Gaia-X Federated Data Infrastructure: The Future of Data Management

An overview of Gaia-X, a cross-centre initiative that aims to create the next generation of data infrastructure by providing a secure, federated system based on the highest standards of digital sovereignty and innovation.

✓ Key Points

- Gaia-X creates the foundation for a sovereign, federated, open data infrastructure based on European values.

- It connects centralised and decentralised infrastructures into homogenous and user-friendly system.

- It creates an open, transparent, federated catalogue of data sources and data services where stakeholders can access data in a consistent, secure and trustworthy manner.

- Gaia-X provides a solid legal and technical foundation to better manage data across European Member States.

Gaia-X creates the foundation for a sovereign, federated, open data infrastructure based on European values. It is a strategy that can facilitate a collaboration between different stakeholders to build a data space using the framework.

In particular, health data space includes a set of shared capabilities and federated health data spaces, where data is granularly and selectively accessible. These federated data spaces will exist on regional, national and European level. The goal is to ensure the data space contributes to the care delivery processes for patients as well as allows the secondary use of data on a cohort or population scale. This will create a data value chain between data holders and data users across the broad and complicated health domain ecosystem.

Mission and Goals of the Data Space

Gaia-X's mission is to allow access and sharing of data securely and confidently. Our primary goals include:
• Providing the means to link currently isolated data and disparate applications, between citizens, care providers and other stakeholders, within countries and across borders, in a transparent manner, adhering to international interoperability standards.
• Providing a framework to implement the data spaces at scale in a compliant, secure, and trustable manner.
• Enabling the storage and access of personal and non-personal information in trusted and collaborative cloud infrastructures, with elasticity to scale and with a proper legal basis (consent, anonymisation, etc.).
• Implementing clear governance for the use of data on personal, regional, national and European level and for delivery of care, for research, for commercial and governmental use.

Key Challenges

Healthcare is facing many challenges including the increasing burden of chronic disease, an ageing population, significant lag in digitisation and a lack of integrated and longitudinal views when managing a single patient or a patient population. The health ecosystem is fairly complex with many stakeholders, regulated processes and multiple sources of funding. This has slowed down the innovation and transformation process in healthcare. The COVID-19 pandemic has further exposed these issues and has clearly demonstrated the fragility of global healthcare systems.

When it comes to patient data, it is collected and stored by disparate systems with low interoperability. Clinicians have difficulty accessing this data as well as share/exchange it. Therefore, they are unable to benefit from it. Data is distributed between many stakeholders, including healthcare providers, insurance companies, companies in the secondary health market and individual patients. For healthcare to progress and innovate, it is important to ensure data from various sources is easily available while meeting the highest standards of privacy, security, transparency and control.

Gaia-X data space should contribute to the care delivery processes (primary use of data) for the individual patient or resident (for healthy living, prevention, diagnosis, treatment and home care, leading towards value-based healthcare) as well as to secondary use of data on a cohort or population scale (for research, innovation, crisis preparedness and crisis management, public and population health). At present, there is significant duplication of effort in the management of healthcare data. Achieving interoperability among these systems is critical for clinical care as well as clinical research. This can also facilitate virtual healthcare services (e.g., e-consultations, e-interventions, telehealth, teleradiology, remote care management, and other aspects of telemedicine), as well as promote digital health science using both trial- and real-world-data.

Data Space – A Holistic Solution

The health domain comprises four essential contributors: data holders & data users, application and service providers, data space governance and operating entities, and cloud service providers. All these stakeholders can benefit from data-driven applications and services built on a trusted, safe and secure cloud infrastructure. Gaia-X data space also enables an equal playing field for application and service developers. Small, medium and large companies, together with governmental and non-governmental institutes, can take the lead in developing agile and innovative solutions that can transform healthcare.

The future landscape for innovative applications is incredibly rich and may include lifestyle improvement using various tracking devices and (non-)human coaching, symptom and vital sign tracking, continuous outcomes measurement, real-world-data trending and analysis, decision and referral guides, appointment scheduling, clinical data trial management, advanced image and genomic data analysis, health resource and capacity optimisation, emergency and triage management, and many more.

Data applications and services require common standards and common components. This will require participation from both data governance and operating entities. The data space initiative will co-create the framework with Member States and the European Commission. The data space will also work in collaboration with operating entities as they are the connectors that link isolated data in a transparent, secure and audited manner.

Finally, the data space is created on a foundation of cloud services and edge components. Gaia-X will create an ecosystem of trusted, safe and secure cloud and edge infrastructures in Europe to allow health data to flow securely and in line with Europe's privacy provisions and values.

Gaia-X Components

Core Gaia-X infrastructural components include:
• Identity & Trust: federated identity management for individuals and organisations

• Federated Catalogue: to publish the registration, consent and query services
• Sovereign Data Exchange: to manage registration, consent, cloud/edge services and data query and access services
• Compliance: rights management, onboarding and certification

Benefits of Data Space

The data space offers the following benefits to healthcare:
• It provides the required technical infrastructure that enables secure storage of large data sets and powerful computing architecture and methods for the complex analysis of data at petabyte scale.
• It ensures data access is secure and GDPR compliant, with datasets, analytics pipelines, and compute provided by different actors in the health care and research sectors.
• It provides extensive and published datasets through secured interfaces.
• It offers automated deployment and roll-out capability of the interoperable and virtualised methods of other secure top locations in the healthcare sector.
• With its ability to integrate data across individual domains (e.g., image data, clinical information), the Gaia-X network offers the potential to realise more complex integrative analyses within personalised medicine for the benefit of patients.
• It enables sharing of documents between different departments, private physicians, clinics, long-term care, acute care with different clinical IT systems.
• It provides a unified approach to support patient rights, clinical care and research.
• It enables secure, and GDPR-compliant access to patient data through compliance with Gaia-X provided policies.

Recommendations

The following are recommendations for the healthcare industry:
• Converge the ongoing or projected initiatives towards the common Gaia-X framework.
• Improve transparency of the existing initiatives and deploy a process and tooling to maintain this transparency.
• Build the federated data and services catalogues to improve the availability of and access to data.
• Commit to privacy and security and to the principles of data sovereignty by design.
   Recommendations for policymakers and society include:
• Align the EU initiatives with Gaia-X: European Health Data Space development and roadmap for cross-border data exchange, the TEHDAS Joint Action, the European Alliance for industrial data, cloud and edge, and the EU Cloud Rulebook with the Gaia-X initiative.
• Obtain buy-in and support from the Health Ministries of the European Member States with the Gaia-X framework and architecture.
• Solve electronic identity management across Europe:

continue and accelerate the deployment of federated eID solutions (eIDAS) for natural and legal persons.
• Agree on data trust and data governance: where/who defines the standards and the guidelines for ethical use of personal health data. Harmonise these principles and guidelines across the EU Member States.
• Provide adequate level of funding, while keeping limited local variations where required (national healthcare sovereignty) to stimulate adoption of the Gaia-X framework and to kick start the deployment of a Gaia-X compliant data space.
• Encourage the European Data Protection Board to provide clear and updated guidelines on the concepts of personal data and non-personal data and on anonymisation techniques.
• Reduce fragmentation of local conditions on data processing for scientific research purposes. Reduce fragmentation of local data protection/healthcare rules applicable to health data, in particular in the field of cross-border transfers of health data within the EU.

Future Outlook

It is evident that the healthcare sector is multi-faceted, with a complex interplay between patients, care providers, insurers, governments and various other stakeholders. There is an increasing demand for better care delivery processes, efficient use of data and overall improvement in prevention, diagnosis, and treatment. There is also a demand for better utilisation of secondary data for purposes of research, innovation, vigilance, public health and population health.

A health data space creates an open, transparent, federated catalogue of data sources and data services, where demanding stakeholders can obtain access to select data in a consistent, secure and trustworthy manner. It is a compliant data management solution with a strict and open set of policies, rules and standards, giving full control to the respective data holders and creating an essential environment of trust.

The growth speed of global healthcare data is stunning. The application areas of big data in healthcare are developing at a massive speed and have further accelerated because of the pandemic. The current state of data collection, exchange and processing in Europe is limited by scale, gaps in standardisation and by legal challenges and national boundaries. Gaia-X provides a solid legal and technical foundation to better manage healthcare data and to ensure European Member States innovate at the required pace to sustain their healthcare systems and compete with the other large-population countries. ■