission Control Protocol (TCP) is used more frequently than other protocols, such as the User Datagram Protocol (UDP), to ensure reliability. These networks do not provide sufficient performance for the workloads of geographically distributed databases and storage systems that are required by the use cases in the IOWN era.

(2) Intra-DC network quality

The quality of the networks used by hyperscale

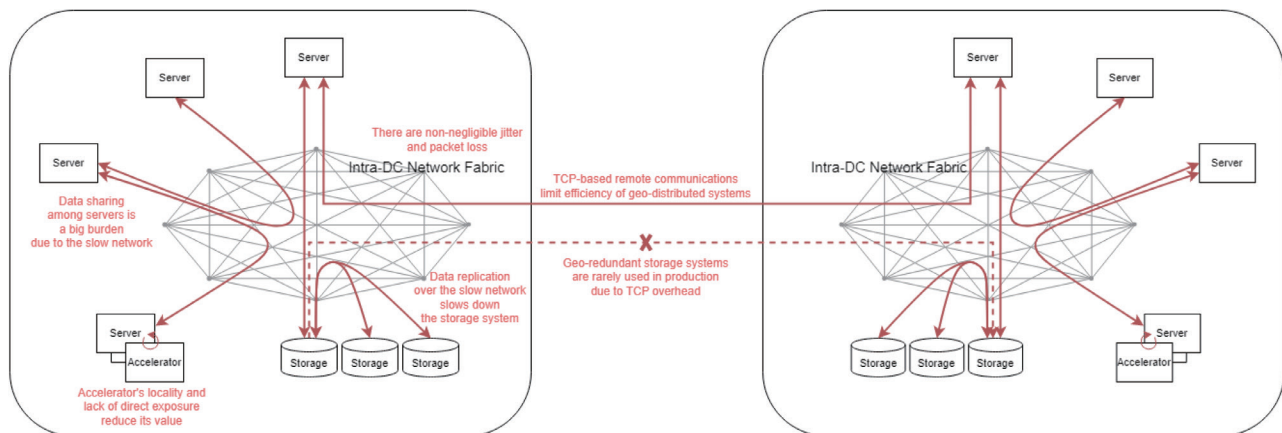


Fig. 1. Origins of gaps in today's technology landscape.

DCs that make up a cloud are not high enough. For example, one-way latency can exceed 1 ms, a latency sufficiently high to cause packet loss to exceed 1% during severe network congestion. The bandwidth is similarly unstable. This is essentially due to oversubscription, a scheme that allows reservation of bandwidth beyond the capacity of the network devices. Therefore, the network that connects computing resources in a DC cannot guarantee all the bandwidth allocated to individual resources.

(3) Storage performance

To ensure data persistence, data are typically synchronized across multiple servers. However, synchronization requires a large workload, rendering performance very unstable. Thus, storage systems that are connected via today's cloud-based storage area networks tend to be slower than directly connected storage devices. For example, the response time of storage systems can exceed dozens of milliseconds, making such systems unsuitable for typical database workloads or redundant storage systems.

(4) Data-sharing performance

In cases where data are processed cooperatively across multiple servers, a certain amount of data exchange is required between servers, for example, to redistribute data. This exchange significantly reduces overall performance. To speed up such data servers, some implementations connect servers using remote direct memory access (RDMA) fabric connections. However, since the use of RDMA in production environments imposes strict requirements, such as no packet loss and no packet reordering, RDMA is used only for very limited purposes in today's cloud environments.

(5) Accelerator usage

Accelerators, such as graphics processing units (GPUs) and field programmable gate arrays (FPGAs), have the potential to achieve high performance, cost-effectiveness, and energy efficiency at least 10 times greater than software-based processing on general-purpose central processing units (CPUs). However, such an accelerator must be deployed within each server and cannot directly access external clients for the purpose of streamlining data loading. Therefore, with current network and DC technologies, it is not possible for multiple general-purpose computing resources to share accelerator resources at high speed and with low overhead.

2.2 Today's implementation model and inherent gaps

Considering the aforementioned issues, we must accept the following constraints if we are to use today's technology to provide data services such as the IDH.

(1) Completion of data processing within one place

If processing efficiency could drop sharply due to poor quality of the network that connects DCs, it would be difficult to implement geographically distributed data services. Therefore, current data-processing services are designed to be implemented within a single DC or within nearby availability zones. In fact, today's clouds are an aggregation of such services. The size of cloud DCs has become so enormous that the number of available cloud regions is very limited, 30 or so at most worldwide. This means that, regardless of where data were generated, all data must be transferred to one of the cloud DCs

and all data processing must be completed there, making the system's energy use highly inefficient.

- (2) Asynchronous data replication as a foundation to build a scalable transactional distributed relational database

In building a scalable transactional relational database (RDB) system, it is a common practice to place replicated read-only data near each RDB server. Such data replication tends to be asynchronous because the network is slow. To reduce the load imposed on slow networks by data replication, often only the transaction or change logs, rather than the data, are propagated. This improves performance but requires reconstruction of table data from the log data on each server, which inevitably increases the system cost and power consumption.

- (3) Sharded data processing to build a scalable key-value store, message broker, and/or analytical distributed RDB

Sharding is a technology for reducing response time and improving scalability by grouping related data together and distributing it across multiple databases. It is used to build scalable systems, such as a key-value store (KVS). However, it does not eliminate data transfer needed for data relocation and server-to-server communication. Rather, the volume of such data transfer and communication may increase in advanced digital transformation (DX) services that need to handle data from many different angles. Since sharded systems require a certain amount of data transfer between scale-out and scale-in operations, data services in the cloud today experience constraints on dynamically changing scalability.

- (4) Usage of a cache layer

The constraints imposed by low-speed networks and storage systems often make it necessary to manage data on a process-by-process basis and rely heavily on asynchronous distributed processing. To streamline such implementations, various cached data management services are provided in today's cloud. However, such services increase end-to-end latency and cause the same data to be copied unnecessarily multiple times in the cloud, which is undesirable in terms of cost and energy consumption.

- (5) Accelerator in use only in a laboratory

One possible approach to solving the aforementioned problems is to use accelerators, such as GPUs and FPGAs, to speed up queries and other data processing in distributed RDBs. Software has been developed and validated for such purposes, but in actual practice, the use of such software rarely improves and often diminishes performance and cost-

effectiveness. This is mainly due to the need to pre-load data at the startup of the accelerator.

The IDH must be designed to address the fundamental and inherent challenges of today's cloud system and must be further developed to adapt to use cases of the wide-area and real-time CPSs that will be increasingly demanded in the future. To this end, it is necessary to verify the IDH iteratively in actual use cases and reflect the results to refine the required functions and design methodology of the IDH.

3. PoC activities on the IDH

In October 2022, the IOWN Global Forum formulated the IOWN Data Hub PoC Reference [5] to help implement the IDH. The use of the Open APN will result in a quantum leap in the speed of hybrid and multi-cloud connections, which in turn will cause a shift from concentration in hyperscalers to the use of enterprise DCs and edge clouds. Thus, it will be necessary to innovate data sharing architecture so that it can accommodate this shift. The IDH disaggregates the functions and processing that have previously been executed by hyperscalers in the central cloud and relocates them to edge clouds. To validate and deploy the IDH architecture, in which connections are set up by the Open APN and DCI, in many use cases, the IOWN Global Forum has released this PoC Reference and is seeking participation from a wide range of organizations, not limited to IOWN Global Forum members.

3.1 The IDH PoC Reference

This PoC Reference has been formulated to validate the IDH architecture that meets the stringent requirements of various DX services. Smart factory, smart grid, and metaverse have been selected as use-case examples and five PoC scenarios have been defined.

To build a metaverse service, for example, it is typically necessary to collect information about the participants, represent them as avatars in the virtual space, and allow the avatars to interact with each other. Thus, provision of a high-quality metaverse service requires the following.

- Collect motion data from up to millions of participants and reflect them in the movements of their avatars in cycles of dozens of milliseconds.
- Move avatars in the virtual space, which contains many virtual buildings and other structural elements, and enable them to communicate with each other.

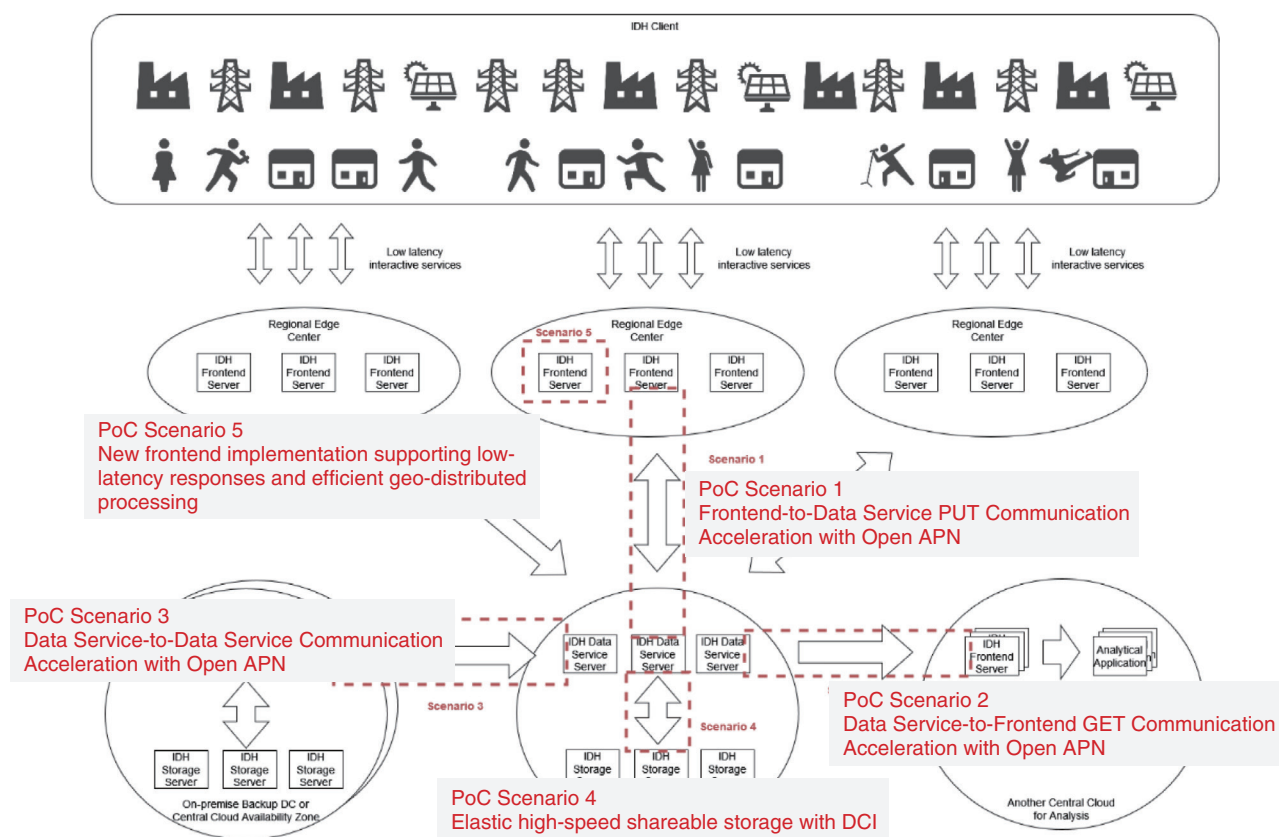


Fig. 2. Geo-distributed IDH adoption model and IDH PoC scenario mapping.

- Render virtual space scenery from each avatar's perspective with motion-to-photon latency of less than dozens of milliseconds.

To demonstrate the IDH architecture for supporting such use cases, the PoC Reference defines a test environment that reflects the geographically distributed deployment of assumed IDH services and identifies the items that PoC projects should validate (**Fig. 2**).

In this model, a subgroup of front-end servers of the IDH architecture is deployed at a regional edge center to provide low-latency services to connected Internet of Things (IoT) devices. The data service servers and storage servers, which are part of the IDH architecture, are deployed at remote DCs or the central cloud to support data persistence and usage.

On the basis of this model, the PoC Reference defines the following five PoC scenarios:

- Scenario 1: Frontend-to-Data Service PUT Communication Acceleration with Open APN
- Scenario 2: Data Service-to-Frontend GET Communication Acceleration with Open APN
- Scenario 3: Data Service-to-Data Service Com-

munication Acceleration with Open APN

- Scenario 4: Elastic High-speed Shareable Storage with DCI
- Scenario 5: New frontend implementation supporting low-latency responses and efficient geo-distributed processing

For each scenario, the PoC Reference describes an overview, required functions, variable conditions, and expected benchmarks to enable the readers to proceed with their PoC projects (**Table 2**).

IOWN Global Forum member companies are currently conducting PoC projects for their respective scenarios on the basis of this PoC Reference. Thus, the IDH is making steady progress toward its social implementation.

4. Activities in the NTT Group

To contribute to the implementation of the IDH, the NTT Software Innovation Center is developing virtual data lake, which is defined as a set of functions in the IDH reference architecture.

Table 2. Benchmark requirements for each scenario.

	Selected features	Variable conditions	Expected benchmark
PoC Scenario 1 Frontend-to-Data Service PUT Communication Acceleration with Open APN	<ul style="list-style-type: none"> • PUT Communications over Open APN • Data preprocessing before data transfer 	<ul style="list-style-type: none"> • Effect of the network type, latency, jitter, bandwidth, packet loss, and packet order change • Network-protocol parameters • Data-object size and presence/absence of data compression 	<ul style="list-style-type: none"> • Throughput per unit system resource • Energy consumption per unit system resource
PoC Scenario 2 Data Service-to-Frontend GET Communication Acceleration with Open APN	<ul style="list-style-type: none"> • Federation function in cloud DC • Get Communications over Open APN • Data filtering and preprocessing in the on-premise DC 	<ul style="list-style-type: none"> • Effect of the network type, latency, jitter, and bandwidth • Network-protocol parameters • Data-object size • Ratio of relevant data 	<ul style="list-style-type: none"> • Time for required data transfer
PoC Scenario 3 Data Service-to-Data Service Communication Acceleration with Open APN	<ul style="list-style-type: none"> • Inter-server communications over Open APN • Continuous data ingestion and data • Pull-access of data • Push-delivery of data • Synchronous data replication across multiple data-service servers of message broker, distributed RDB, and KVS • Scalability 	<ul style="list-style-type: none"> • Effect of the network type, latency, jitter, and bandwidth • Effect of the transaction size • Effect of the data-record size 	<ul style="list-style-type: none"> • Data-hub response time and jitter • Throughput per unit system resource • Energy consumption per unit system resource • Scaling factor
PoC Scenario 4 Elastic High-speed Shareable Storage with DCI	<ul style="list-style-type: none"> • Establishment of communication paths between data-service servers and storage servers, which may go through the gateway server, made by DCI • Usage of standard protocols such as iSCSI • Shared storage • Scalability 	<ul style="list-style-type: none"> • Effect of the network type, latency, jitter, and bandwidth • Effect of the data-block size • Random/sequential access • Ratio of read/write operations • Storage system topology 	<ul style="list-style-type: none"> • Storage-system response time • Throughput per unit system resource • Energy consumption per unit system resource • Scaling factor • Data-hub performance
PoC Scenario 5 New frontend implementation supporting low-latency responses and efficient geo-distributed processing	<ul style="list-style-type: none"> • Data Ingestion through Open APN • Convergence of message broker and other data hub services • In-memory-based data processing • Memory-space sharing across message-broker (buffering) role and other data-hub roles • Continuous data ingestion and subsequent data processing • Resource controls (both test environments) • Scalability 	<ul style="list-style-type: none"> • Effect of the network type, latency, jitter, and bandwidth • Effect of the transaction size • Effect of the data-record size 	<ul style="list-style-type: none"> • End-to-end response time and jitter • Throughput per unit system resource • Energy consumption per unit system resource

There are various stakeholders in both the real world and virtual world, which is built by Digital Twin Computing. To enable the feedback between the two worlds to loop in real time, data must not only be transmitted at high speed but also be securely exchanged between stakeholders who have different standpoints. Therefore, it is essential to enable stakeholders to handle the latest data as if they were at hand, while ensuring permanent governance of the shared data.

The virtual data lake being developed by NTT Software Innovation Center virtually aggregates and centralizes omnipresent data that are managed by

different organizations and geographically dispersed across multiple locations, enabling data users to acquire and use only the needed data efficiently and on-demand. For this purpose, the virtual data lake provides functions that enable data users to search for and discover the data they need from among a vast amount of data based on metadata (semantic and format information of data) and functions that maintain governance by enabling data users to display and use only those items of data they are permitted to use based on the policies set by data providers. These functions enable organizations to mutually use a large volume of widely varied data quickly, easily, securely,

and reliably across organizational boundaries.

To make the IDH platform fit for practical use, we will continue to conduct demonstration tests of the platform by referring to the IDH PoC Reference and other reference materials.

5. Future development

As a “database is a network” platform that supports Digital Twin Computing that runs on the Open APN and DCI, the IDH is expected to be used for mission-critical use cases such as wide-area automated driving. Together with many partners, NTT will verify various IDH implementation models that satisfy the requirements imposed by use cases in the IOWN era

and promote its social implementation.

References

- [1] T. Inoue, “Study of Storage Services at IOWN Global Forum,” NTT Technical Review, Vol. 20, No. 5, pp. 32–36, 2022. <https://doi.org/10.53829/ntr202205fa4>
- [2] IOWN Global Forum, “Data Hub Functional Architecture,” Ver. 2.0, 2023. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Data_Hub_Functional_Architecture-2.0.pdf
- [3] IOWN Global Forum, “Cyber-Physical System Use Case Release-1,” Ver. 1.0, 2021. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-CPS_Use_Case_1.0.pdf
- [4] IOWN Global Forum, “AI-Integrated Communications Use Case Release-1,” Ver. 1.0, 2021. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-AIC_Use_Case_1.0.pdf
- [5] IOWN Global Forum, “IOWN Data Hub PoC Reference,” Ver. 1.0, 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-IDH-PoC-Reference_1.0.pdf



Kunihiko Minamihata

Senior Manager, IOWN Development Office, R&D Planning Department, NTT Corporation.

He received a B.E. and M.E. in information and knowledge engineering from Shizuoka University in 1998 and 2000. He then joined NTT WEST Corporation in 2000 as a system and security engineer for corporate business. He is engaged in spreading IOWN technologies and supporting IOWN Global Forum operations in his current position.



Kazuyuki Takaya

Senior Research Engineer, Supervisor, Social Information Sharing Research Project, NTT Social Informatics Laboratories.

He received an M.E. from Waseda University, Tokyo, and joined NTT in 2000. He is engaged in the research and development of technologies and platforms for data sharing.



Tomohiro Inoue

Director, Data Sharing Infrastructure Project, NTT Software Innovation Center.

He received a B.E. and M.E. from the University of Tokyo in 1999 and 2001. He joined NTT in 2001 and is currently engaged in the research and development of data hubs, trusted data-sharing infrastructure, and global data-space interconnection as the leader of Information Sharing Architecture Group.