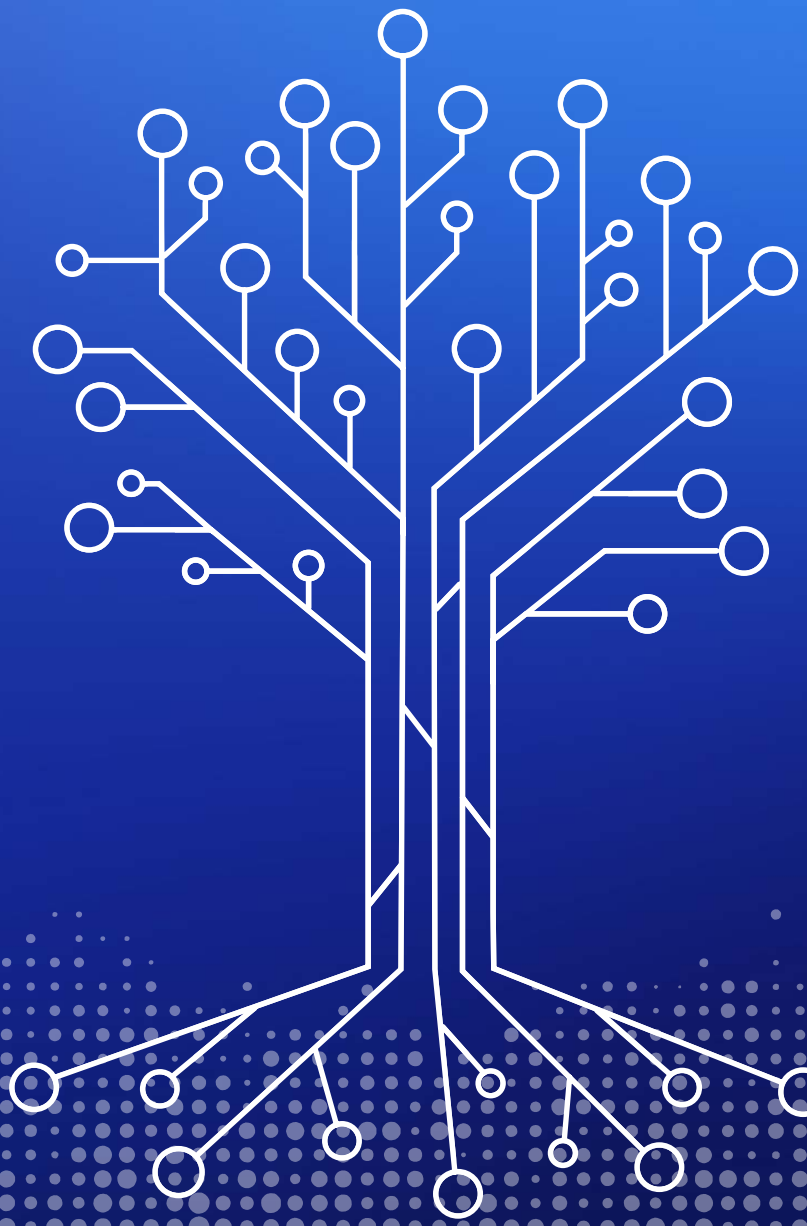
GAIA-X

GAIA-X: A Pitch Towards Europe

Status Report on User Ecosystems and Requirements



Imprint

Publisher
Federal Ministry for Economic Affairs and Energy (BMWi)
Public Relations
11019 Berlin
www.bmwi.de

Status
May 2020

Design
PRpetuum GmbH, 80801 Munich

Illustrations
BMWi / title, p. 8, p. 15, p. 25, p. 32
Shomiz / Getty Images / p. 4

This publication as well as further publications can be obtained from:
Federal Ministry for Economic Affairs and Energy
Public Relations
Email: publikationen@bundesregierung.de
www.bmwi.de

Central procurement service:
Tel.: +49 30 182722721
Fax: +49 30 18102722721

Contents

1. Summary and Outlook

At the Digital Summit on 29 October 2019, we presented a concept paper entitled 'Project GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem'. This launched the project and signalled the start of this review of the user perspective, of the 'User Ecosystems and Requirements' workstream and of the GAIA-X project.

Scaling on the user side is a fundamental requirement for a successful data infrastructure. We want to demonstrate how GAIA-X can contribute to the establishment of a vibrant European ecosystem and which requirements the technical concept must meet for this.

Establishing vibrant European ecosystems

1. We have compiled use cases that illustrate the breadth and diversity of use and that may inspire innovative business models (see Chapter 3). We started in October 2019 with twelve use cases relating to four domains. We now have more than 40 use cases. The current issue of coronavirus is also being looked at with two use cases. GAIA-X will enable opportunities to respond to an extreme medical situation such as the current pandemic with fast and sensible measures through the linking of different data sources (including geodata, reporting data and patient data). A full overview of the use cases is available on the website: www.data-infrastructure.eu/gaia-x-from-the-user-perspective.

2. As the number of use cases rose, so too did the breadth of application. We initially started with four user domains. There are now eight: Industry 4.0/SMEs, Health, Energy, Finance, Public sector, Mobility, Agriculture and Smart living.

The participants in the GAIA-X project are currently working on use cases in working groups for eight domains:

- Energy
- Finance
- Health
- Industry 4.0/SMEs
- Agriculture
- Mobility
- Public sector
- Smart living

3. GAIA-X is a European project. We cover the areas in which the European Commission wants to establish European data spaces almost exactly with the domains mentioned above. The aims that the European Commission meets with its data strategy are congruent with the aims of the GAIA-X project (see Chapter 2).

4. The number of people working in the workstream has also grown strongly in parallel with the number of use cases and the domains. There are now more than 170 people from around 150 companies, research institutes, associations and institutions working in the domains mentioned. There are also increasing numbers of participants from other European countries such as France, the Netherlands, Switzerland, Spain and even Japan.

Requirements for the technical implementation of a federated infrastructure

1. We have used the use cases to identify the requirements for the technical concept of GAIA-X. We expect that a large proportion of the requirements (around 80%) will be identical across the domains.

2. It is key that the understanding of the technical requirements is the same in all domains, i.e. that we

speak the same language. We have therefore developed a fractal model, that illustrates the communication relation within that use cases and to which the use cases will be applied in abstract form.

3. On this basis, we have developed the fundamental and general requirements for a GAIA-X layer, which must be met across all use cases and domains (see Chapter 4). A GAIA-X layer is the central connecting element between different elements of a federated data infrastructure. The GAIA-X layer is fully compatible with the 'X graphic' (see Figure 7), as used by the workstream developing the technical foundations for GAIA-X. For further details, see for instance the document 'GAIA-X: Technical Architecture Release June, 2020', (www.data-infrastructure.eu/gaia-x-technical-architecture).

4. At the same time, every domain will have requirements that are not the same for others or do not have the same weight. In this case, we talk about a 'domain delta'. We have set out an example of a domain delta using selected domains (see Chapter 5).

Outlook and next steps

1. We want to establish data spaces in various domains in Europe. That is the aim of the European data strategy. GAIA-X is a key component for the establishment of these data spaces, as GAIA-X can help to break up data silos and avoid data lock-ins.

   Our aim is therefore for GAIA-X to be successful across Europe and beyond. To achieve this, all relevant stakeholders and initiatives need to be involved.

   Our initiative, which we have organised as a GAIA-X hub, is open to all parties who share the goals of GAIA-X. We are actively inviting new participants to get involved in defining the requirements and to benefit from our network.

   We would like, in particular, for GAIA-X hubs and anchor centres to be established across Europe (and beyond) as nuclei for GAIA-X. There is already a variety of initiatives that are driving forward the development of new business models and data spaces in the EU and its Member States. Together, they form a network that is driving the implementation of GAIA-X regionally, nationally and internationally. This network enables specific national or sectoral features to be addressed and cross-border collaboration to be organised. It will become easier to integrate small and medium-sized companies into the GAIA-X ecosystem. We would therefore like to encourage the European partners to each establish their own GAIA-X hub structure, as shown in Chapter 6. The GAIA-X hubs are the basis for collaboration in Europe and beyond.

2. We want to develop the domain-specific and cross-domain requirements with greater granularity.

   GAIA-X will develop an architecture of standards, as proposed in the paper 'GAIA-X: Policy Rules and Architecture of Standards' (www.data-infrastructure.eu/gaia-x-policy-rules-and-architecture-of-standards). Each domain (user group) has its own standards, interfaces and processes, which should be used for successful implementation of GAIA-X. These will be listed in the next step.

   We want to consolidate the requirements and the domain-specific standard architecture in order to reach a definition of a domain-specific demonstrator as a prototype, which will enable the existing application examples to be implemented quickly. This development will be done in collaboration with technical experts from the GAIA-X project.

3. Expansion to cross-domain functionality is the medium- to long-term aim for the implementation of GAIA-X. We want to move forward step by step, by progressively expanding the GAIA-X layer with additional sub-functionalities.

2. Introduction to GAIA-X

2.1. GAIA-X as a European Project

Sovereignty is a core challenge for Europe. This does not imply that all of the technical components need to be produced independently in Europe. Rather, it is a matter of full control over existing technologies and, in particular, the question of data sovereignty. Data sovereignty is the foundation for data-driven business models, the creation of sectoral and cross-sector data ecosystems, innovations, economic competitiveness and, ultimately, prosperity in Germany and Europe. The objective is therefore to safeguard and expand the industrial competitiveness and, subsequently, the prosperity of the European community by reducing dependence and fostering competition, thus enabling the stakeholders to exercise their sovereignty in data-driven business models and data ecosystems.

Europe's digital infrastructure currently lies in the hands of a small number of major non-European corporations: Europe has no notable operating system developers, no relevant search engines, no global social network and no competitive cloud infrastructure.

Essential data and analysis infrastructures are also provided by companies from outside Europe. European alternatives do not offer any comparable market capitalisation, scalability or breadth of applications; they are active in specialist niches at best. There is a risk of European data being stored outside of Europe or on servers in Europe that belong to non-European companies and will be subject to a so-called lock-in. Innovation and further value creation within Europe may not be feasible in this way.

A variety of stakeholders at national and European level have now recognised the need for action. Our answer is the **GAIA-X project**, which aims to achieve an open, federated, secure and trustworthy data and cloud infrastructure for Europa as the basis for a digital ecosystem.

This approach is indispensable in building a distributed, federated and open data infrastructure based on European values. The approach builds on existing infrastructure solutions. GAIA-X uses an open ecosystem of providers of various size that follow a common GAIA-X reference architecture. The reference architecture ensures that the cloud infrastructures are technically interoperable and create opportunities for collaboration and scaling of business models. Such a network of different cloud infrastructures operated by companies (virtual hyperscaler) is the central basis for the emergence of data-driven ecosystems in the most diverse fields of application, from industrial automation and use in the healthcare sector, through to e-government and many other scenarios. We strive for an ecosystem that distributes sovereignty and benefits among business, science, the state and society in equal measure. Involvement is open to all market participants – including those outside Europe – who share the goals of data sovereignty and data availability defined by GAIA-X.

GAIA-X is a **European project** and is based on European values. The GAIA-X project supports the goals of the European data strategy, which the EU Commission published on 19 February 2020. The strategy emphasises the economic and social opportunities of data as well as European values and EU law (data protection, fundamental rights, laws regarding [cyber-] security/openness, fairness, diversity, democracy and trust). The required trust must be generated among citizens in order to harness the opportunities of digitalisation. This trust is to be guaranteed by implementing European values and EU law. The aim is to enable better decision-making through the use of data for business, science and the public sector, as well as for citizens.

The GAIA-X project is one of the few initiatives at Member State level to be specifically mentioned in the EU data strategy. The objectives stated as part of the Commission's specific measures (e.g. federated cloud,

a rulebook, cloud service marketplace, European data spaces) are likewise integrated in the GAIA-X project. Our aim is for GAIA-X to become the central project in driving forward the 'cloud federation' at European level and beyond. We are also in close contact with our European partners from France, Italy, Spain, the Netherlands, Sweden, Finland and Austria, as well as with the European Commission.

Just like the European data strategy, the GAIA-X project aims to support digital ecosystems in the various sectors and the development of these. The EU Commission wishes to support the development of the following shared European data spaces in particular, which largely coincide with the domains in the GAIA-X project:

1. Industrial data space
2. Data space for the European Green Deal (exploit potential of data in the relation to the environment and climate)
3. Mobility data space
4. Health data space
5. Financial data space
6. Agricultural data space
7. Data spaces for public administration
8. Skills data space (to reduce the skills mismatches between current education and training on the one hand and the labour market needs on the other)

We invite European partners to get involved in this Workstream. This has already happened to a certain extent. In order to reflect the heterogeneous regional economic structures in Europe, we would like to see the emergence of GAIA-X hubs and anchor centers for GAIA-X. There is already a variety of initiatives in the EU and its member states which are driving the development of new business models and data spaces. Together they will form a network driving the implementation of GAIA-X regionally, nationally and internationally. Within this network, national or sectoral specifics can be addressed on the one hand, and cross-border cooperation can be organized on the other. The integration of small and medium-sized enterprises into the GAIA-X ecosystem will be facilitated. The GAIA-X entity will benefit from this. Hence, we encourage the European partners to each establish their own GAIA-X hub structure.

2.2. Guiding principles of GAIA-X

Taking European values as the starting point, seven principles have already been set out in the first concept paper to guide the establishment of a federated and sovereign European data infrastructure. These have been further developed and their meaning has been jointly defined since the publication of the first concept paper. The guiding principles, to which all of the initiative's stakeholders commit, not only support the implementation of GAIA-X, but also reflect the added benefit of a European data infrastructure.

1. **European data protection**
   Data protection is a core element of German and European digital policy. GAIA-X enables European legislation to be implemented, while at the same time permitting case-specific adaptation of data protection scenarios in various domains and for different protection classes.

2. **Openness and transparency**
   Self-describing GAIA-X nodes[1] are used to promote transparency, as well as to create new business and usage models between participants (e.g. such as intermediary activities for data or services).

   Digital ecosystems, which GAIA-X aims to establish, encourage the growth of marketplace solu-

1 GAIA-X nodes are elements of a data infrastructure that meet the requirements developed by the workstream that deals with the technical implementation of GAIA-X.

tions in the user domains. Standardised contracts and procedures reduce transaction costs, data markets can emerge and data availability is improved.

3. **Authenticity and trust**
   Independent and automatable certification and contracting of a GAIA-X ecosystem participant will ensure compliance with the GAIA-X rules (with regard to IT security, data sovereignty, service levels and framework contracts). Full transparency will also be provided through the self-description, for instance, about certified data protection and regulatory criteria met for the products and services offered. Clear conditions for participation in GAIA-X and for collaboration, and common rules for cross-company authentication and access management will strengthen the underlying level of trust, bring down the obstacles to participation and reduce the amount of work involved in bilateral coordination between individual stakeholders.

4. **Digital sovereignty and self-determination**
   A further important added benefit of GAIA-X is the guarantee of data sovereignty: Each user decides for themselves where their data is stored, as well as who may process it and for what purpose, based on the user's own data classification.

5. **Free market access and European value creation**
   The use of technologies with transparent security and openness in an open ecosystem encourages competitiveness, in particular when compared internationally.

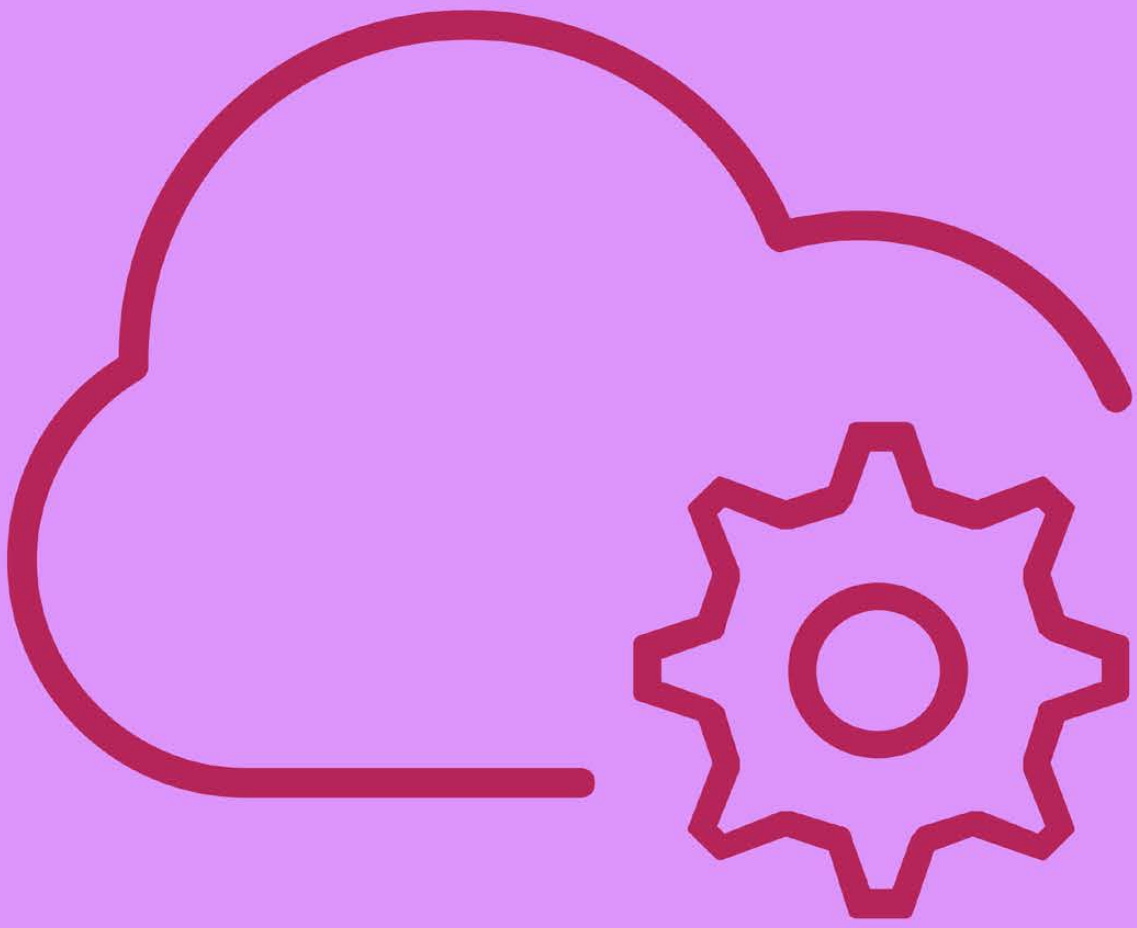
   Users can use the trustworthy decentralised data infrastructure of GAIA-X to access AI applications and data pools. Based on standardisation rules and the different options for managing and controlling the transfer of data, data can be exchanged between companies, organisations, institutions, research institutes and associations, linked with other data, processed, evaluated, and monetised in value creation networks. Centralised and decentralised cloud infrastructures can be linked with each other, enabling data and algorithms to be used safely and data to move along the value creation chain to the applications. These opportunities of data and service sharing can promote innovations, harness synergies and enable new business models to be developed and scaled up in Europe.

6. **Modularity and interoperability**
   GAIA-X gives users access to a broad, relevant and specialised range of products and services from cloud providers, thus enabling the use of tailored solutions. GAIA-X facilitates the portability of data between cloud infrastructures and the combinability of data from different cloud infrastructures. Specialist providers and small providers can also take advantage of this and be successful in the market through modular offerings.

   A high level of interoperability – in terms of technical and semantic standards, as well as interconnectivity at a network, data and service level between edge or cloud instances – simplifies the management of IT interfaces through a federated GAIA-X infrastructure, can avoid lock-in effects, and enables the emergence of data silos to be prevented.

7. **User-friendliness**
   GAIA-X is intended to be clear and intuitive to use for all stakeholders. The provision of centralised services, which the ecosystem needs for secure and user-friendly operation (e.g. authentication), will enable even relatively small organisations to use secure infrastructures.

3. Added benefit, mobilisation and integration of the user perspective



3.1. Added benefit of GAIA-X from the user perspective

Workstream 'User Ecosystems and Requirements' represents the user perspective of GAIA-X and supports the broad and sustained mobilisation of this perspective. The initiatives and the workstream thus offer users from business, science and the public sector the opportunity to be involved in developing the solution and to contribute their requirements for a European data infrastructure. As a result, the requirements from the participating companies, research institutes, institutions and associations are integrated into the technical implementation of GAIA-X, ensuring that a data infrastructure that is focused on users and requirements will be created.

The overarching goal is to create a cloud infrastructure solution that will enable participating companies to become more competitive and better equipped for the future. At the same time European data sovereignty should be guaranteed. The adherence to European values (for example with regard to data protection) will lay the foundations for data-driven operating models and also guarantee trustworthy data exchange. At the same time, GAIA-X enables sectoral and cross-sector networks to be established between companies, irrespective of their size, thus revealing and exploiting economic potential and synergies between them. It also enables collaborative and digital working models, in which innovations are discovered, fully developed and supported. GAIA-X essentially lays the foundations for this and meets the requirements for establishing a data infrastructure in the various domains in future. In addition, the transformation of domains is being encouraged and digitalisation is being driven forth in the various areas. Not only individual companies will benefit from the results of the GAIA-X project. Because networks are established, innovations are promoted and collaborations across sectors are made possible or supported. The increased data availability and improved collaboration within a network will result in new potential for scaling up AI applications, as new AI applications will be able to be used across different sectors.

3.2. Integration and mobilisation of the user perspective for GAIA-X development

The aim of Workstream 1 'User Ecosystems and Requirements' is to achieve the stated added benefits through a broad and sustained mobilisation of the user and demand perspective. This occurs through continual identification, integration, development and implementation of domain-specific use cases, which illustrate the need for and the added benefit of a sovereign European data infrastructure. Representatives and experts from the user perspective are also involved in domain-specific working groups as part of GAIA-X, in order to define, pool and contribute their requirements. These form the basis for the growth of data infrastructures and are integrated into the concrete development of GAIA-X.

The first twelve use cases from four domains were presented in the concept paper at the Digital Summit on 29 October 2019. Since the publication of the paper, the user base has grown: We now have more than 40 use cases and more than 170 people from around 150 different companies, research institutes, associations and institutions are now contributing. These representatives of the user perspective are organised into eight domain-specific working groups ('Industry 4.0/ SMEs', 'Healthcare', 'Finance', 'Public sector', 'Smart living', 'Energy', 'Agriculture' and 'Mobility'). Each domain is led by an expert patron, who ensures that the example uses are well grounded in fact and coordinates the identification of domain-specific requirements. The expert also aids communication and coordination with the other domains so that shared cross-domain requirements can be defined.

The following domains have so far been involved in the creation of GAIA-X and underpin its design with a range of different use cases:

Energy

This domain is working to identify intelligent solutions for power generation, energy storage, power transmission and consumption monitoring. It addresses and supports all areas of potential generated by a federated data infrastructure throughout the value creation chain, from power generation to energy consumption.

Finance

This domain includes stakeholders from financial markets, regulation, supervisory authorities and financial intermediaries such as banks and insurance companies. A federated data infrastructure offers new opportunities to improve processes and efficiency, to collaborate and to create new business models. The collaboration of the relevant stakeholders is highly important from a regulatory perspective in particular. Open digital platforms and the use of AI methods enable a more efficient interplay to emerge between, for instance, the stock exchanges, the supervisory authorities and companies.

Health

Federated and secure use and exchange of data is also extremely important for this domain – in particular as patient data is highly sensitive. The domain deals with the use of AI in relation to health and care and examines the new opportunities in this heavily regulated and decentralised area. Development of the appropriate solutions entails a major workload to satisfy the requirements of the various stakeholders involved. A high level of data protection must also be guaranteed alongside this. The Federal Ministry of Health (BMG) is part of the project GAIA-X and welcomes the establishment of a secure and trustful data infrastructure for Europe. BMG supports the possible contribution of GAIA-X to creating a European Health data space, especially with regard to improving the secure use of healthcare data.[2]

Industry 4.0/ SMEs

This domain taps into the added benefit from the opportunities that arise from the linking and use of data in production environments. The initiative of the 'Plattform Industrie 4.0' formed the basis for developing a federated data infrastructure founded on European values. The aim of the domain is to create an ecosystem for usable value-added services in heterogeneous production environments and so achieve a breakthrough for the broad implementation of Industry 4.0.

Mobility

The Mobility domain is concerned with the services and opportunities that enable energy-efficient, convenient and cost-effective mobility and that can be used intelligently by users. Intelligent services and AI-based business models also depend on links with data from related system environments in AI-based applications. This demands end-to-end data recording, processing and networking and so requires a suitable cloud environment.

2 BMG is not responsible for the content and processes demonstrated in GAIA-X Health Use Cases.

Public sector

As the state is increasingly becoming a user in the digital environment, this domain is also highly relevant. It highlights the added benefit of GAIA-X through the need for security, reliability, trust and transparency. Applications play an important role in public administration and in science, which is also included in this domain. The applications aim, for example, to improve government services or to provide city planners and decision makers with concrete and effective solutions. Such applications depend primarily on a high-capacity infrastructure and broad data availability.

Smart living

This domain concerns the use of data from intelligent components, devices and functions in consumers' private settings to provide various applications and services that make life safer, lower impact, more convenient and simpler. *Smart living* is therefore more than an intelligently networked home – this domain also focuses on the applications and services that this enables.

Agriculture

The domain of agriculture is currently in the launch phase.

The workstream has thus achieved the goal of a broad and sustained mobilisation of the user and demand perspective. The users are key to the creation of data ecosystems. To demonstrate the benefits of GAIA-X, new use cases that illustrate the need for and added benefit of a sovereign European data infrastructure continue to be identified and integrated.

Use cases outside of the GAIA-X work structure should also be developed and implemented.

The following graphic provides an overview of the submitted use cases from Germany, France, the Netherlands Switzerland, Spain and Japan.

Requirements for the technical concept are also being collated from the use cases. Our approach is to pool the requirements for specific domains and also across different domains, bringing together requirements that apply across domains and form a common GAIA-X layer and where domain-specific expansions (domain-specific delta) may be necessary.

There is a fundamental consensus that GAIA-X must not interfere with existing business models. In this respect, the technical concept of GAIA-X must take up and implement already existing domain-specific solutions.

The detailed descriptions of the use cases shown can be found online at www.data-infrastructure.eu/gaia-x-from-the-user-perspective.

Overview of the submitted use cases from Germany, France, the Netherlands, Switzerland, Spain and Japan

Energy

- Infrastructure data for new business models

Agriculture

- Agri-Gaia

Health

- Smart Health Connect
- Berlin Health Data Space – AI against acute kidney failure
- Recupera REHA
- Digital Twin
- Patient Empowered, Privacy Secured
- Improve Chronic Heart Failure Patient Management
- KAMeri – Cognitive Occupational Safety for Human-Machine Interaction
- CarePay
- COVID-19-Dashboard & Data Hub
- Surgical Platform for AI-based Risk Identification
- KIKS – Artificial intelligence for clinical studies
- Research Platform Genomics
- Future Care Plattform
- Image Sharing for Medical Professionals and the Citizen
- Framework of medical records in Europe
- EMPAIA (EcosysteM for Pathology diagnostics with AI Assistance)
- Differential diagnosis
- Medical crisis management and research platform 'UNITY'

Industry 4.0/ SME

- Smart Manufacturing
- Supply Chain Collaboration in a Connected Industry
- Collaborative Condition Monitoring
- Shared Production: cross-factory and cross-company production as a showcase
- IIoT Platform with out of the box MES Applications
- Smart Predict
- Connected Shopfloor

Mobility

- Data interoperability with data sovereignty
- The Testbed Lower Saxony is ready for GAIA-X
- Digital parking space management – Seamless Parking

Public sector

- Space4Cities
- Digital administration chatbot
- High-performance and Quantum Computing as a Service
- Smart Infrastructure Management
- Quality infrastructure digital (QI digital)
- IntraX – Transportation Infrastructure
- Open Source Orchestration Framework

Smart Living

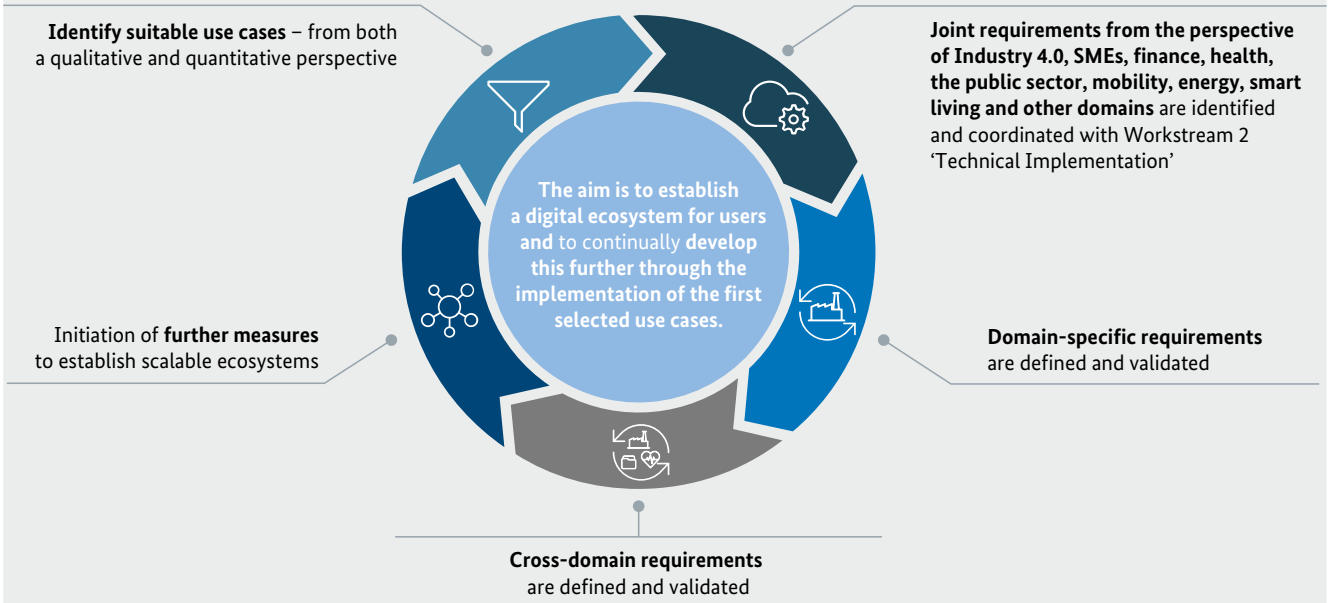
- Smart Living
- Energy Efficiency
- Everyday life-supporting assistance solutions for Smart Living
- Security

Finance

- Financial Big Data Cluster
- Creation of a secure basis for increasing data sovereignty
- Sustainable Finance
- Optimised networked techniques in the prevention of money laundering
- Research of new methods to increase market integrity
- Improving the database for test and simulation environments for monetary policy decisions
- Stable Supply Chain Finance

Use cases submitted after the Digital Summit

Figure 1: The mission of Workstream 1 'User Ecosystems and Requirements' is broad and sustained mobilisation of the user perspective for GAIA-X



3.3. Benefits of GAIA-X

The use cases illustrate that GAIA-X can form the basis for ecosystems that effectively integrate the strengths of different participants and promote cooperation.

Because GAIA-X[3]:

- Simplifies the management of IT interfaces and integration, especially regarding multi-cloud strategies and data-pooling; this is done by a high level of interoperability of compatible products. In this respect GAIA-X helps to avoid lock-in effects. In addition, domain-specific data silos, which could not be connected and evaluated so far due to a lack of data interfaces, are broken up. In this respect GAIA-X helps to avoid lock-in effects. These three components enable or facilitate tailor-made solutions for the application. In addition, an important contribution can be made to the acceptance of the application of AI for particularly sensitive data.
- Integrates existing digital and cloud-based state-of-the-art products and services. GAIA-X makes it possible to integrate additional modular offerings, for example from specialist- or small-scale providers, to meet specific needs.
- Offers full transparency by providing authentication of verified data protection and regulatory criteria for the products and services offered. It provides transparency regarding the level of con-

3 See BMWi project 'GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem', 2019, p. 15.

fidentiality of all the ecosystem's participants through self-description. This is reflected in the guarantee of data use monitoring (data sovereignty).

- Makes it possible to store data where users consider it useful to do so, in light of the respective data classification. This means that the user can retain command over particularly sensitive data, while simultaneously sharing other data with partners for joint use.

- Creates the preconditions for optimising the users' data strategies. Decentralised and/or centralised cloud infrastructures, can be linked up with one another. This link-up generates options regarding how data and algorithms can be used securely. So, for instance, various cooperation partners along the value creation chain are also given the possibility of migrating data to the applications. To protect intellectual property, users can thus retain their own algorithms and data, for instance.

- Makes an important contribution to the emergence of digital ecosystems in the various user domains, by enabling them to make the transfer from bilateral individual-project solutions to marketplace solutions. Standardised contracts and procedures reduce transaction costs, data markets can emerge and data availability is improved.

We assume that a large part of the requirements will be similar across domains. For the analysis, we have firstly developed a common model, the so-called fractal model, based on the submitted use cases and the represented domains. This model represents the communication relationships specifically named in the use cases in an abstract form. We have found that we can apply this model across domains. Secondly, we have derived a GAIA-X layer from the model, which embodies the basic, cross-domain requirements for a European data infrastructure. This defined GAIA-X layer is shared by all domains.

This layer initially covers the cross-domain requirements. Beyond these requirements, however, there are also domain-specific requirements for GAIA-X. The common underlying GAIA-X layer and the additional domain-specific requirements (exemplary for selected domains) are explained further in the following chapters.

4. Shared/cross-domain requirements: what does GAIA-X need to offer across different domains?

4.1. The GAIA-X ecosystem

An infrastructure provides a basis on which data can be exchanged, digital services can operate and added value can be created. For data exchange, value-added services require a basic infrastructure that provides access and transport protocols, services and rules ('mechanisms') that can be used together (referred to as 'basic services and mechanisms', 'basic services', 'basic mechanisms' or 'basic framework' in singular), to organise the exchange of data between participants. Existing IT systems need to be able to build on this basic framework and establish new value-added services on it themselves.

The joint requirements of the aforementioned use cases can be consolidated in a GAIA-X layer. They may be part of a general and cross-domain basic infrastructure or ecosystem. This is not an exhaustive list; additional requirements may be described depending on the usage domain and use case and may be needed to build on this ecosystem.

An implementation of a digital ecosystem that enables shared use and is based on a shared reference architecture would be desirable to enable added-value offerings to be scaled and to allow value creation that builds on digital basic services and mechanisms to be easily applied. This makes the basic services more universal beyond the individual use case, in the form of a basic framework shared by all application domains and so makes implementation easier for users. Such a shared ecosystem can therefore drive its own distribution through scaling and thus become an enabler of digital value-added services in various application domains.

4.2. Participants in the GAIA-X ecosystem

Ecosystem participants ('entities') can essentially be data providers, data consumers, or simultaneously consume data and generate the same or new data. In this abstract way of looking at this, a data provider can be seen as a supplier and a data consumer can be seen as a customer. An 'asset' may have a physical form or be an immaterial (software) service. From a data perspective, the 'supplier' role is a data provider, the 'customer' role is a data consumer and the combined role of 'customer/supplier' is both a data consumer and data provider. The generic form is the same for all the roles.

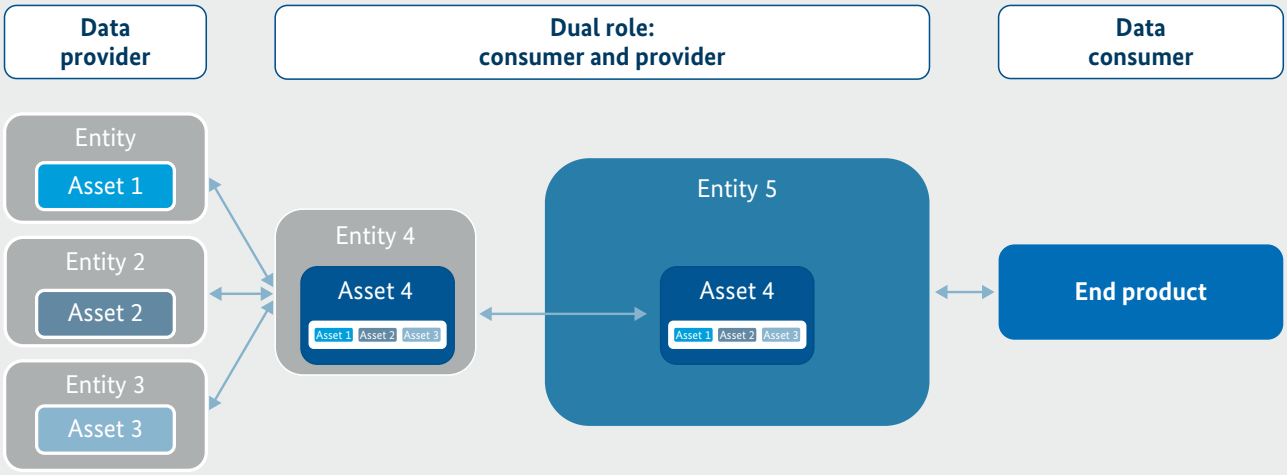
We can use a three-level value creation chain with various entities with the following roles (supplier, customer/supplier and customer) as a simplified example of the process for an operational ecosystem:

1. Multiple suppliers (entity 1-3) each supply or link an asset to the next participant (see Fig. 2).

2. Entity 4 performs a dual role as a customer and supplier and in the second stage, creates a new asset (4) from the assets supplied by entities 1-3, as a physical device or immaterial software (app) or service. This structure may also occur with mixing across multiple levels.

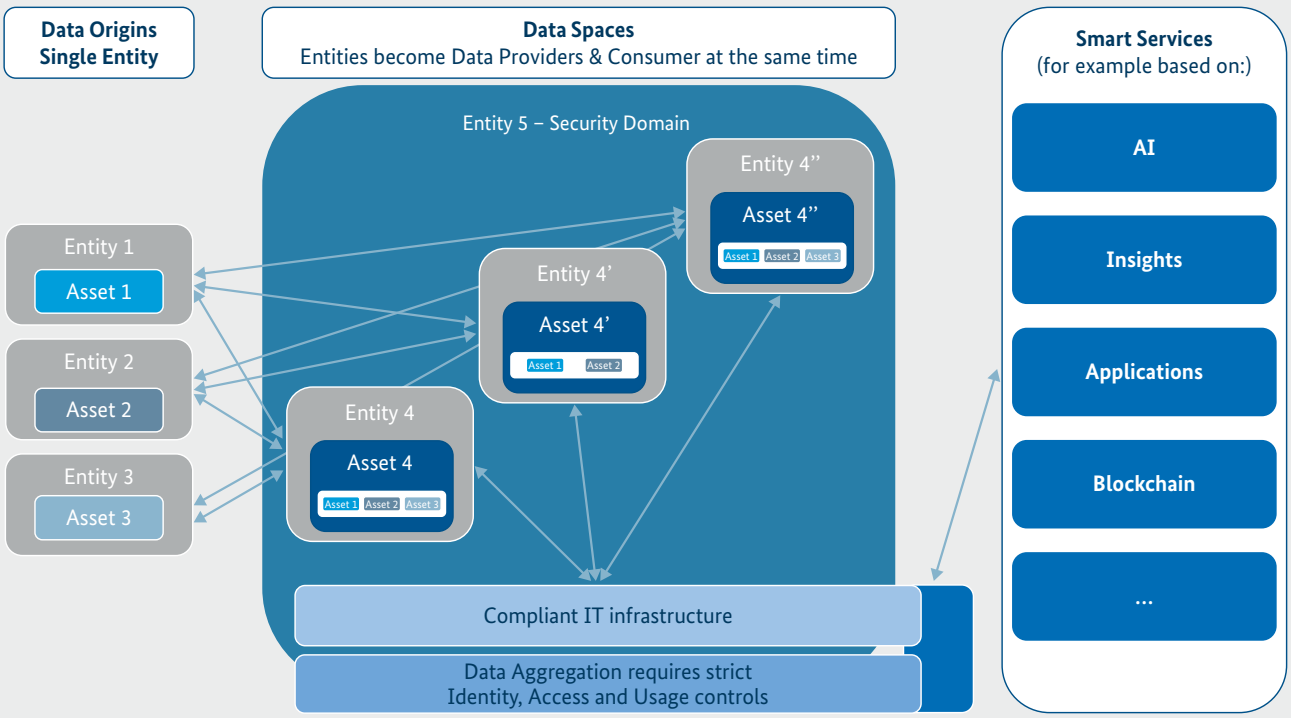
3. In the third stage, asset 4 from entity 4 takes up its function in the environment of a following entity (customer) in the chain (entity 5), which in turn may itself be a data consumer and provider. The product itself (as a physical device or immaterial software (app) or result of a service) may be handed over to an end consumer. This last instance is also a data consumer in the simplest case.

Figure 2: Manufacturing Industry 4.0 – three-part fractal for communication



Source: BMWi

Figure 3: Communication relationships – three-part fractal

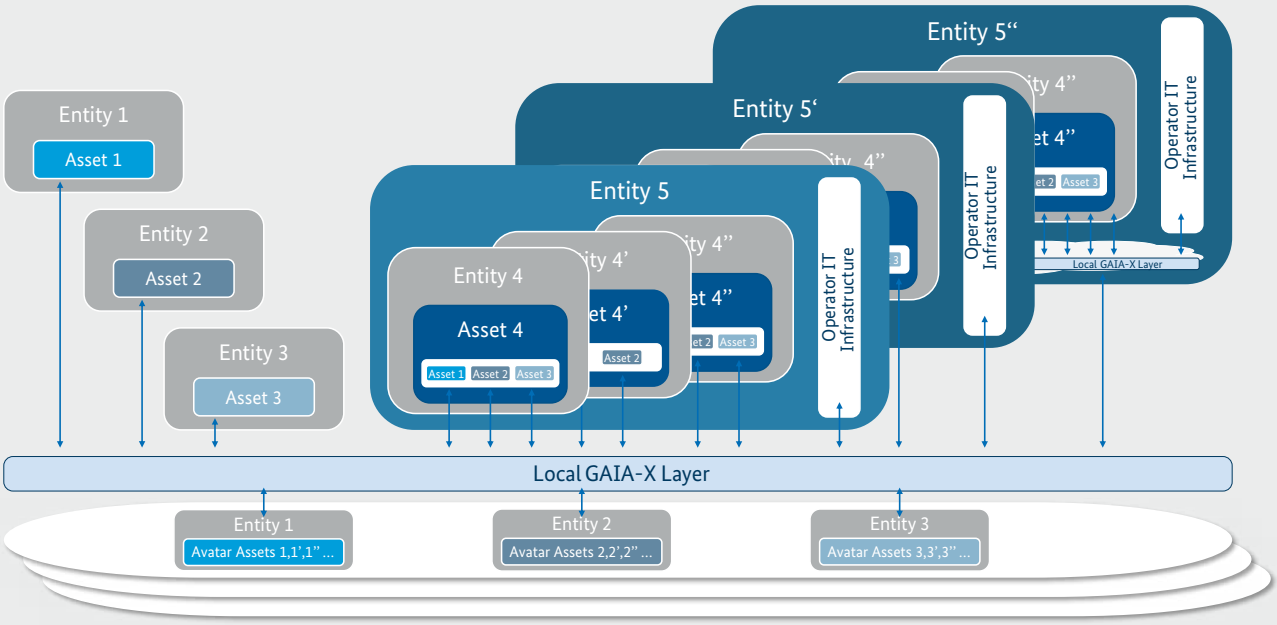


Source: BMWi

The described three-stage structure (supplier, customer/supplier, customer, data provider, consumer/provider and consumer) is there treated as a possible fractal[4] of a multilateral structure. It should then be simple to scale this to value creation chains, networks and ecosystems by combining the three-part fractal as a component of chains and networks in which the subsequently described findings and laws continue to apply (see Fig. 3).

The communication relationships shown in Figure 4 can be transferred to many user domains. Manufacturers (here: entities 1-3) produce data generators or consumers (assets), e.g. components. Using a GAIA-X infrastructure, manufacturers communicate with "their" assets, while these are built into other assets (asset 4 - 4"), e.g. machines. These are used in the value chain of the superordinate entity (entity 5-5"), e.g. factory, hospital or geodata service. Using the assets

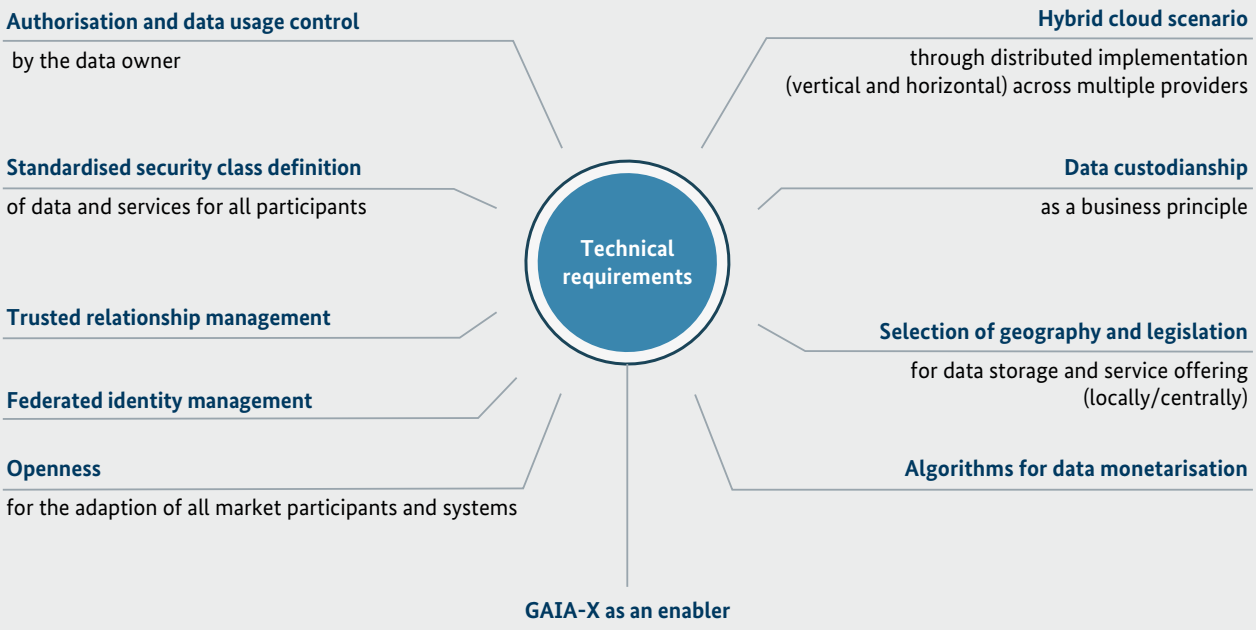
Figure 4: Abstract representation – embedding a three-part fractal into a data ecosystem with participants in multiple security domains 5, 5', 5", GAIA-X layer centralised or decentralised, cloud and 'on-premises'



Source: BMWi

4 The smallest possible fractal of a multilateral structure is a three-part structure – shown here as a provider, consumer and combined provider/consumer. It should then be simple to scale this to value creation chains, networks and ecosystems by combining the three-part fractal as a component of chains and networks in which the subsequently described findings and laws continue to apply.

Figure 5: Key elements of a data ecosystem framework



Source: BMWi

communicating via the GAIA-X infrastructure, all entities can build and offer smart services for their own purposes. Communication in the GAIA-X ecosystem transforms this from a complex mesh and network structure (see Figure 3) for all assets into a simple point-to-point communication with a GAIA-X interface.

Figure 5 shows the key requirements desirable from the user's point of view, which the technical concept of GAIA-X should fulfil. We discuss these cross-domain key requirements in chapter 4.3. In chapter 5 we describe the so-called domain delta or the application of these key requirements to individual domains.

4.3. Key requirements for a GAIA-X ecosystem

The key requirements were derived from manufacturing use cases. However, they are valid for all domains:

Openness, hybrid cloud', basic framework, open vision

This describes the requirement for the existence of an open ecosystem that spans domains and value-added services in the form of a network that exchanges data and services and in which value-creating actions can be performed. It needs to have centralised, decentralised and heterogeneous organisation in terms of geography, legislation and technology.

It needs to be easy to provide, link up and integrate data from non-homogeneous IT systems from different stakeholders, with minimal barriers for SMEs (user friendliness, usability of familiar interfaces and the ability to use it without extensive technical expertise are specified for example) and it should be possible to aggregate data across many 'operators' (participants and data sources). It should be possible to operate nodes ('participant', 'access') in a local device (manufacturing use case: industrial control), in an edge computer, in local on-premises clouds of an operator of data providers and on centralised cloud systems with different geographies and legislation. Data exchange, enhancement, linking, analysis and evaluation need to be possible between the various security domains. It also needs to be possible to store data with distribution in all aforementioned applications. In the event of 'mixed' data allocations, the protection classes and confidentiality rules mentioned below are also to be satisfied for a datum (data set). These elements can be used in the GAIA-X basic system to build a shared data space.

Existing protocols and interfaces are to be supported, potentially meaning a wide variety of standards depending on the domain.

The GAIA-X basic system is to offer a defined process for restoring functionality within defined time frames in the event of an incident or, more generally, a failure of functionality. Emergency concepts of participants, especially SMEs, shall be supported in a resilient way. They must be able to continue their own value-added processes in crisis situations in a robust, structured and at least subsistence-preserving manner. It should offer legal, technical and procedural transparency so that users of GAIA-X can integrate this into their own emergency management processes with reliability.

Several user domains require real-time processing, as well as demand-based computing or storage capacity. There are different requirements for real-time processing depending on the usage domain. Profiles for certain use cases will need to be defined in a solution domain. Requirements in this area will also influence the deployment of the GAIA-X nodes, for example (cloud vs. edge cloud vs. on-premises edge).

GAIA-X should therefore define a framework (cf. Policy Rules and Architecture Standards "PRAAS") whose implementation by the individual participants qualifies them for integration into the ecosystem. The interested parties should be able to choose supported compliance levels, which are independently verified. GAIA-X members should be able to further develop the framework in an open process and a controlled life cycle in committees. Services with compliance status should be declared via a directory or catalogue and also – if requested by the participants – communicated.

'Ecosystem for algorithms and methods of data monetisation'

New, previously unknown business models will emerge, synergies in existing and new value-added networks of stakeholders will be exploited. The ecosystem is intended to be used as a marketplace for data monetisation and to create incentives for data exchange, spanning across different stakeholders in the ecosystem.

The purpose of the ecosystem also includes linking and enabling the use of data and services, including for algorithms such as those needed for AI applications and data pools for training these applications.

Primary data with different security requirements for specific domains should lead to secondary data with lower security requirements, which will become available for monetisation – including across different domains. Domain-specific releases and approvals or releases shall be storable and retrievable with the primary data.

Companies should have the opportunity to develop and market integration components for GAIA-X such as software libraries. These companies will then have access to all necessary information, including authentication mechanisms etc. The developed products will then be owned by the companies (open source licence conditions are not mandatory) and can be put on the market as intellectual property (IP).

Data custodianship, authorisation and data usage monitoring, ability to select geography and legislation

Data owners want to decide for themselves which data is shared with which users and with which access rights, and the purpose for which the data is processed. Data sovereignty is central to the ecosystem. In particular it should be possible to select and control the localisation of (particularly) confidential data, for example with regard to individuals, expertise or production secrets. There also needs to be the ability to individually select which data is to be shared at each level of the value creation chain for each participant, from the data provider to the data processor and data consumer (see below). It should be possible to select privacy at every level for the data in individual cases according to the data owner's decision, to pass it to certain dedicated consumers for processing ('sharing' with named data consumers) or to make it publicly visible and/or usable for all. Data usage provisions may be more complex than described here. The use of the data and services should be traceable, for example through a logging mechanism that the data owner can view at any time for instances of access right down to the level of a single data record. A role-based authorisation concept and access logging that can be viewed should also be provided. These roles shall be able to satisfy regulatory requirements and it shall be possible to define access rights for specific times and geographical areas.

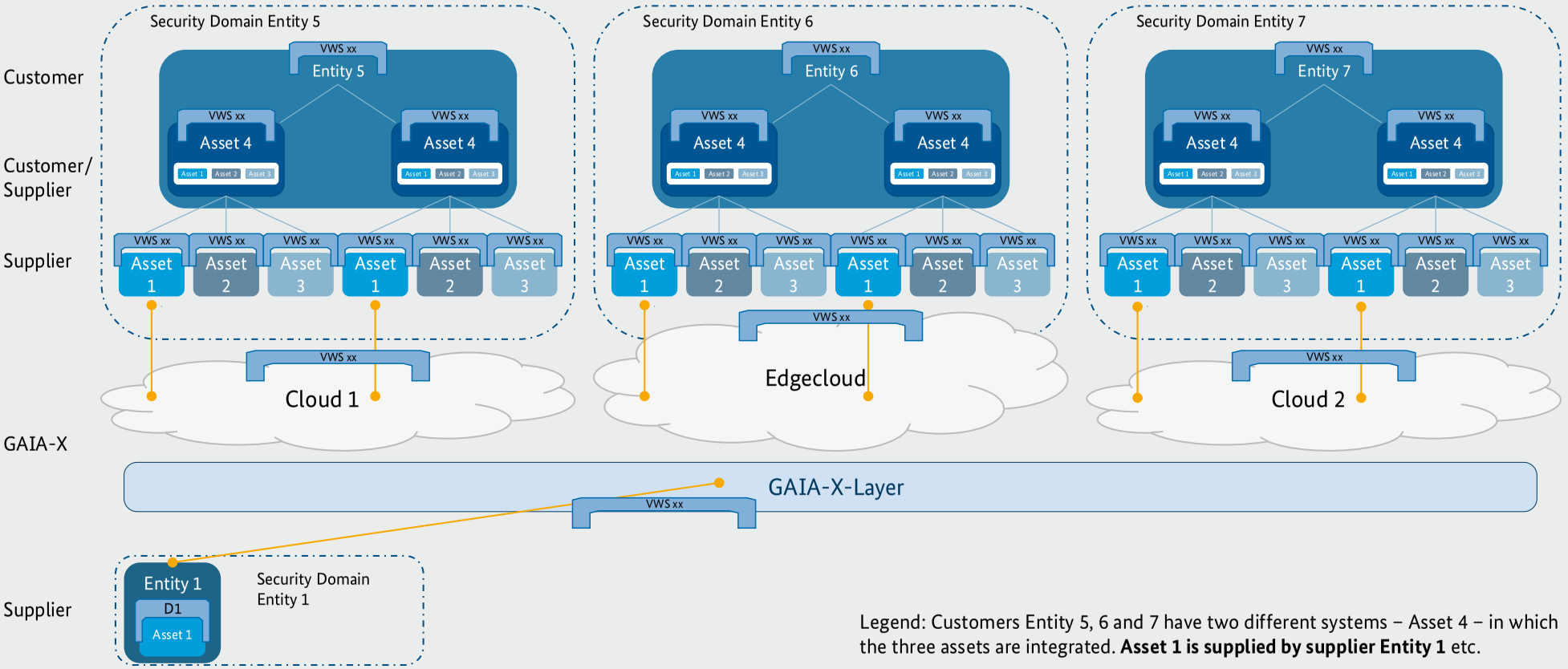
In such an ecosystem, it should be possible to agree on all necessary data exchange agreements in a legally compliant manner according to the purposes and roles described.

Federated identity management, trust relationship management

Visitors should be able to get an overview of the GAIA-X Federated Catalogue anonymously and without registration. Participants will need to complete a one-off qualification process for identification purposes when they choose a service and wish to obtain it. Identity management is to be standardised; the authentication of participants and the access control within GAIA-X is to be controlled by the ecosystem. The GAIA-X nodes need to function as 'trust anchors': secure communication is a basic requirement and the ecosystem needs to be a trustworthy infrastructure for data exchange. The provision of clear and traceable measures (rules) for participation and collaboration across company boundaries (security domains) provides the foundation for shared use of data in an ecosystem. It must be possible to standardise the roles of participants and select different options. It must be possible to organise multilateral participant groups.

It is important to us that the basic services of the GAIA-X ecosystem are to have in depth defences and be protected against external and internal cyber threats. For this purpose, a security organization is to be implemented, which runs an ISMS (Information Security Management System) at all times and maintains its certification. Furthermore, this security organisation must also ensure sufficiently effective protection of the integrity, confidentiality and availability of the GAIA-X basic services and – insofar as this lies within the responsibility of the basic services – the relevant framework conditions for processing its users' data, in line with the protection requirements for the data, services and applications. This includes a coherent concept for encryption key management (cryptography), which contains encryption key and certification management controlled by GAIA-X and also offers customers the option to supply and protect this information independently.

Figure 6: View from the user's perspective: Interoperability layer for applications and participants of a data ecosystem framework



Legend: Customers Entity 5, 6 and 7 have two different systems – Asset 4 – in which the three assets are integrated. **Asset 1 is supplied by supplier Entity 1** etc.

Source: BMWi

Standardised definition of protection classes for data and services; standardised semantics; legally compliant agreement

The semantics for data exchange and for the use of services are to be standardised and organised at an overarching level. It should be possible for data to be exchanged between different security domains within a participant (company) and also between participants (companies) in different security domains in the GAIA-X network. Interfaces are to function between security domains of the participants, taking into account the security requirements, and enable secure data exchange. The security requirements are to be scalable. It should be possible to store and use algorithms and data in accordance with the relevant IP rights. Access rights to data are to be retained when data is passed on in the ecosystem and are to remain traceable.

The following shall also be available for interoperability of participants in the basic system:

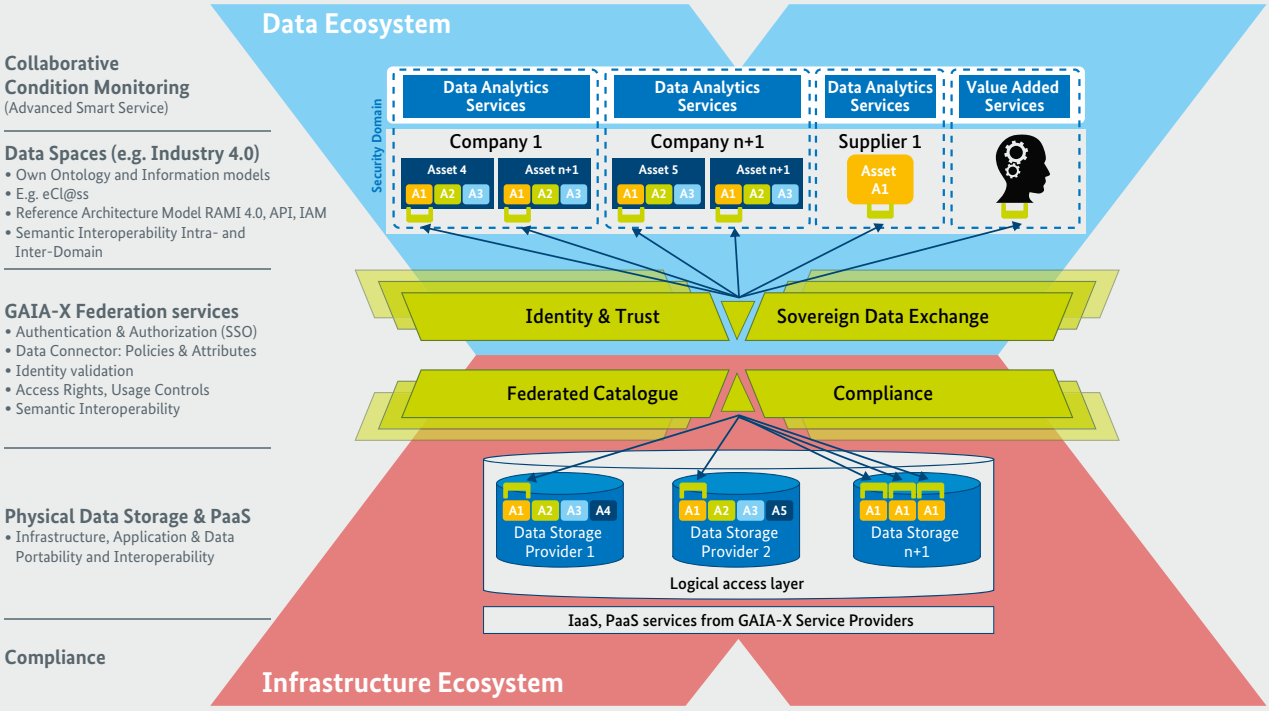
- One of more protocols for transport and interoperability.
- Federated access and trust management (identity and trust; see above).
- At least one basic ontology and semantics for operating the ecosystem and interoperable interaction between participants (together with semantic enablement layer for the definition of further shared semantics).
- A shared definition of basic metadata for identifying participant data that enables interoperable exchange of data and applications (services) across domain boundaries.
- At the same time, support for existing domain-specific standards shall be enabled.

API, container, administration shell[5]

GAIA-X is designed to enable cross-domain interoperability. In a first step, the transport of domain- or application-specific data streams shall be enabled. In order to simplify entry into or participation in GAIA-X, data transport or data exchange is to be made possible only between the participants in a so-called "grey channel". Especially for SMEs and existing systems, participation in GAIA-X would then be easy. In a next step, the interoperability layer is to be expanded using the above-mentioned common properties and elements of the basic system, thus enabling interaction between participants of different applications and domains.

5 The administrative shell is the implementation of the digital twin for Industry 4.0.

Figure 7: Workstream 2 Draft: key elements of a data ecosystem framework and application of the basic communication fractal



Source: BMWi

5. Domain-specific requirements (domain delta) and comments from selected working groups

5.1. GAIA-X health domain

This section was created in close consultation with Working Group 6 of the Plattform Lernende Systeme (Germany's Platform for Artificial Intelligence).

Barely any other area of society has such a directly positive benefit for citizens as the health sector. The potential enabled here by digital technology such as networking via cloud infrastructures or artificial intelligence is still a long way from being fully exploited. There are strong reservations about the analysis of personal health data, and with good reason. However, the use of modern digital technology can generate medical benefits for society as a whole through the analysis of health-related information in compliance with data protection aspects.

The availability of cloud solutions is a key element in achieving digital advances in the health sector as storing data in a trustworthy and centralised cloud service enables broader analyses to be conducted than with isolated solutions in the networks of individual health partners, whilst complying with stringent security standards. For the innovative medicine and healthcare of the future, health data needs to be available from different sources and it must be possible to combine and process this whilst also meeting the highest standards of security. The integrative process will enable better diagnostic methods and tailored therapies that benefit all patients.

To achieve this vision in Europe, we need an infrastructure such as GAIA-X to enable distributed research and health systems to be used with trust. These systems need to fulfil different requirements depending on the particular problem. Each GAIA-X healthcare application centres on state-of-the-art tools for various tasks. This includes data protection, which has other requirements depending on the protection classes of the health data; data trusteeship; identity management for access control and authorisation; openness for the addition of further stakeholders and data types; and the opportunity to implement distributed health data systems in all EU countries in accordance with the law.

The GAIA-X health domain group is currently analysing a variety of digital health applications that are to be implemented with GAIA-X. The use cases cover the diverse nature of the health sector, with cases from care, research and the health industry. The current issue of coronavirus is also being looked at with use cases. The opportunity to respond to an extreme medical situation such as the current pandemic with fast and sensible measures will be made possible by GAIA-X through the linking of different data sources (including geodata, reporting data and patient data).

The following requirements can be abstracted from the various use cases as general requirements for the health sector:

Openness in the health sector means that new players (care providers, researchers, manufacturers and service providers) and data types can be added (for example for the combination of genomics and pathology data) and innovative technologies (such as support for decision making) can be made available for citizens swiftly. It is important to establish shared semantics and shared standards for communication and data access to enable useful exchange of data and the integration of data (interoperability).

Hybrid cloud scenario: Concepts for distributed storage and processing need to be implemented in open-access clouds and in local health infrastructures. For example, analysis results are pooled centrally in a public cloud for integrative analysis, but clinical treatment data with high security requirements is analysed in the data centre at a hospital and the anonymised results are then transferred to the centralised solution.

Ecosystem, including algorithms for data use: Many successful health projects have so far been publicly financed research projects. From the perspective of the health sector, algorithm monetarisation therefore means establishing clear rules for how the data basis and the output can be used by economic stakeholders (licensing, infrastructure sharing, open source algorithms, open data).

Selection of geography and legislation: From the perspective of the health sector, the ability to select the geography and legislation means that it will in future be possible to use big data and AI efficiently in the health sector across the entire EU. It needs to be easier to expand research projects from one region to another, and from one EU country to the others. In Germany, region-specific laws (regional hospital laws, etc.) also need to be observed.

Data custodianship, authorisation and data usage monitoring: In the health sector, custodianship is often assumed through data access committees. The heterogeneity of the committees and commissions could be brought together and professionalised with a standardised digital platform. The health sector has the specific need of checking that the relevant consent has been granted (and obtaining this if necessary) – thus ensuring and monitoring concrete authorisation for access to and use of data.

Standardised definition of protection classes (for data and services): The health sector requires a fine-grained protection class concept with records of patient consent. Access needs to be logged where necessary. Specifically, it is not only access to individual data that must be protected, but any linking must also be subject to special protection if applicable.

Federated identity management, trusted relationship management: In the health sector, data is mostly pseudonymised locally in a very large number of different trust centres. Identity management that spans trust centres is needed for patients (FAIR principles) and for analytical access (AAI).

5.2. 'Public sector' domain/here: 'geodata' sub-domain

The results of the work by bitkom AK Geo serve as reference. These are documented in the position paper 'Geoinformationen: Nutzung optimieren durch vernetzte Geodatenplattformen' (Geoinformation: optimising usage through networked geodata platforms) (https://www.digitalwahl.de/sites/digitalwahl/files/2019-09/190916_pp_geoinformationen.pdf).

There are currently numerous obstacles when it comes to processing public geoinformation, which limit effective use. Geodata from public authorities is generally not straightforward to combine and process with automated processes. The currentness of the data and the structures of the data models from public sources differ widely. The data is largely static, with dynamic data remaining rare in this area. GAIA-X will help to overcome this situation.

GAIA-X shall supply geodata as a key element of the digital infrastructure: References to location and time are indispensable for many data analysis processes. Public geodata and geodata services should therefore be supplied in accordance with open data principles in a way that is open to technology, is comprehensive and is application-neutral for science, administration, business and society. The trends in Germany towards a platform economy should be particularly taken into account. These specific requirements for geodata are set out in detail in the 'Public sector' domain.

Creating a platform-based digital infrastructure for geodata: In future, public geodata services need to be as quick and easy to use in applications as apps are on a smartphone. This requires distributed and networked geodata platforms. Public sector organisations compile geodata in particular for planning purposes and public services. GAIA-X will provide the administrative authorities with the functions they need to use this geodata for specific situations in a collaborative real-time environment.

Providing federated geodata catalogues: Geodata is currently already catalogued digitally and made available through special geodata or open data portals. It must be ensured in future that every data set or data service can be found through any data portal. GAIA-X can, for example, support interdepartmental coordination and interdisciplinary collaboration with suitable search and processing functions.

Taking into account open standards and openness to technology for geodata processing: A greater focus on users and customers requires the broad use of geodata/geodata services. This includes openness to technology, as well as open standards and defined APIs. GAIA-X will take particular account of standards for the general IT and Internet industry.

Ensuring machine readability and automated analysis for geodata: Geodata from the public sector should be made available through GAIA-X in an automated process that enables transfer, transformation and analysis for direct use. Variable processes between data sources and processing will enable the results to be competitive when entering the market via GAIA-X and an increasing number of new combinations to be used for specific situations. The provision of corresponding services should enable data to be analysed at the storage location without requiring high-volume geodata to be downloaded.

5.3. Smart living domain

The smart living ecosystem is becoming a strong growth market. The development of digital applications is advancing swiftly in the areas **convenience and security, energy efficiency and management and for independent living in old age**. There are 23 million existing homes in Germany that could be fitted with smart home and smart building solutions with intelligent devices.

Smart Living comprises different, today often still fully separate areas. These are the housing sector, the electrical sector, the energy industry, healthcare and the skilled trades. Driven by research and development in the field of digitalization of buildings and infrastructures as well as artificial intelligence (AI), these areas work together. For successful collaboration, it is therefore important that data on these assets can be exchanged from the building, apartment, central building services engineering right down to the individual household or smart home device or an individual meter (smart meter).

A housing sector survey revealed that 38% of those companies asked currently have access to important data in a digital and analysable format. However, to develop revenue models, it must be possible to exchange this data without major manual interventions in data processing. This is currently the case for just 11% of the surveyed companies. Among other things, the implementation and application of GAIA-X will play an important role here in closing these digital data gaps.

GAIA-X aims to build a secure, scalable, high-performance, European cloud environment in the smart living sector, which includes local edge devices to avoid high latency times and enables simple and trustworthy access to a multifunctional cloud environment in the area covered by the GDPR.

The GAIA-X 'Smart living' working group is compiling and analysing smart living applications that can be implemented with GAIA-X. **The first use cases have already been published** in the areas of **artificial intelligence, convenience and security, energy efficiency and management**, and **independent living in old age.**

In the smart living sector, **openness** means making the data available for all involved entities, such as between the housing industry and energy suppliers. For example, the next generation of intelligent **power-to-heat-systems (intelligent heat pumps)** or **power-two-mobility-systems (intelligent wall boxes)** can be integrated into the network in a network-serving and cost-saving manner via intelligent measuring

points **(smart meters)**. All energy flows can be optimised by intelligent **control and forecasting systems**.

To do this, smart living requires **hybrid cloud scenarios** for the digitalisation of buildings and infrastructures and for the use of artificial intelligence (AI), as already highlighted in the **'ForeSight'** research project. In principle, every asset from the washing machine to the room thermostat becomes a provider of data, which needs to be integrated into the local building infrastructure through the technology and also at a more abstract logic/semantic level in cross-building applications from service providers and manufacturers. The heterogeneous cloud and on-premises systems that currently exist can be transferred into these hybrid cloud scenarios and further developed into a federated GAIA-X cloud infrastructure for smart living.

The aim of the smart living ecosystem and its entities such as housing companies, service providers and manufacturers is to not only guarantee technical interoperability of all smart living assets. The aim of the ecosystem is also to create semantically interoperable data infrastructures in accordance with international standards. The focus is firstly on the semantically interoperable representation of the various assets as 'digital twins'. Secondly, standards for secure data exchange and data monetization must be set in order to establish sovereign smart living platforms in the GDPR area. This involves linking data from various sectors/domains to implement savings (such as in the energy sector) and guarantee optimum communication across all entities and assets.

From the perspective of the GAIA-X 'Smart living' working group, **geography and legislation** refers to the European judicial area, in which smart living projects are to be conducted. In our opinion, frameworks for data sovereignty and the use of Artificial Intelligence (AI) shall be established on a European level. By freedom of choice of geography and legislation we mean being able to choose from entities and assets for this European judicial area.

Data custodianship, authorisation and data usage monitoring: The smart living domain may have the highest requirements for security and privacy. In many applications, personal and therefore highly sensitive data subject to the GDPR are processed. GAIA-X must ensure that the data owner always has control of their data and is able to grant third parties (data consumers) access to the data for specific purposes. GAIA-X must therefore ensure the sovereignty of the data producer. Such authorisation management must have a very granular structure for the smart living sector that must be able to extend down to the sensor level. Implementation of these requirements demands comprehensive identity and access management down to sensor level in an edge cloud architecture.

Standardised definition of protection classes (for data and services): Smart Living data can be very sensitive if, for example, it contains information about infrastructures that are subject to the German regulation for determining critical infrastructures according to the BSI law. In this case, data storage must meet very high security standards and may only take place in data centres certified by the German Federal Office for Information Security (BSI). The heterogeneity of smart living data means that processes must be established to determine the protection requirement in order to assign the smart living data to the various protection classes. A standardised method is needed for identifying individual data so as to differentiate between 'public' and 'non-public' smart living data. Consideration must be given to whether a security management system will be offered for consumers and small companies. The data owner can use this method to exercise their data sovereignty and pass on parts of their data for further use by third parties.

Federated identity management, trust relationship management: Within the smart living domain, a data ecosystem is created from a wide range of different entities and assets, which must each be protected through intelligent identity and access management. The data providers include, for example, residents,

owners, service providers and manufacturers. In addition to the building and apartment and the data relating to these, their assets also include other assets such as electric charging points and building access systems. Particular consideration must be given to the needs of the individual stakeholders – such as residents, landlords, nursing services, suppliers and technicians – in the area of safety, with different types of authorisation. Comprehensive identity and access management must be set out from the perspective of security and safety.

5.4. Finance domain – the financial big data cluster (FBDC)

Based on the five following problems, the SafeFBDC concept is used to illustrate that new findings can be obtained from large and heterogeneous volumes of data with the use of artificial intelligence and that answers can be found to complex questions in the area of financial markets.

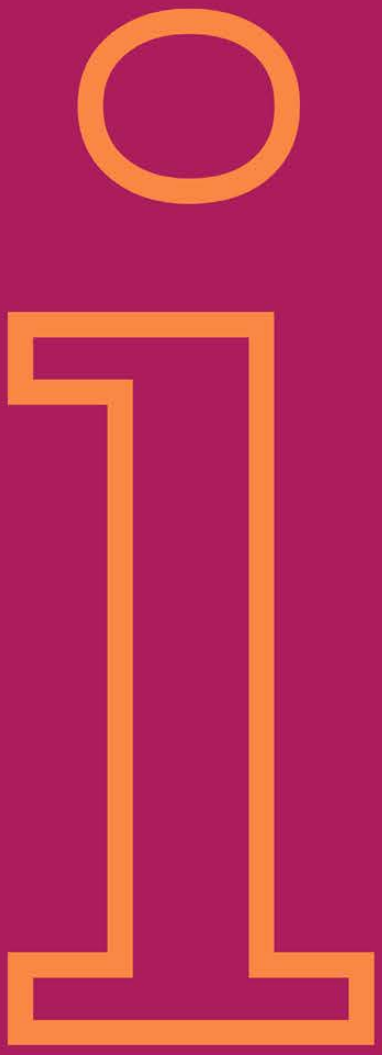
- Methods for analysing climate risks in the risk management of financial institutions.
- Methods to combat money laundering.
- Methods to identify and prevent market manipulation.
- Methods for data-driven assessment of the results of monetary policy decisions.
- Methods for designing new risk management and financing instruments along complex physical supply chains.

These examples show that a large number of different stakeholders (financial institutions, companies from the real economy, supervisory authorities, central banks, stock exchanges, start-ups, etc.) will access the FBDC with a range of different remits. For some of the examples mentioned, there are requirements for real-time availability, performance and scalability, for instance, which do not exist to the same extent in other domains. These framework conditions lead to FBDC-specific requirements for data structures and the suitability of secure and sovereign exchange and access architectures. The necessity of different protection classes and access rights in the system and also the need to mirror legal issues that relate specifically to financial transaction data must be taken into account in the GAIA-X infrastructure and in the GAIA-X Identity Service.

Further domain-specific requirements include:

- Tool-based linking of various data sources, combined with the option for data to be loaded into the system flexibly in different input formats using different ETL tools and for different data to be described with a data catalogue.
- Storage in accordance with layer model: separation of storage and computing components that enable defined access to storage in a provisioning layer.
- The option of recording data history and data availability in the system. Data governance and regulatory requirements demand particular attention to data history and data availability in the system (with online and offline access/retention periods).
- Option for flexible aggregation and drill-down into the microdata.
- Use of AI for different purposes.
- Anonymisation components for sensitive data.
- Monitoring of data usage, which makes it possible to identify who accessed which data at which time.
- Use of the data in real time with relevant implications for scalability and performance.

6. Further information



Invitation to participate

We also apply the GAIA-X project's guiding principle of *openness and transparency* to involvement in the project and the *User Ecosystems and Requirements workstream.*

All interested parties are invited to participate in the workstream, by submitting new use cases (1) and through active and ongoing involvement in a domain-specific working group (2).

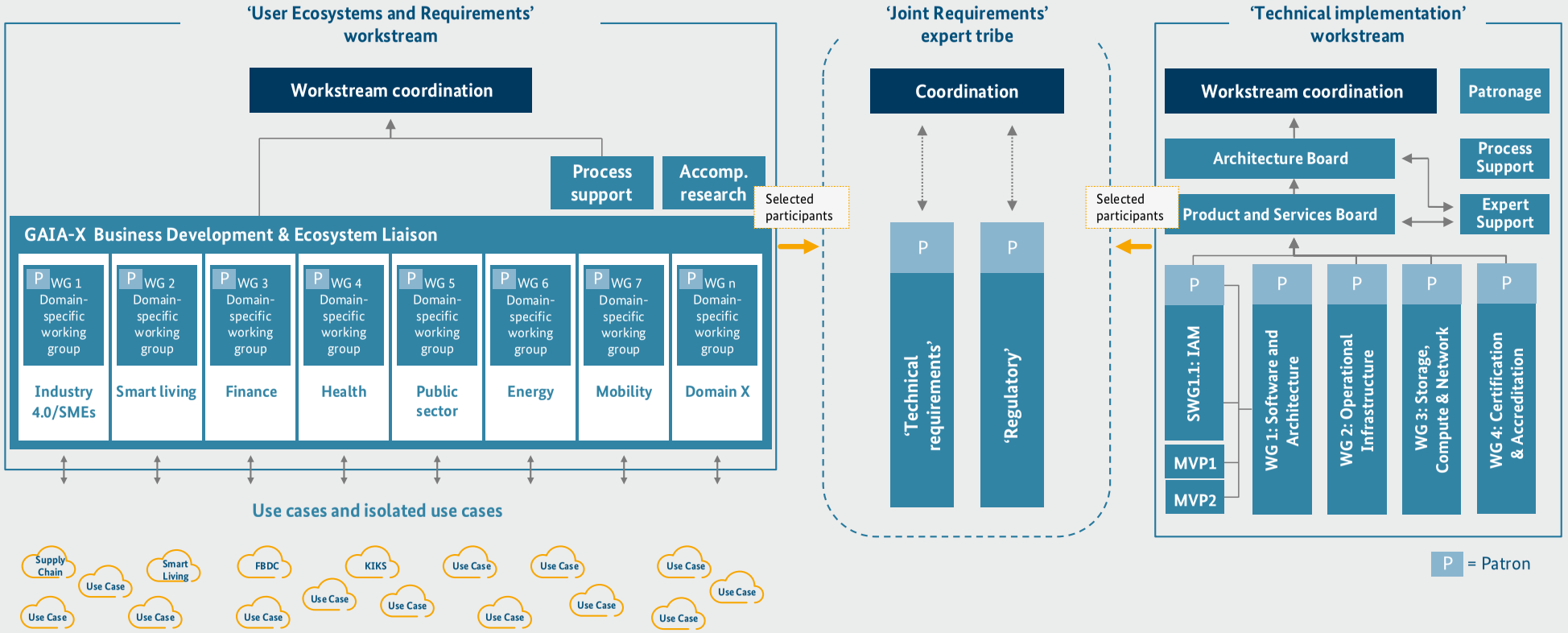
(1) If you would like to provide a use case for the (further) development of GAIA-X, please use the questionnaire via www.data-infrastructure.eu/gaia-x-use-case-questionnaire. We of course welcome further use cases from the aforementioned domains, but are also open to new examples from new fields of application. Use cases outside of the GAIA-X work structure may of course also be developed and implemented.

(2) The project is open to participation from interested European partners who wish to develop and expand the project with us. The aim is to transfer the data infrastructure into a live system. We welcome interested parties with specialist expertise who proactively support our work in the individual working groups and share our values and aims of data sovereignty and data availability.

Linking with European initiatives and beyond Europe's borders

There is already a large variety of initiatives dealing with the creation of data ecosystems in Europe and the Member States. GAIA-X wishes to link up with these initiatives and help to tap into the great potential and expertise in Europe. A large number of international companies and organisations have already been successfully integrated into the GAIA-X project. In the 'User Ecosystems and Requirements' workstream, we want to drive forward the efforts to link up with other European initiatives even further. In the next phase of the project, it is envisaged that this will also lead to GAIA-X hubs being established in the participating Member States, as shown in Figure 8. These hub structures are intended to help the GAIA-X project to become established in all EU Member States and beyond EU borders and to coordinate cross-border projects. The aim is to build a strong European network that will accelerate the distribution of GAIA-X at regional, national and European level.

Figure 8: Structure of the GAIA-X hub. The project structures of GAIA-X are flexible and support agile collaboration and overarching exchange



Source: BMWi

7. Contributors

The status report of Workstream 1 “User Ecosystems and Requirements” was developed by the following persons:

- Gerd Hoppe (Beckhoff Automation GmbH & Co. KG)
- Peter Kraemer (acatech – National Academy of Science and Engineering)
- Fabian Biegel (SAP SE)
- Prof. Susanne Boll (University of Oldenburg | OFFIS Institute for Information Technology)
- Dr. Stephan Bredt (Ministry of Economics, Energy, Transport and Housing Hessen)
- Prof. Dr. Klemens Budde (Charité – Universitätsmedizin Berlin and Germany’s Platform for Artificial Intelligence)
- Prof. Dr.-Ing. Gerd Buziek (Esri Deutschland GmbH)
- Susanne Dehmel (Bitkom e.V.)
- Prof. Dr. Roland Eils (Charité – Universitätsmedizin Berlin)
- Thomas Feld (Strategion GmbH)
- Dr. Friedrich Gröteke (Federal Ministry for Economic Affairs and Energy)
- Michael Jochem (Robert Bosch GmbH)
- Prof. Dr. Michael Laskowski (innogy SE)
- Dr. Christina Schmidt-Holtmann (Federal Ministry for Economic Affairs and Energy)

and under the contribution of:

- Daniel Abbou (KI Bundesverband e.V.)
- Sven Ahlfeld (COUNT+CARE GmbH & Co. KG)
- Craig Atkinson (Roche Pharma AG)
- Robin Barclay (Roche Pharma AG)
- Álvaro Barrón López-Para (GMS Management Solutions Germany GmbH)
- Jochen Bauer (Friedrich-Alexander-Universität Erlangen-Nürnberg)
- Ingo Beimdiek (ROSEN Technology and Research Center GmbH)
- Beko Deutschland GmbH
- Yassin Bendjebbour (BDEW Federal Association of the Energy and Water Industry)
- Jeanine Berg (GDV Dienstleistungs-GmbH)
- Maria Christina Bienek (IoTOS GmbH)
- Finn Brüning (e.V.)
- Association of German Banks e.V.
- Majlinda Caci (DEW21 – Dortmunder Energie- und Wasserversorgung GmbH)
- Geschäftsführer José David da Torre Suárez (COUNT+CARE GmbH & Co. KG)
- DATEV eG
- Dataport AöR
- German Research Center for Artificial Intelligence (DFKI)
- Dr. Björn Dietrich (Stromnetz Hamburg GmbH)

- Dr. Julian Dörr (DIE FAMILIENUNTERNEHMER)
- Jürgen Eils (Charité)
- Andreas Eisenreich (T-Systems International GmbH)
- Dr. Thomas Esch (DLR – The German Aerospace Center e.V.)
- Jochen Fehse (Robert Bosch GmbH)
- Thomas Fischedick (KEO GmbH)
- Lars Francke (OpenCore GmbH & Co. KG)
- Daniel Frank (Gasnetz Hamburg GmbH)
- General Association of the German Insurance Industry e.V. (GDV)
- Anne Grondke (GDV Dienstleistungs-GmbH)
- Dr. Jörg Habetha (Philips)
- Uwe Hartmann (Software AG)
- Dr.-Ing. Dr. med. Karsten Hiltawsky (Drägerwerk AG & Co. KGaA)
- Dennis Hoppe (German Association of Towns and Municipalities and High Performance Computing Center Stuttgart (HLRS))
- Hospital Clínic Barcelona
- Prof. Dr. Peter Hufnagl (Charité – Universitätsmedizin Berlin)
- Haimo Huhle (ZVEI – Germanys Electrical Industry e.V.)
- Anke Hüneburg (ZVEI – Germanys Electrical Industry e.V.)
- Dr. Philipp Huppertz (Power System Insight UG)
- Dr. Claudia Ivascu (Roche Pharma AG)
- CIO Harald Joos (Deutsche Rentenversicherung Bund)
- Jürgen Jost (Philips GmbH Market DACH)
- Dr. Markus Junginger (OHB SE)
- Kai Kalusa (VDMA e.V.)
- Dr. Elsa Andrea Kirchner (DFKI GmbH)
- Lukas Klingholz (Bitkom e.V.)
- Dr. Bastian Koller (Municipalities and High Performance Computing Center Stuttgart (HLRS))
- Henrik Kortum (Strategion GmbH / German Research Center for Artificial Intelligence)
- Gunther Koschnick (ZVEI – Germanys Electrical Industry e.V.)
- Prof. Dr. Frank Köster (DLR – German Aerospace Center e.V.)
- Prof. Dr. Ingo Kunold (Institute for Communication Technology FH Dortmund)
- Stefanie Lang (Fujitsu Technology Solutions GmbH)
- Christian Lawerenz (Charité – Universitätsmedizin Berlin)
- Prof. Dr. Sebastian Lehnhoff (OFFIS e.V.)
- Sönke Liebau (OpenCore GmbH & Co. KG)
- Dr. Kai Lindow (Fraunhofer Society for the Promotion of Applied Research e.V.)
- Dr.-Ing. Frank Lochter (Ministry of Economics, Labour and Energy Brandenburg)

- Prof. Dr. Alexander Löser (Research Center Data Science, Beuth-Hochschule für Technik Berlin Universtity of Applied Sciences)
- Dr. Benoit Marchal (Roche Pharma AG)
- Jens Mazzega (DLR – German Aerospace Center e.V.)
- Dr. Sven Meister (Fraunhofer Institute of Software and System Technology (ISST))
- Dipl. Ing. Lars Nagel (International Data Spaces e.V.)
- Prof. Dr. Thomas Neumuth (Universität Leipzig, Innovation Center Computer Assisted Surgery)
- Klaus Ottradovetz (Atos SE)
- Stefan Pollmeier (ESR Pollmeier GmbH Servo-Drive Technology)
- Katharina Psiuk (innogy SE)
- Markus Quicken (SupplyOn AG)
- Dr. Sebastian Ritz (German Edge Cloud GmbH & Co. KG)
- Thomas Rössler (Stadtwerke Gießen AG)
- Marilen Ronczka (Power Plus Communication AG)
- Markus Schaal (Marquard & Bahls AG)
- Dr. Matthieu-P. Schapranow (Hasso-Plattner-Institute for Digital Engineering GmbH)
- Ingenieur Jan-Willem Scheijgrond (Royal Philips B.V.)
- Dr. Thomas Schmidt (acatech – National Academy of Science and Engineering and Germany's Platform for Artificial Intelligence)
- Simon Skaznik (COUNT+CARE GmbH & Co. KG)
- Dr. habil. Florian Thiel (Physikalisch-Technische Bundesanstalt (PTB))
- Frank Trautwein (RAYLYTIC GmbH)
- Dr. Marco Ulrich (ABB AG German Corporate Research Center)
- Dr.-Ing. Mathias Uslar (OFFIS – Institute for Computer Science)
- Verband kommunaler Unternehmen e.V.
- Bert Verdonck (Philips)
- Luke Voutta (Federal Association of Medium-Sized Businesses (BVMW), Unternehmerverband Deutschlands e.V.)
- René Walter (Federal Ministry of the Interior, Building and Community)
- Stefan Weingärtner (LandesCloud GmbH)
- Dr. Andreas Wierse (SICOS BW GmbH)
- Lisa Witte-Stremmel (Federal Ministry of Health)
- Dr. Johannes Winter (acatech – National Academy of Science and Engineering and Germany's Platform for Artificial Intelligence)
- Stephan Wrede (DB Systel GmbH)
- Julian Zeidler (DLR – German Aerospace Center e.V.)
- Dr. Jochen Zimmermann (easierLife GmbH)
- Ole Zipfel (Mainzer Stadtwerke AG)