Towards the development of a network provisioning platform for data exchange in the health data space

Nikolaos Petroulakis*, Petros Zervoudakis*, George Nomikos*, Alexandros Kornilakis*, Panos Chatziadam*,
Dimitrios Laskaratos†, Vlontzou Maria - Eleftheria†, Zarogianni Eleni†, Vasileios Theodorou†,
*Institute of Computer Science, Foundation for Research and Technology-Hellas (FORTH-ICS), Greece,
‡Intracom SA Telecom Solutions, Greece
{npetro,zervoudak,gnomikos,kornilak,panosc}@ics.forth.gr, {dlaskaratos,mevlon,ezarog,theovas}@intracom-telecom.com

***Abstract*—The deployment of network provisioning mechanisms across data locations in the health domain requires a secure environment to enable reliable, trustworthy, and sovereign data exchange, addressing all legal and regulatory considerations. A data space solution can facilitate secure, interoperable data exchange among diverse stakeholders, fostering collaboration while ensuring data privacy and compliance. This paper aims to tackle the challenge of data exchange in the health domain by outlining the requirements, intended uses, key functionalities, and proposed services for implementing secure network sharing mechanisms within the data space ecosystem. Additionally, it investigates and compares existing minimum viable data spaces to identify the most suitable mechanism for developing a data space platform. The paper presents the development of an architectural framework for a data space-enabled platform designed for network-provisioned data exchange within the health data space. Finally, to validate the proposed data space architecture and services, several realistic scenarios are presented.**

***Index Terms*—data space, network provisioning, data sharing, security and sovereignty, health domain, distributed data systems.**

I. Introduction

In the rapidly evolving digital landscape, data spaces play a crucial role in facilitating secure and efficient data exchange, especially in sensitive sectors like environmental health. The increasing volume of data demands a trusted environment where stakeholders can share, access, and utilize data while ensuring security and regulatory compliance. A data space ecosystem provides this framework, enabling collaboration among providers, research institutions, and technology developers by ensuring interoperability and seamless integration. Data space technology aims to create a trusted environment where organizations within a business network can securely exchange data. Through mechanisms governed by data owners, access rights are declared and verified, allowing data owners to retain full control and adjust usage terms as needed.

The European strategy for data aims to create a single market for data, ensuring Europe's global competitiveness and data sovereignty [1]. The primary goal of the Common European Data Spaces initiative is to establish a unified framework for secure data sharing across sectors such as healthcare, finance, energy, manufacturing, mobility, public administration, and skills, thereby fostering innovation and collaboration while safeguarding data sovereignty and privacy [2]. The European Health Data Space (EHDS), introduced by the European Commission in 2020 [3], seeks to promote health data sharing across the European Union.

Several EU projects and initiatives integrate the concept of health data space, such as Health-X dataLOFT [4], which offers a federated cloud infrastructure aligned with Gaia-X standards [5], and the HealthData@EU Pilot [6], aimed at overcoming existing barriers to the full utilization of digitalized health data. Additionally, the NextGEM project [7] explores reliable, scalable, and secure data sharing between distributed storage locations within the health data space. The NextGEM Innovation and Knowledge Hub (NIKH) platform [8] aims to create a centralized knowledge hub for Electromagnetic Fields (EMF) and Health, providing a standardized approach for European regulatory authorities and the scientific community to store and assess project outcomes. NIKH also supports Findable, Accessible, Interoperable, and Reusable (FAIR) data practices by establishing a secure environment that addresses existing legal and regulatory challenges.

Based on the work carried out, Section II provides a comprehensive overview of the background and related works highlighting the requirements of the data space ecosystem, as well as existing initiatives and data-sharing architectures. Section III describes the NIKH platform architecture and proposes a set of data space services to enhance flexibility and scalability according to different user needs. Section IV includes a comprehensive analysis of existing minimum viable data space solutions to establish a secure and reliable data exchange environment. Section V details the development of the NIKH data space platform as an architectural framework supporting network provisioning between different storage locations, with a focus on implementing a health data space ecosystem through robust access control and techniques for secure and interoperable data sharing. Section VI validates the developed platform in realistic scenarios, where the envisioned data governance mechanism of the NIKH platform leverages data space technology to offer a trusted platform for secure data exchange and data. Finally, Section VII concludes the paper with a summary of outcomes and future directions.

II. Background and related works

Various data lakes architectures have been proposed to facilitate the ingesting, storing, and sharing of heterogeneous data, enabling users to explore and analyze data [9]. The

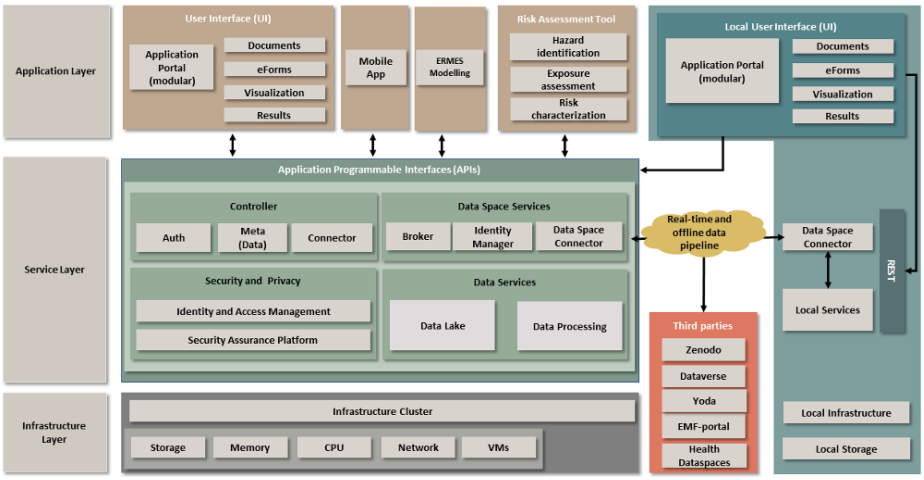


Fig. 1. Conceptual architecture of NIKH platform

authors in [10] introduce a platform designed to address the need for data sharing of heterogeneous scientific data related to the environmental and agricultural research. The platform employs an on-premise solution utilizing the data lake concept to provide cloud services with both institutional authentication and open data access, deploying a catalog service to make data findable by their metadata according to data governance standards. In [11], authors propose the HEALER, a data lake architecture that effectively performs data ingestion, storage, and access with the aim of providing a centralized repository for efficient storage of heterogeneous data in health, enabling the analysis and querying of the data regardless of their native format and type. In addition, authors in [12] propose a blockchain-based solution that can provide a more intensive mechanism for data access control, enabling data owners to specify the access policies, and the schemes achieve fine-grained access control over data. However the above solutions do not ensure a data-sharing schema that achieves security, interoperability and sovereign data exchange while data ownership is kept to the data owners.

To overpass such challenges, the data space appears to be a suitable structured environment where data is stored, exchanged, and accessed across a network of users and platforms in a secure and interoperable manner. The International Data Space (IDS) [13] promotes a virtual environment based on established standards, technologies, and governance models. By offering an architecture for secure data exchange and trusted data sharing, the IDS contributes to designing enterprise architectures for commercial and industrial digitization. Additionally, the European Initiative GAIA-X focuses on developing a federated and secure data infrastructure in for European level [5]. Integration and data exchange across various platforms and networks require connectors that link data providers and consumers with other components. The most widely used and established connectors are the IDS Connector [14] and the Eclipse Data Space Connector [15]. The IDS Connector incorporates the Information Model and employs the Messaging Service for essential IDS functionalities and message handling. In contrast, the Eclipse Data Space Components (EDC) provide an open-source framework for sovereign, inter-organizational data sharing. EDC include an IDS-compliant, modular, and highly extensible framework that supports alternative protocols, offering a distinct separation of the control and data planes, allowing each to operate independently. Finally, an overview of all available data space connectors can be found in [16].

III. Data Space Services in the NIKH Platform

The NIKH platform serves as a central hub for facilitating knowledge, research, active collaboration, and knowledge sharing among all stakeholders involved in order to maintain compliance with safety standards, proceed to risk assessment and increase citizen's awareness on EMF and Health research [8]. Its architecture is organized into three main layers, Infrastructure, Service, and Application, with a clear separation between each (Figure 1). In addition, the architecture includes a Local Premises Layer, which manages infrastructure, storage, and application functionalities within local environments. The Infrastructure Layer provides the foundational hardware and networking capabilities necessary to support the platform's robust and reliable functions. The Service Layer includes all applications and services that enable the platform's core functionalities, such as authentication and access control, data ingestion, storage, and inter-component communication. At the center of this layer is the Controller, which acts as a gateway between the platform and third-party applications utilizing its services. Each request to the platform is routed through the Controller, which orchestrates and manages internal workflows, supporting metadata management, data uploading, and data exchange processes. Finally, the Application Layer is the user-facing layer of the NIKH platform, containing applications and services that users interact with directly. This layer provides intuitive interfaces and tools tailored to user needs

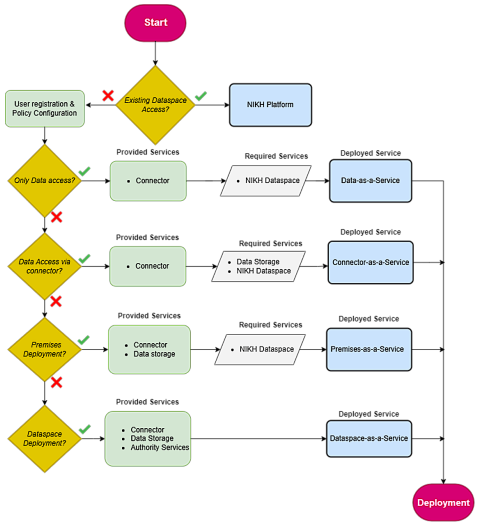


Fig. 2. The relation between all 'as-a-Service' scenarios

and enabling seamless interaction with the underlying system components and the platform's functionalities.

The data space functionality of the NIKH platform is considered one of its key applications. Through Data Space Services and the integration of Data Space standards and protocols, the NIKH platform enables data to remain decentralized, transferring only upon request, and includes mechanisms for premises identity validation, data usage restrictions, and metadata cataloging. In this regard, decoupling the connector and data storage within a data space architecture introduces a versatile approach that provides distinct solutions, referred to as as-a-service (Figure 2). This conceptual separation is designed to offer organizations and stakeholders enhanced flexibility and scalability, allowing them to tailor data management strategies to their specific needs and requirements by introducing four distinct services: *Data-as-a-Service*, *Connector-as-a-Service*, *Premises-as-a-Service* and *Data-Space-as-a-Service*.

Data-as-a-Service (DaaS): DaaS provides end-users and organizations with a comprehensive solution for accessing data resources within a data space ecosystem. A set of web interfaces allows end-users to seamlessly interact with data space functionalities, supporting workflows for efficient data retrieval and metadata management.

Connector-as-a-Service (CaaS): The primary function of CaaS is to seamlessly enable data exchange capabilities among various entities within the data space ecosystem. Using data space connectors, end-users can easily explore and retrieve heterogeneous data across the ecosystem. Consequently, CaaS provides the NIKH platform's connector to facilitate collaboration among users, allowing data sharing without the need to switch between different tools or platforms.

Premises-as-a-Service (PaaS): PaaS provides stakeholders and organizations with virtual premises, including reliable storage connected to a data space ecosystem. This PaaS solution offers participants a range of storage benefits, such as seamless scalability, high availability and security, eliminating the need for their resources. Participants can confidently store, manage, and access their data assets, assured that their data is securely hosted and managed by a trusted service provider.

Data-Space-as-a-Service (DSaaS): DSaaS leverages the advantages of the data space ecosystem by providing the flexibility to create a fully managed environment tailored to specific needs, with a centralized data space authority to maintain complete control over the data space. With DSaaS, the setup process is streamlined and focuses on the capabilities of the deployed data space, driving innovation and collaboration within the newly created data space ecosystem.

IV. Investigation of Minimum Viable Data Spaces

To begin developing data space services on the NIKH platform, it is essential first to investigate existing data space solutions. In this regard, the Minimum Viable Data Space (MVD) approach has been explored as a minimal deployment focusing on the core functionalities needed to establish a functional data exchange environment. This section presents four different minimal configurations as primary deployments for implementing the NIKH data space platform.

A. International Data Space MVD

The IDS MVD consists of essential components that establish a data space with the minimum required features for secure and sovereign data exchange, as outlined in the IDS Reference Architecture [17]. The IDS connector was locally deployed, utilizing the provided Swagger-UI for exploring and testing various resource offering and usage control scenarios. In addition to the locally deployed provider-consumer IDS Deployment, Kubernetes deployment via Helm was also tested. Within the IDS Connector, data are made available as text or JSON files through a URL, allowing consumers to download content directly. Additional components for identity management, including the Dynamic Access Provisioning Service and Certificate Authority, were also evaluated. With the active maintenance of the IDS Connector discontinued, exploration shifted toward the EDC connector to identify functionalities best suited to the needs of NIKH Data Space Services.

B. Single-Connectors MVD

The Samples EDC repository comprises a collection of scenarios that demonstrate the basic functionalities of configuring and running an EDC connector, as well as the steps for executing various data exchange examples [18]. Initial testing was conducted through a centralized local deployment, where efforts were made to extend the Samples' functionalities to support file exchange within the local filesystem. Simultaneously, the two connectors from the examined Samples were deployed on separate host machines to support the distributed model envisioned for the NIKH platform. These connectors were run as containers deployed in Kubernetes, with their addresses exposed to enable communication between them. JSON data was successfully exchanged between the provider and consumer, with records logged on an HTTP logging server. However, limitations in the Samples—such as the

TABLE I
COMPARISON OF CONNECTORS DEPLOYMENT

MVD Name	Core Connector	Deployment	Files supported for exchange	Data Storage	Authority and Additional Services	Federated Catalogue
Single-Connectors MVD	Eclipse Data space Connector	Centralized Non-containerized	REST response with a JSON object	Local file system	No authority services	No
IDSA MVD	IDS Connector	Centralized Containerized	Actual dataset exchange	Local file system	No authority services	No
Sovity MVD	Eclipse Data space Connector	Centralized Containerized	REST response with a JSON object	Local file system	No authority services	No
Eclipse MVD	Eclipse Data space Connector	Centralized Containerized	Actual dataset exchange	Object storage service (Azurite)	Registration Service / Identity Hub / DID Server	Yes

lack of support for actual file exchange (beyond *HttpData*) and identity management—emphasized the need for a more comprehensive connector implementation.

C. Sovity MVD

The Sovity MVD is a demonstration deployment that showcases the capabilities of the EDC connector [19]. The project provides a range of Docker images with certain extensions, including an open-source, development edition with basic features, as well as a user-friendly interface that supports management operations and facilitates interactions with the connectors. Initial testing of the Sovity deployment was conducted in a local environment, where the additional Policies and user interface functionalities were evaluated. Additionally, the Sovity connectors were deployed on separated host machines in a distributed setup, with addresses exposed to enable communication between them, supporting a distributed data space configuration. However, limitations such as the lack of actual data exchange functionality and the absence of an Identity Management mechanism highlighted the need for a more comprehensive implementation.

D. Eclipse MVD

The Eclipse MVD is a demonstration deployment showcasing the capabilities of the EDC connector, providing a practical example of how abstract data space concepts can be implemented in real-world scenarios and demonstrating features essential for supporting a data space [20]. Built on the EDC connector, it includes object storage capabilities via Azurite [21] local blob storage. A key component of the Eclipse MVD is the Data Space Authority Services, which comprises a registration service for managing participants within the data space, a federated catalog listing all participants' contract offers, and a DID server for decentralized digital identity representation. To evaluate the Eclipse MVD, transfer capabilities between participants in a decentralized data space were tested, where the Contract Offer feature allowed consumers to negotiate asset transfers before the data transfer. The evaluation confirmed that the Eclipse MVD's core functionalities, such as i) the discovery and access assets through the federated catalog, ii) the contact negotiation, iii) the secure data transfer between participants and iv) the provision of authority services for registration and identity management, could be used for the NIKH data space platform

E. Comparison of connectors deployments

Based on the exploration of various deployment implementations, a comparison of the existing features and functionalities offered by each deployment has been conducted. This comparison considers different configuration setups tested, the types of data the provider can offer, supported methods for file exchange, the presence or absence of identity management functionalities, and available policies. For the NIKH platform and its users, the distributed paradigm is essential to enable the offering and exchange of actual data files, likely through object storage. In terms of identity management, support for these functionalities is critical for the authentication and authorization of connectors participating in the data space. A detailed comparison of different data space solutions, including the support or absence of these features in each MVD implementation, is presented in Table I.

V. DEVELOPMENT OF NIKH DATA SPACE PLATFORM

The NIKH data space platform establishes a comprehensive framework that enables secure and sovereign information exchange between participants through data ownership and distribution policies. To this end, various standards and protocols have been implemented to support the integration of diverse services necessary for creating a unified platform. Specifically, the proposed data space architecture includes configurations and deployments for secure data exchange within the health data space. To establish a functional data space ecosystem, the NIKH platform incorporates functionalities such as data space connectors, data storage, service orchestration and access control, all deployed on a container orchestration system, as presented in Figure 3 and detailed below.

The NIKH platform's data space ecosystem encapsulates its main components in container images, independently deployed as Docker containers. These container images are automatically built within GitLab CI, based on configuration files (Dockerfiles and Docker Compose files) provided by component owners. The containerized NIKH software components are deployed in Kubernetes, using appropriate configuration files (Kubernetes deployment YAML configurations). The Eclipse Data Space Connector, selected based on the investigation in the previous section, is built upon the Eclipse MVD and serves as the most suitable solution for representing organizations and data providers/consumers across various locations. Acting as a middleware layer, it

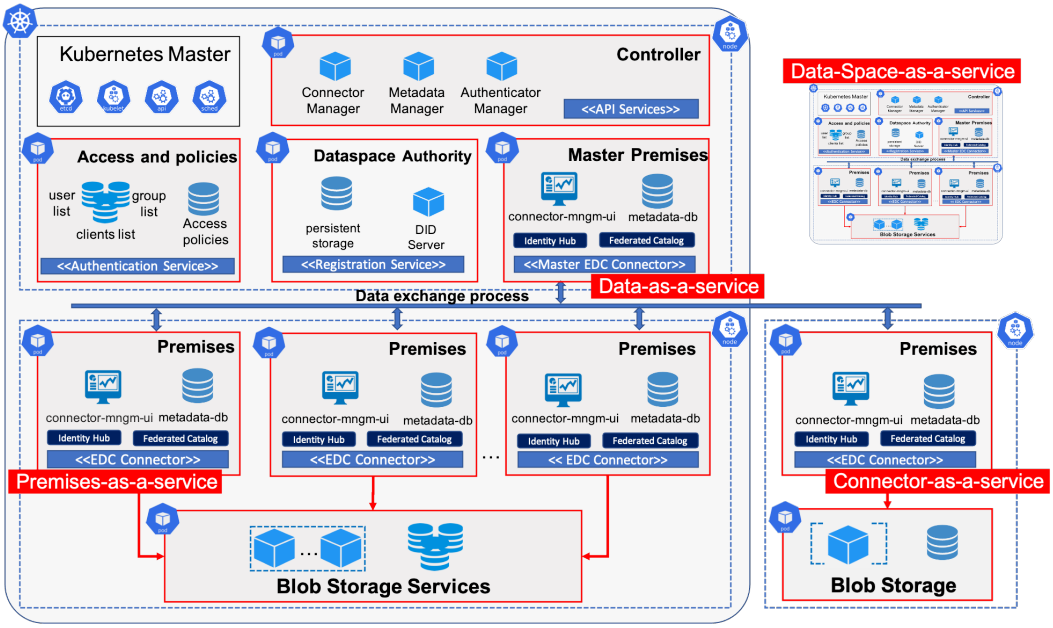


Fig. 3. Deployment of NIKH data space ecosystem

enables seamless integration and interaction between data storage environments and the platform, offering enhanced data management capabilities. Data offerings become visible on the platform automatically once a participant's connector is registered in the data space. The ultimate goal is to keep data with or under the control of the data owner. For the NIKH platform, the Azurite emulator has been utilized to replicate Azure Blob storage service functionalities within local development environments, supporting object storage capabilities through a comprehensive feature set.

To enable efficient access control policies, a harmonized mechanism was deployed for access control using Keycloak [22] and access policies through the EDC. Specifically, the Keycloak open-source tool is used for identity and access management, providing customizable features for user roles, groups, and permissions. It has been deployed as a container on the same Kubernetes cluster, and communication with Keycloak is facilitated through RESTful APIs, primarily using JSON as the data format. To streamline user access, a registration request process has been introduced on the platform, allowing new user accounts to be created within a Keycloak realm, where users can log in to the NIKH platform. This setup enables a centralized authentication and authorization service for secure communication and interaction with NIKH's services. For access policies, the EDC allows users to create metadata records of their data and link them to specific usage policies through contract definitions. Contract offers can be negotiated with users who request access to this data, based on the defined access policies. The available policies are: i) *public*, where actual data is fully accessible with no restrictions; ii) *public after embargo*, where data is publicly available after a specified embargo period; iii) *restricted access*, where data may require negotiation; and iv) *closed*, where both data and metadata are inaccessible to other participants.

The Controller of the NIKH platform is designed to orchestrate all platform interfaces, providing REST API endpoints for the connector, metadata and authentication manager. It integrates the deployed Connector Manager, which enables access to the data space by acting as a wrapper for key functionalities, including file sharing between data space connectors and supporting the creation and updating of policies, contracts and assets by utilizing the APIs provided by the data space connector service. The Controller offers functionalities for authentication and manages the metadata catalog, including the creation, editing and deletion of entries. It also automates the process of new data uploads by making the necessary API calls to the Data Space Connector. All sub-components of the Controller have been containerized and are managed within a Kubernetes container orchestration environment.

VI. Evaluation of NIKH Data Space Platform in Realistic Scenarios

To evaluate the practical application and adaptability of the proposed data space solution, we investigate its effectiveness, storage capabilities, and data governance based on a data-oriented approach on various deployment scenarios. These scenarios emphasize secure, flexible data exchange between stakeholders, addressing the unique data management needs of different organizational environments within the NextGEM consortium. Each scenario highlights the operational requirements, challenges, and essential functionalities necessary for effective deployment within a health data space ecosystem.

A. NextGEM Realistic Scenarios

To validate the applicability of the NIKH data space capabilities within the developed NIKH platform, a variety of

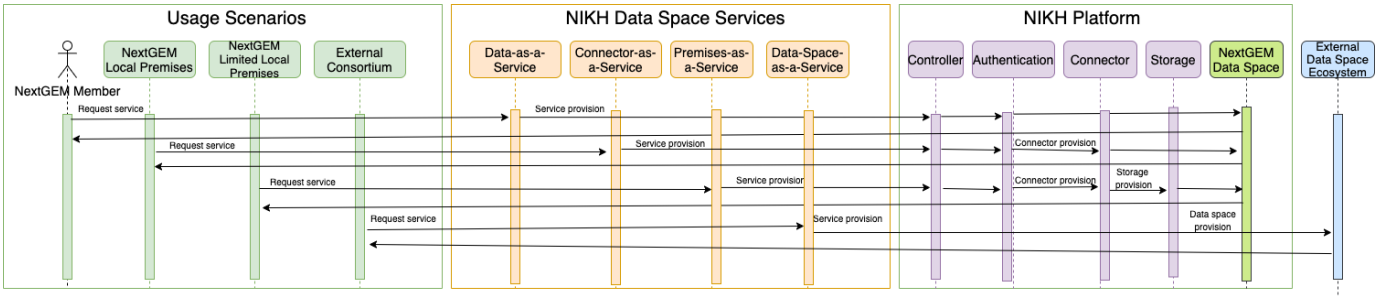


Fig. 4. NIKH Platform data space services in NextGEM realistic usage scenarios

realistic NextGEM usage scenarios are proposed, in which the functionalities of data space services are evaluated (Figure 4).

Usage Scenario 1 - NextGEM Members: The first usage scenario focuses on users within the NextGEM consortium who aim to retrieve data through a secure, permission-based platform. A NextGEM Member is any registered and authorized user on the NIKH platform, with access capabilities varying according to predefined permissions. To support these requirements, the *Data-as-a-Service* (DaaS) model enables seamless data access for users, eliminating the need for on-premises infrastructure or dedicated connectors for each NextGEM Member. Members are registered via the Keycloak access control mechanism to gain access to the available data through the NIKH controller. Access is granted through the centralized data space connector of the NIKH platform, which aggregates requested data from all network participants in compliance with data space protocols. This setup provides streamlined access and supports all operations without requiring additional data storage, while ensuring high availability, scalability, and reliable access for NextGEM Members.

Usage Scenario 2 - NextGEM Local premises: The second usage scenario involves deploying necessary services on both the NIKH platform and the organization's infrastructure to facilitate data exchange between data providers and consumers. The local premises infrastructure ensures the protection of sensitive data, supporting seamless data transfer and sharing among all involved partners. This approach allows stakeholders to maintain secure and protected data within their own infrastructure while fostering efficient collaboration and sharing practices. In alignment with the proposed architecture, the *Connector-as-a-Service (CaaS)* model offers the deployment of a dedicated container running the EDC connector, along with additional containers running authority services, pre-deployed to support all local premises needs. A storage service is also deployed within the local infrastructure to enable persistent storage of scientific data. Consequently, this model demands sufficient storage capabilities on the local premises infrastructure, while data availability, scalability, and reliability depend on both the local premises infrastructure and the NIKH platform. Data isolation further ensures data governance, as organizations retain full control over their data.

Usage Scenario 3 - NextGEM Limited Local premises: The third usage scenario involves entities deployed on the NIKH platform, enabling the exchange of scientific data from a data producer to other organizations through a shared infrastructure. Limited local premises utilize a common pool of resources that allow for the deployment of a dedicated connector and other necessary services for data collection and transfer, including an object storage service for storage needs. This approach ensures that organizations using this model can maintain secure and protected data within the NIKH platform. In alignment with the proposed architecture, the *Premises-as-a-Service (PaaS)* model deploys a dedicated container running the EDC Connector, with essential services—such as authority services and a common object storage service for persistent storage—already deployed on the NIKH platform. Consequently, this model does not require individual storage capabilities, as the NIKH platform ensures high availability, scalability, and reliable data access. However, the centralized storage of data affects data governance, as organizations do not retain full control over their data space.

Usage Scenario 4 - External Data Space Ecosystem: The final usage scenario involves a group of entities or organizations, such as members of a project consortium or initiative, that need to deploy their own data space ecosystem to facilitate data exchange in alignment with data space principles. The *Data-Space-as-a-Service* (DSaaS) model allows for the creation of an external data space ecosystem tailored to the interests of these entities. Consequently, the service provider can replicate the proposed platform (including the Kubernetes cluster, controller, authority services, premises, storage, connectors, etc.) according to the specific needs and capabilities of the stakeholders involved. This approach enables the NIKH data space platform to be re-deployed in a centralized or remote location based on entities' requirements (e.g., number of participants, storage capabilities) to support the establishment of a fully functional data space ecosystem.

B. Discussion

As previously discussed, the NIKH platform provides data-sharing functionalities across various realistic scenarios, leveraging developed data space mechanisms to accommodate organizations with differing levels of infrastructure and needs. The primary goal of this Data Space topology is to enable controlled, sovereign, and secure data exchange and sharing among stakeholders. Within this framework, data sovereignty

TABLE II
VALIDATION OF DATA SPACE SERVICES IN REALISTIC SCENARIOS

Usage Scenario	Service	Storage	Availability	Scalability	Reliability	Data Governance	Isolation
US1 - Members	Data-as-a-service	-	High (NIKH)	High (NIKH)	High (NIKH)	No	-
US2 - Local Premises	Connector-as-a-service	Depends on premises	Depends on premises	No	Depends on premises	Yes	Physical
US3 - Limited Local Premises	Premises-as-a-service	Limited (NIKH)	High (NIKH)	High (NIKH)	High (NIKH)	No	Virtual
US4 - External Data Space	Data-Space-as-a-service	Limited (NIKH)/ Depends on external space	High (NIKH)/ Depends on external space	High (NIKH)/ Depends on external space	High (NIKH)/ Depends on external space	Yes	Virtual/ Physical

and trust are upheld, as each participant can apply usage restrictions to their data and monitor transactions through continuous tracking. Security is further reinforced through identity verification for each participant. Additionally, the framework supports metadata storage and query functionalities, allowing participants to search for relevant data sources and request access to specific datasets. In all scenarios, ecosystem members can access, publish, or acquire datasets under specified rules. Organizations or users wishing to utilize a dataset published within the ecosystem can request access, relying on data governance policies to validate the request under various conditions and locations. The different functionalities offered by NIKH data space services, along with the operational and technical requirements for end-users within these scenarios, are summarized in Table II. Consequently, with standardized implementation procedures, the NIKH data space platform provides relevant stakeholders with an efficient and streamlined way to access the data space, update information, and provision data to involved parties.

VII. CONCLUSION

This paper outlined the network provisioning mechanisms across distributed data locations and reported on the development status of the NIKH platform. It presented the various implementation activities and frameworks utilized during the design phase of the NIKH platform, establishing a secure environment for health data sharing in line with data space mechanisms. Within this scope, an overview of the platform's services was provided, highlighting tailored data management strategies. Additionally, an investigation into the most efficient minimum viable data space mechanisms was conducted, identifying the most suitable one for the platform. Development details of the NIKH data space platform were also provided, along with an evaluation of its applicability under realistic scenarios. Future steps will include refining and expanding NIKH data space functionalities related to network provisioning among NextGEM partners, facilitating the exchange of experimental results across the various NextGEM case studies.

ACKNOWLEDGMENT

This work is part of the European Union's Horizon Europe research and innovation programme under grant agreement No 101057527 (NextGEM). Funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] A European strategy for data, European Commission, https://digital-strategy.ec.europa.eu/en/policies/strategy-data
[2] Common European Data Spaces, European Commission, https://digital-strategy.ec.europa.eu/en/policies/data-spaces
[3] European Health Data Space, https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
[4] HEALTH – X dataLOFT Platform, https://www.health-x.org/en/platform
[5] Gaia-X: A Federated Secure Data Infrastructure, https://gaia-x.eu
[6] HealthData@EU Pilot, https://ehds2pilot.eu/
[7] N. Petroulakis, et al. "NextGEM: Next-generation integrated sensing and analytical system for monitoring and assessing radiofrequency EMF exposure and health." International Journal of Environmental Research and Public Health 20.12 (2023)
[8] N. Petroulakis, et al., "Designing NIKH: the NextGEM Innovation and Knowledge Hub to Access Next Generation Radio Frequency EMF Exposure and Health Data," IEEE CSCN, Thessaloniki, Greece, 2022.
[9] Sawadogo, P. & Darmont, J. On data lake architectures and metadata management. *Journal Of Intelligent Information Systems*. **56**, 97-120 (2020,6), http://dx.doi.org/10.1007/s10844-020-00608-7
[10] Sarramia, D., Claude, A., Ogereau, F., Mezhoud, J. & Mailhot, G. CEBA: A Data Lake for Data Sharing and Environmental Monitoring. *Sensors*. **22** (2022), https://www.mdpi.com/1424-8220/22/7/2733
[11] Manco, C., Dolci, T., Azzalini, F., Barbierato, E., Gribaudo, M. & Tanca, L. HEALER: A Data Lake Architecture for Healthcare. *EDBT/ICDT Workshops*. (2023), https://api.semanticscholar.org/CorpusID:258559134
[12] Wang, S., Zhang, Y. & Zhang, Y. A Blockchain-Based Framework for Data Sharing With Fine-Grained Access Control in Decentralized Storage Systems. *IEEE Access*. **6** pp. 38437-38450 (2018)
[13] International Data Space Association (IDSA), https://internationaldataspaces.org
[14] J. Pampus, J. Brian-Frederik, Q. Ronja. "Evolving data space technologies: lessons learned from an IDS connector reference implementation." ISOLA 2022.
[15] Eclipse Dataspace Components (EDC), https://projects.eclipse.org/projects/technology.edc
[16] Data Connector Report, Report, No 16, September 2024, International Data Spaces Association, https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Data-Connector-Report-84-No-16-September-2024.pdf
[17] Minimum Viable Data Space (MVDS) International Data Spaces Association https://docs.internationaldataspaces.org/knowledge-base/mvds
[18] Eclipse Dataspace Components, Samples, https://github.com/eclipse-edc/Samples
[19] Sovity, Data Sovereignty as a Service, https://github.com/sovity
[20] Eclipse Dataspace Components, Minimum Viable Dataspace https://github.com/eclipse-edc/MinimumViableDataspace
[21] Azurite, Microsoft Azure, https://github.com/Azure/Azurite
[22] Keycloak, Open Source Identity and Access Management, https://www.keycloak.org/