Gaia-X Domain Geoinformation

Position Paper Version 1.0 2021

Contents

Mission and Goals of the data space

The development of the world economy and its industries - especially those in Europe, depend highly on the availability and accessibility of data. As in early studies on public sector information in Europe stated, more than 80% of data is somehow related to space. Spatial relationships of nearly every data type will lead to extensive leverage effects for economies and markets in different perspectives of sustainable development. Thus, it makes sense to develop and maintain a general basis, better: a network of geo platforms, geo services, and methodologies, to serve different industry demands on a technology open, common and harmonized geo data and geo information basis.

Europe is aware of the high impact of spatial data, information, and relationships: this is proven by the revision of the Open Data Directive of the European Commission, as most of the High Value data Categories (geospatial, earth observation and environment, meteorological, statistics, companies and company ownership, mobility) are clearly related to space.

During the past decades, the gathering, maintenance, and provision of geo information was achieved by several different spatial information systems, like cadaster, utility networks, roads, topography, environment, space, meteorology, water, shipping, defense, and so forth. Those systems serve lots of administrative processes in rural and urban planning, smart city development, property registration and ownerships, environmental monitoring, tourism, land surveying, and mapping.

Beneath the public Spatial Data Infrastructure (SDI) approaches, like the Infrastructure for Spatial Information in Europe (INSPIRE), the Data Information Access Services (DIAS nodes) of the Copernicus programme, and several open data portals on each administrative level (Europe, Federal, State, Local), geo information systems being operated within private sector industries, like real estate, insurance, architecture, engineering, construction, transportation, logistics, mobility, smart cities, energy, and utilities, too.

To build and support geo related infrastructures, a worldwide geo spatial industry has emerged over the past 50 years. It is based on sectors like positioning, geo information systems (GIS), spatial analytics, earth observation sensors, and 3D. Due to a study on Copernicus by the European Commission from 2018, the overall market potential for the EMEA region was expected by 5.016 billion USD in 2020 with an average growth rate of 9.33%. Having this in mind, Gaia-X geoinformation domain's mission is to follow a technology open geo based digital tin approach. Intention is to unlock the power of spatial relationships through geo spatial enabling technologies by integration in value adding business processes. The characteristics of Gaia-X, known as trust, governance, data sovereignty, data portability, and openness will lead to additional value within geo data dissemination, provision, integration, and usage.

From a visionary point of view, a "horizontal", and value driven network of interconnected, private, and public geo platforms serve the demands of other "vertical" Gaia-X domains based on Gaia-X principles of value, trust, interoperability, transparency, and openness. The overall benefit and leverage effect emerge through a "geoinformation spillover" into other Gaia-X domains. This "cross domain approach" fosters geo information sharing, creates business

value, enforce crowd sourcing, collaboration, innovation, and makes the application of spatial technologies more economical and efficient.

Challenges addressed

As stated above, the European geo spatial market is a developed market driven by user demands, private cloud infrastructures, public spatial data infrastructures, public open geo data, free and open earth observation data, commercial products, open-source components, and the societal demands to create a sustainable future on the basis sharing and collaboration principles. This leads to more understanding of the natural and built environment to precede smart actions for an increased sustainability.

Following challenges are foreseen

- to reach out Gaia-X principles into a widely developed geo spatial market
- to address the geospatial market and identify the business value of Gaia-X principles
- to align successful business models of the private industry and public sector developments with Gaia-X architecture, values, and governance principles
- to adapt commercial software products to Gaia-X and vice versa to make specific public information systems (e. g. cadaster, high resolution images) accessible with Gaia-X principles
- to implement use cases without having Gaia-X core services available
- to agree with the domain industries and communities on common governance principles, and implement those
- to consider and develop current relevant standards and initiatives, such as OGC, ISO, GDI-DE and INSPIRE within a modern and future oriented Gaia-X based standardization process
- to develop interoperable data models, which allows the spatial data providers on local, regional, and national level to be integrated into the specific Gaia-X services via standardized spatial data services
- to have billing/pricing principles available to sell commercial data and services into/out of Gaia-X
- to have early available a set of generic Gaia-X services to foster the development of applications and solutions (e.g. administration, industries, Infrastructure as a Service approach)
- to enforce the availability of information as a service, and actionable data/Information/services generation (value adding services to generate ready-to-use products and solutions)
- to implement a clear und transparent lifecycle management (research & development & economy/productization)
- to make high valuable datasets widely accessible

- to have a testbed/sandbox environment available early to build PoC or show case studies for evidence-based standardization/processing or using federated catalogs such as tools, data, or services

Solution: Data space description in a holistic view – detailed view on the endeavour

Partners of the ecosystem

- GIS Software vendors
- Building Information Modeling (BIM) Software vendors
- Private and public geo data providers
- Platform and service providers, PaaS, SaaS providers
- Governmental agencies, public sector; including cadastral and land management information/systems
- Private Sector/Private companies, professional service companies, system integrators
- All levels of Public administration (European, National, State, Local level)
- Citizens (to benefit from open geo data and end-to-end services)
- Researchers, Space administration, SDI initiatives
- Other domains and data spaces
- NGO (disaster and risk management, World Bank/Development Organizations, UN) using data with global dissemination
- International institutions (Development Banks, EEA, WHO, WFP ….)

Use-cases (scenarios) within the data space and the current state of development

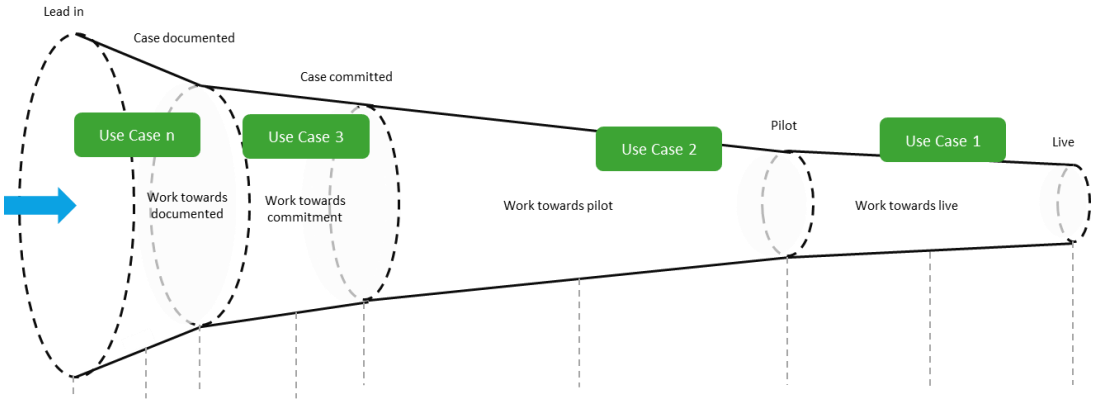

Figure 1

Use cases play an important role in the domain geoinformation. Since the onboarding process of domain members is ongoing, new use cases will appear continuously over time. However, not all of them will succeed as this depends on commitment and voluntary engagement at the beginning. Since business models must be developed, and the reach out to customers/users, too, some funding has to pave the way in the second half of the shown "Use Case Funnel". Successfully implemented cases will have a "going live" to show and point out the domain geoinformation itself and cross domain benefits. As geoinformation is not abstract data and can be depicted very vivid and clear on digital and interactive maps, it is an excellent basis for public Gaia-X value communication.

Description Use Case 1 – Space4Cities

Solution

Space4Cities provides solutions for the joint exploitation of earth observation, machine learning and modern information and communication technologies to provide decision makers from science, economy and policy with precise, up-to-date and ready-to-use information on the highly dynamic built environment – from local to global scale.

Problem solved

Cities are one of the most rapidly evolving systems on Earth. Hence, effective management and development of the built environment requires detailed, timely and holistic knowledge about the status and changes of the urban system. At the same time the digital transformation taking place in all areas of life has led to a massive increase of digital data – in particular, related to the places where and the ways how we live. It is here, where the enabling technologies developed in Space4Cities allow users to effectively access, process and analyse the petabytes of available big city data in order to gain innovative knowledge and insights that help making the human habitat more sustainable, livable and resilient. At the same time mechanisms for the integration of citizen sciences, crowdsourcing and open knowledge foster the inclusive, creative, and transparent character of the applications and services of Space4Cities.

Partners/Ecosystem

The Space4Cities activity includes and serves a broad spectrum of partners and users from science, urban and environmental planning, economy, and policy. A particular focus lies on the active involvement of international institutions (e.g., United Nations, World Bank, Asian Development Bank) in order to substantially support the Sustainable Development Goals, contribute to the monitoring of global change, and improve disaster and risk management.

Main technology/Gaia-X components

The main technological components of Space4Cities and related Gaia-X instances include instruments for big data collection, structuring and processing, earth observation, and smart

data analytics. The related technical solutions cover the automated information extraction from large-volume and heterogeneous data sets (big earth data) by means of machine and deep learning, empowered by distributed computing infrastructures for high performance processing. These approaches are supplemented by modern information and communication technologies, which allow the autonomous control of data access, management, and processing (automation). Key components of the Space4Cities framework have successfully been applied in the global urban monitoring initiatives of the Global Urban Footprint (GUF) and the World Settlement Footprint (WSF), whose information products are already used by over 600 renowned institutions worldwide.

Concrete benefits

The value of the Space4Cities solutions lies in the synergistic evaluation of databases of earth observation, open data, and public administration to break down data silos and offer - GDPR-compliant -tailored information products for city planning and development and for new digital products and business models (e.g. in the area of sharing economy or local public transport). Ultimately, Space4Cities is supposed to help improving the quality of life for citizens and save costs by making processes and decision more effective and sustainable. Therefore, Space4Cities finally supports cities in stepping toward a digital metropolis that is fit for the future.

Description Use Case 2 – Smart Infrastructure Management

Solution

Gaia-X enables the early engagement of organization as well as the fusion of various data sources in a trustful data sharing ecosystem towards one information base in order to plan, execute and run complex Architectural, Engineering or Construction (AEC) projects in an interdisciplinary, collaborative, and timely manner. Additionally, Gaia-X enables the provision of a crowd participation service for the affected citizens to ensure early community involvement.

Problem solved

Planning and approval procedures have reached a very high degree of complexity in the last two centuries, making it very difficult for decision-makers in planning offices, authorities, associations, and courts to master these procedures. Moreover, construction planning needs to consider a large number of interests and stakeholders, including for example property owners, transport ministries, land surveying offices, environmental protection agencies and the general public.

Currently, the various stakeholders are engaged successively. This is partly due to the different restrictions in place regarding the purpose and use of data, as well as questions of competence and partly due to isolated approaches that make it impossible to merge relevant data. At the moment, there is no uniform platform for collaboration and the secure provision of data with

all the necessary functionalities and rights to access and use data. As a result, processes are mostly limited to simple data provision, leaving the great potential of real-time cooperation untapped, including advantages in terms of flexible and targeted data use. In addition, objections from communities and associations often go unheeded in the planning phase, leading to later changes and additional project costs for the planners.

Partners/Ecosystem

The use case combines several stakeholders such as road construction authorities, planning companies, responsible developers for instance bridge structures, Ministry of Transport, Environmental Protection organizations, surveying offices, affected municipalities and citizens, providers of critical infrastructure (e.g. energy suppliers), SMEs and the general public.

Main technology/Gaia-X components

This use case impressively demonstrates that an ecosystem of central services based on the Gaia-X Cloud can be established to enable secure and user-friendly operation for relevant use cases. The use case shows that the secure provision of data always includes the necessary functionality to process this data in a relevant business process. A common ecosystem of users and providers is created. Therefore, component such as Catalogue, Identity Access Management, and control of digital identities are crucial.

Concrete benefits

The 'Smart Infrastructure Management' use case lays the foundation for better interdisciplinary and user- friendly cooperation in planning beyond the boundaries of organizations and authorities – while ensuring their self-determination and sovereignty. According to modularity and interoperability, Gaia-X ensures the continuous availability of various data sources as well as integration of additional existing solutions or data bases. Based on the digital data infrastructure Gaia-X, which ensures data exchange across national boundaries, this use case accelerates the procedures and quality of project planning – and therefore contributes to the European added value.

Description Use Case 3D Planning Platform

Solution

The use case describes a solution for using a digital twin of a city as the backbone for planning projects and processes. Projects can be visualized, observed, and checked throughout their entire life cycle. Planned developments can be examined within a 3D model under the existing real conditions in the respective city. It is also possible to streamline the process of creating and sharing 3D zoning and land use plans. The use case enables citizens, real estate developers and other community stakeholders to participate in the planning and development process.

Problem solved

The planning, information and participation process is digitized and made more efficient thanks to a common 3D planning platform, the Digital Twin. Existing 3D context data from the surveying offices form the basis of the planning platform and can be used for this and many other applications, resulting in benefits for city and planning offices, regional planning authorities and the population. Cities and states get easy access to the relevant data and software to guarantee a sustainable and citizen focused development.

The planning and development processes are open for various stakeholders (government, public, research, AEC, real estate). In this way, the community can be asked for feedback at an early stage. Stakeholders can use the Smart Urban Planning platform to access a web-based view of all plans and projects if they are publicly approved.

Partners/Ecosystem

The use case allows various stakeholders (government, public, research, AEC, real estate) to participate in planning and development processes via a digital twin of a city.

Main technology/Gaia-X components

Gaia-X provides the necessary infrastructure for the integration of the Digital Twin and its data space, ensures secure data exchange, and allows data to be made available without media disruption. The solution is offered as SaaS and considers the pursuit of modularity and interoperability. Specific software (e.g. ArcGIS Enterprise) is run on Gaia-X for access to the 3D models or the digital twin services. Gaia-X is also used for the hosting of the 3D services.

Concrete benefits

Gaia-X provides a fast and secure environment to host 3D services and to run user specific (planning) software. Gaia-X offers stakeholders like cities or states a user-friendly and easy access to the relevant data and software. According to the concerns about political data protection regulations and preferences Gaia-X provides the planning processes based on legal foundations such as land use and development plans. The calculation and comparison of future scenarios and their impact on capacity indicators (jobs, energy consumption, etc.) is a time-consuming task and requires considerable expertise. The Digital Twin serves not only for planning purposes but for many processes in smart cities (e.g., AI, IoT, green infrastructure).

Description Use Case Automatic 3D Spatial Content Generation

Solution

The use case describes the automation of processes to create and update 3D surfaces and features in high detail and accuracy with methods of artificial intelligence.

Problem solved

The concepts and technologies of Open Data Science are utilized to generate or update 3D surfaces and features from geospatial incoming raw data automatically. Additionally, market available software products are used to prepare the training data and to handle the incoming geospatial raw data. It exists a trained algorithm of artificial intelligence to do the job based on various incoming data sources, e.g., aerial imagery or LIDAR point clouds.

Partners/Ecosystem

The use case is aimed at companies, SME, research institutes. Potential end users are citizens, Industry 4.0, smart cities, agencies, security organizations - for planning and operational purposes.

Main technology/Gaia-X components

Along the chain of finishing the training of algorithm of artificial intelligence, a lot of cloud resources are essential. Therefore, Gaia-X will put the safe usage of artificial intelligence over the edge and provides a certified environment of trust. The intellectual property of the labeled training data and the final algorithm must be protected. Gaia-X will enable the protection. In addition to authenticity and trust, components such identity access management are crucial.

Concrete benefits

Gaia-X enables reducing costs of creating and updating 3D content and improves the frequency of such data updates. In addition, it is possible to save worthful human resources to manage the information instead of having to produce it manually.

Description Use Case Earth Observation Federate Data Access

Solution

Earth Observation (EO) data federated access system is the solution for any stakeholder who wants to contribute or to use EO satellite data in complementary and direct way, online and secure from one comprehensive source. Proposed cloud-based system will be able to efficiently deliver data to Destination Earth tasks. Our system could also be a perfect solution for not only satellite but aerial, health image diagnostics, IoT or any other big data challenge and partners.

Problem Solved

Problem definition:

- Instant, online, and immediate accessibility of huge amount of EO data.

- Availability and easy use of enormous data sets – Copernicus and other EO data may be augmented with some in-situ data. Data sets of the size 50PB or more – a challenge, too.
- Online, immediate access to long time and space series of data for any cloud or HPC infrastructure in Europe
- Data easily discovered – for interactive and m2m applications
- Data tagged for potential AI applications

Rapidly growing amount of collected Earth Observation data should be provided to the users in predictable and reliable manner. The vast amount of public data is generated in an open access Copernicus - European Earth Observation programme. Nature of the data sources ecosystem consist of growing time series of freely available satellite products (e.g. Sentinel), increasing number of data-based services (e.g. Copernicus Services) and of the commercially available VHR data. Successful use of this wealth of information will depend on the possibility to access them in a fast and reliable way. It is as well to be expected that the number of the sources (data repositories) to be consulted in each project will grow with time even if some data sources will federate. Similarly, the size of individual repositories will increase. Already full Sentinel Data repository takes well over 30 PBytes of data with daily growth exceeding 25 TBytes per day and new missions are planned. It is thus necessary to design robust, open source solution standards allowing data providers (data repositories) and data users to 'meet' and to process the data in cloud environment. It is a challenge itself to manage 50 PBytes or even more of heterogenous data (EO, in-situ and statistic) nevertheless this use case/solution can solve this enormous challenge.

EO data federated access solution includes and serves a broad spectrum of partners and users from commercial, scientific, and administrative worlds.

We propose to build cloud-based federated data access service using API and interoperability standards to provide access to a very large amount of EO data from various sources while keeping data traceability and accounting but protecting user rights. In particular, developing and testing definition semantic metadata standards for AI data labelling are needed in order to include AI-ready data stores. Project will include specialized tools for data providers to make their collections discoverable.

Partners/Ecosystem

Partner ecosystem would consist of data and information providers, cloud operators, data users and information end users including VHR data providers, specialized associates for data processing algorithms development, science and research institutions, federation partners. Since Earth Observation and Services usually address specific challenges (geohazard, water or coast monitoring) the project will involve already existing ecosystems focused on their specific area of interest.

Main technology/Gaia-X components

The project would build on the already existing European solutions such as:

- Open source based public clouds
- Open source based Big Data storage
- Existing Open data access standards
- Big Earth Observation data repositories (e.g. CREODIAS with over 21 PBytes of the data, WEkEO, other DIAS-es), other data and information sources
- PaaS solutions – data discovery, on-demand processing
- AI tagging – open source plus necessary developments

Challenge – scalability of storage, efficiency of data access (more than 2 PB daily delivered), federation of existing data sources and development of PaaS elements (in terms of new functions and increased efficiency). Offered data will be consolidated, indexed, tagged, and prepared for potential pre-processing, making the service adjustable to any need. The main Gaia-X components to be used are the Federated service catalog, Identity management and Standards of interoperability. The solution while secure and robust should be flexible and open for evolution.

Concrete benefits

Federated, stable, comprehensive, efficient, and easy access to Earth Observation data for instant use and/or further processing for every customer, maximizing the use of available data on European platforms are the main benefits. Fully controlled European solution with European competencies and infrastructure, located and managed only in Europe and by Europeans harmonizes and constitutes with the main assumptions of Gaia-X.

Added value of the proposed system comes from the limiting carbon footprint by de-duplication of data and keeping it instead of repetitive generation on-demand. In this way maximized use of existing resources will be achieved (no need to copy or reprocess data to make them usable). Elasticity and multitasking - possibility to extend the model to other fields (like telemedicine data) is another example of potential positive impact of our System on environment as well as a community and business.

Thanks to this project Earth Observation data will become a part of an open, transparent digital ecosystem, where data and services can be made available, collated, and shared in an environment of trust.

Maturity indication of the data space, health status

How is the demand side represented?

- by other domains data spaces
- through domain members from public and private institutions

- by customers of private companies
- by representatives from technical and scientific organizations
- European citizens, participation processes, transparency of political decisions if spatial relevant
- through shared knowledge generation as basis for innovation, business
- by decision makers, politicians, and delegates from crisis management centers

How is the supply side represented?

- Software vendors (e.g., GIS, BIM)
- Platform operators and providers
- Public geo data agencies
- Private geo data providers
- Developers
- Startups
- Crowd sourcing communities (business perspective)

Is there an equal representation of demand and supply side to provide a sustainable business model?

- Yes
- Open business models; usage tracking; buy versus use on demand
- Enhancement of business models

Is the story of the data space well documented?

Data space members didn't agree on a common story yet. However, the fundamentals of a storyline are outlined in section 1. As every information is somehow related to space, the geo data space is a common data space that enables business value in other data spaces through spatial referencing by a network of interoperable platforms, data sharing, and collaboration. As most of today's challenges towards a sustainable future needs several data from different industries (big data paradigm), the geo data space provides a common and harmonized spatial reference. On one hand side application neutral data and services is provided (maps, terrain, land use, 3D city models, satellite images etc.), and on the other hand, several spatial methods achieve spatial intelligence to lots of domain specific data. In other words, the geo data space is beneath Gaia-X itself another common nominator of Gaia-X domains to foster efficiency and innovation.

What is the business model and the business mechanics of the data space after the PoC implementation?

Gaia-X provides federated cloud architecture components on an open basis, and common governance principles. Following the rules, Gaia-X geospatial platforms are interconnected and serve as a “platform of platforms” (geo based digital twin approach) other Gaia-X domains with geo spatial data, information, relationships, and location intelligence. In terms of a business model the geo data space creates value through spatial relations and location intelligence to other domains without having extensive geo knowledge by themselves. Interconnected identity management, open API and interfaces, and a system of systems lead to better insights into data, and smart intelligent decisions.

Which components will be certified according to the Gaia-X federation services?

It is worth to mention, that the use of data sets, information layers and intelligent geo spatial functionality depends on the availability of geospatial technology and geo data. The Gaia-X federation services will lead to certified interfaces and APIs to feed data and services into a decentralized, federated and independent catalog that is maintained by many involved actors of Gaia-X. Together with a common and cross domain identity management, and standards of interoperability, this achieves the basis for data sharing, collaboration, billing, and pricing.

What is the potential for adoption of the endeavor and for further scaling?

As geoinformation is a fully developed industry, the adoption to Gaia-X depends on additional value created by Gaia-X. This business value might result from Gaia-X characteristics as far as others did not cover them with their own business models. To scale up, value, benefits and incentives might be necessary to involve software vendors, data, and IT infrastructure providers to adopt to Gaia-X and roll out products with Gaia-X compliant interfaces and APIs.

How can the commitment of the parties involved be proven?

The commitment of involved parties depends highly on a successful governance of Gaia-X. If Gaia-X services and features are required through procurement processes within public sector, public administration and public owned companies, the commitment could be checked by existing Gaia-X compliant IT products. If there is no further “real” market pull, Gaia-X would just be nice to have, and not a must.

Are sufficient resources available to realize the endeavor according to its mission?

To awake the spirit of Gaia-X by resources for further implementation, innovation and business value must go hand in hand to tackle the markets. Resources of companies will be available if

differences to existing solutions and offerings would bring additional business value. This also depends on further funding and tenders.

Evolution of the data space

Roadmap of the evolution

- Onboarding of platform providers and software vendors (GIS, BIM etc.)
- Interconnect with Gaia-X federated services (catalog, identity)
- Agreement to common governance rules
- Implementation of use cases
- Marketing and business development

Quick wins (for 2021)

- Find geoinformation services by a federated Gaia-X catalog service
- PoC to access an identity controlled geoinformation service by “n” other domain/data space
- New innovative capabilities to produce, update and share high quality 3D x/y/z and spatiotemporal x/y/t and x/y/z/t geoinformation in a controllable manner
- Interdisciplinary fusion of relevant geoinformation towards a more realistic digital twin as a basis for BIM or other trustful interconnected processes

Mid-term benefits (2022-2023) building on already-launched or soon-to-be-launched projects

- Running a data hub on basis of federated platforms and identity access management
- Improvement of usability for open data proposals
- Deeper integration of geoinformation and appropriate functionality to manage them into the valued adding process chains of the ecosystem
- Accessibility of non-profit and profit geoinformation for training of AI algorithm in a trust-fair environment
- Availability of a deep, high quality geoinformation database to discover and research phenomena of climate change even better
- Improvement of environmental and biodiversity protection by a better and interdisciplinary accessible data base of geoinformation of high scale
- Connecting with relevant European Spatial Data Infrastructure activities, such as INSPIRE, EarthServer, and DIAS nodes (on European and national level),

Long-term benefits requiring significant investments on the 2021-2025 period

- Gaia-X acting as a living marketplace of data services

- Access to Gaia-X based storage and computing capabilities
- Deployment of geoinformation out of Gaia-X connected platforms as services
- Effective integration of geoinformation into all value-adding processes to achieve a better exploitation of the value of geoinformation by trust-share and digital business models

Actions to be taken and recommendations for industry, politics, and society

- Align public geo spatial data infrastructure with Gaia-X
- Provide geoinformation out of Gaia-X
- Implement technology open value chains for successful market impact
- Provide tools and methods to combine, analyse, visualize geo-based information
- Follow on open data initiatives
- Simplify existing initiative in a more user focused manner (usability)
- Implement government and platform interconnections to exploit the value of existing geoinformation as well as to improve a high-quality basis of accurate and current geoinformation
- Enable trustful-sharing and real-time collaboration for and with geoinformation
- Enrich 4D digital twin initiatives with geoinformation
- Foster relevant standardization process towards open standards in the domain of geoinformation
- Foster business opportunities on all levels, including Public-Private Partnerships