TRAINING MANUAL

Geospatial Knowledge Infrastructure for National Development

10-12 May 2024, Rotterdam, The Netherlands

PRODUCED BY

GEOSPATIAL WORLD
ADVANCING KNOWLEDGE FOR SUSTAINABILITY

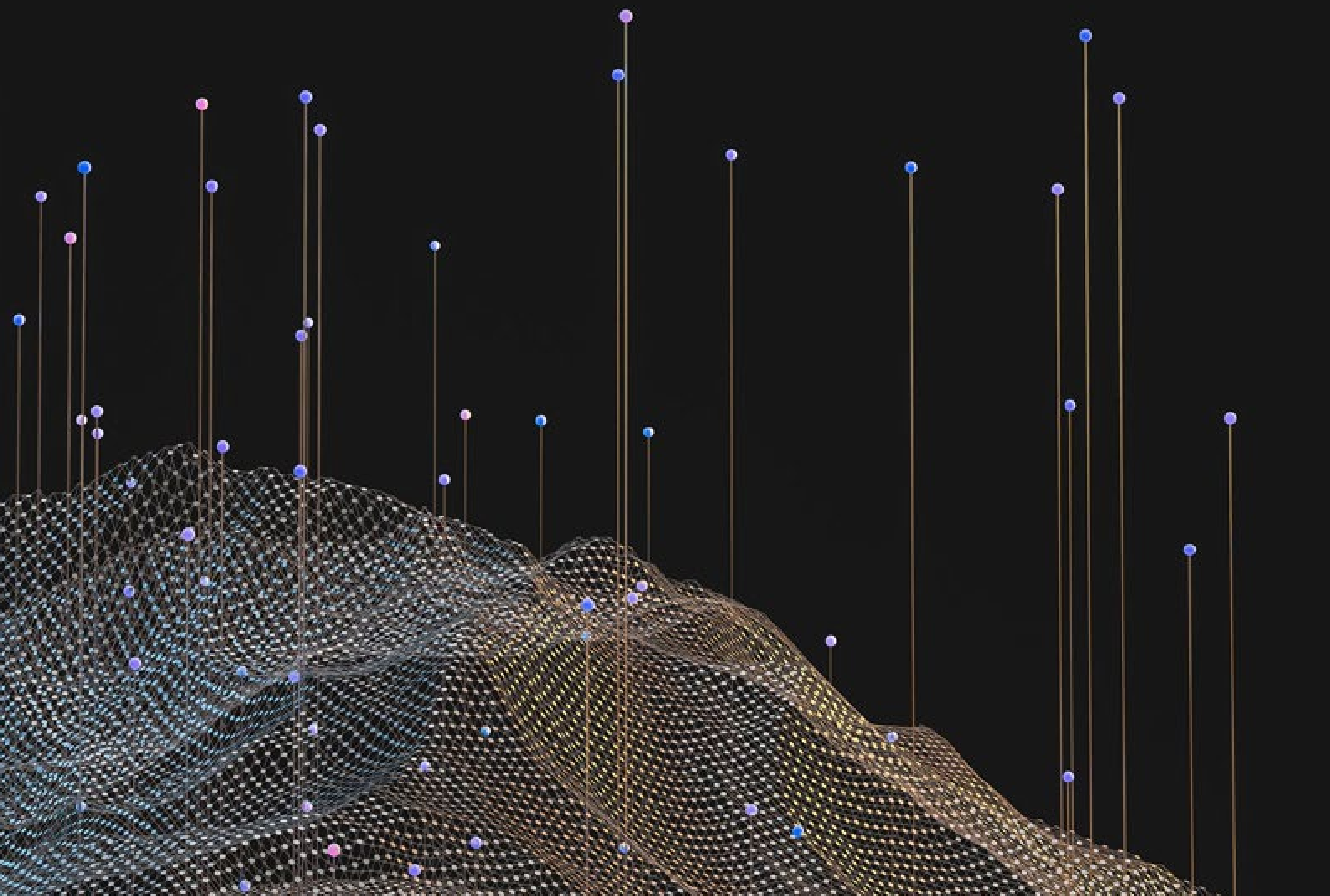


TABLE OF CONTENT

Training Introduction

The future lies in adopting the geospatial way, and it is the most efficient and effective way to collect, process, integrate, and utilize information for overall global, national, and individual development. But, concentrating only on data would not suffice, and it is imperative to move up the value chain from data to knowledge through Geospatial Knowledge Infrastructure (GKI). GKI embodies a concept that champions the evolution of a nation's geospatial ecosystem along the value chain towards knowledge, diverging from the current emphasis on data infrastructures towards the establishment of knowledge infrastructures. National geospatial infrastructures, oriented towards knowledge, serve as structures empowering countries to ascend the knowledge value chain. These infrastructures are essential for national development, and they require partnerships, knowledge sharing, and capacity building to be successful. Countries can adopt it by investing in developing a positive and collaborative approach, including government to government, government to private, government to funding agencies, academia, and research institutes.

The adoption and integration of geospatial knowledge in critical economic sectors is crucial for national development. GKI plays a significant role in this regard by integrating geospatial approaches, data, and technologies to deliver location-based knowledge, services, and solutions. By incorporating geospatial knowledge into user workflows, the potential impact on sectoral growth and development becomes evident.

The three-day training programme on **'Geospatial Knowledge Infrastructure (GKI) for National Development'** is designed to address crucial questions on the growth trajectory of geospatial knowledge, the influence of new-age technology ecosystems, including AI, Big Data Analytics, Cloud Computing, Robotics and Drones on user segments, and the role and relevance of the paradigm shift from data to knowledge in national development. It also intends to help NGAs to improve their data management capabilities by enhancing their understanding of geospatial infrastructure, promoting data sharing and interoperability, increasing efficiency, enabling better decision-making, fostering collaboration and cooperation, and adapting to new technologies.

Training Agenda

Time	Lesson	Topic
Day 1: 10th May 2024		
0900 – 0930	Welcome and Introduction ***Sanjay Kumar,*** CEO, Geospatial World	
0930 – 1045	Lesson – 1	**Relevance and importance of Geospatial Knowledge Infrastructure (GKI)** *Trainer:* ***Ruban Jacob,*** Associate Director - GKI, Geospatial World
1045 – 1100	Coffee Break - 1	
1100 – 1300	Lesson – 2	**The socio-economic impact of Geospatial Knowledge Infrastructure (GKI)** *Trainer:* ***Ms. Ananya Narain,*** Vice President - Consultancy, Geospatial World
1300 – 1400	Lunch Break	
1400 – 1600	Lesson – 3	**Live data streams in geospatial knowledge infrastructure** *Trainer:* ***Prof. Josef Strobl,*** Department of Geoinformatics, University of Salzburg
1600 – 1630	Coffee Break - 2	
1630 – 1730	Lesson – 4	**National Geospatial Portal** *Trainer:* ***Mr. Alan Devenish,*** Solution Engineer, Esri
Day 2: 11th May 2024		
0930 – 1130	Lesson – 5	**Data Access and Interoperability** *Trainer:* ***Dr. Joan Maso Pau,*** OGC Member
1130 – 1200	Coffee Break - 1	
1200 – 1400	Lesson – 6	**Partnerships on Spatial Social Challenges** *Trainer:* ***Dr. Rosemarie Mijlhoff*** – Portier, Geonovum ***Gen. Girish Kumar,*** Former Surveyor General of India, Survey of India
1400 – 1500	Lunch Break	
1500 – 1700	Lesson – 7	**Leveraging Latest Data and GeoAI Technologies for National Mapping** *Trainer:* ***Dr. Kumar Navulur,*** Founder & CEO, GeoCognition
Day 3: 12th May 2024		
0930 – 1030	Lesson – 8	**Fundamental Geospatial Data Themes** *Trainer:* ***Dr. Zaffar Sadiq Mohamed Ghouse,*** Director – Advisory & Innovation, Woolpert – Asia Pacific
1030 – 1100	Coffee Break - 1	
1100 – 1300	Lesson – 9	**Cloud Computing** *Trainer:* ***Dr. Siva Ravada,*** Vice President of Development, Oracle
1300 – 1400	Lunch Break	
1400 – 1600	Lesson – 10	**Geospatial Infrastructure for AEC Industry** *Trainer:* ***Prof. John Wilson,*** Founding Director, Spatial Sciences Institute, university of Southern California
1600 – 1630	Coffee Break - 2	
1630 – 1730	Lesson – 11	**Strategy for Integrated Geospatial Knowledge for Digital Economy** *Trainer:* ***Dr. Mark C. Abrams,*** Chief Scientific Officer, Exquisite Geolocation Systems LLC
1730 - 1800	Guest Address and Certificate Distribution ***Mr. Ronald Bisio,*** Senior Vice President, Field Systems, Trimble Inc.	

Learning Objectives

To gain a greater understanding of:

- The evolving role of national geospatial agencies in digital age.
- The socio-economic impact of Geospatial Knowledge Infrastructure (GKI)
- The geospatial knowledge value-chain and and its role in sustainable development.
- The changing geospatial ecosystem and its convergence with the broader digital ecosystem.
- The growing business models for partnerships and collaborations among various stakeholders.
- The components of integrated geospatial strategies and their significance in national development.

Format of Session

Each session is divided into 4 segments, although presenters may adapt these according to the subject.

Presentation: It will be taken up by the trainers via ppt to present the concept/topic. A case study will be presented by an expert

Q&A: During the Q&A the floor will be open for trainees to ask questions and doubts to the trainers.

Individual Exercise: In this timeframe 5-6 questions (objective + subjective) questions will be asked to the trainees based on the lesson plan aligning it to the national perspective. The answers will be captured via an online platform.

Summing up of Key Points

Training Evaluation

At the end of each session an individual exercise will be undertaken to understand whether the participants have learned what they were supposed to learn. This will provide an understanding whether the learning objectives were achieved or not.

As the closing session a brief feedback exercise will be done to see if the overall expectation highlighted were met.

A short follow-up survey will be conducted within 3 months after the training program. This survey will provide course presenters with more detail about the effective use of the lessons learned.

LESSON – 1

Relevance and Importance of Geospatial Knowledge Infrastructure (GKI)

Trainer's Biodata: Ruban Jacob

Ruban Jacob is the Associate Director – GKI at Geospatial World. He has over 7 years of experience in the policy and governance domain, and around 5 years of experience in geospatial governance specifically. He has been part of the collaborative partner alliance project on GKI since its inception, and has been actively involved in the development of the GKI concept and its elements. He is also actively involved in the development of the GKI readiness index which ranks countries around the world based on the maturity of their national geospatial infrastructures. Through the GKI project and the integration of GKI concept in all activities under the aegis of Geospatial World, he is striving to expand geospatial beyond the traditional geospatial stakeholders, by reaching out to the digital and user domains and enhancing workflow integration to better contribute to sustainable national development.

1.1 Introduction

Humanity's insatiable thirst for knowledge propels us forward, driving advancement, fostering development, and fueling competition. Throughout history, maps have served as indispensable guides, facilitating exploration, facilitating trade, fortifying nations, and weaving the intricate social fabric of our civilizations. Yet, in today's interconnected world, characterized by vast troves of location data and sophisticated analytics, maps represent merely one facet of a burgeoning landscape of discovery and opportunity. This landscape, shaped by the proliferation of Big Data, including remote sensing, Artificial Intelligence (AI), automation, continuous connectivity, Web 3.0, and IoT sensors, is evolving rapidly, driven by innovation and machine-enabled disruption. This convergence, often referred to as the 4th Industrial Revolution (4IR), blurs the boundaries between human, physical, and digital realms, ushering in unprecedented societal transformations. These changes touch upon every aspect of existence, encompassing domains such as living, health, leisure, work, wealth distribution, industry, and the intricacies of our social fabric.

This is not a vision of the future; it is our current reality. Geospatial knowledge, powered by the widespread availability of location data through GNSS, is already transforming our world. However, we continue to grapple with geospatial concepts that remain largely two-dimensional in a four-dimensional reality, focused on products in a data-driven era, and providing information rather than fostering true understanding. Part of this challenge stems from the historical emphasis within the geospatial sector on data alone, despite the fact that data, in isolation, holds limited value. The GKI sought to broaden the discourse surrounding geospatial knowledge and spark discussions about the necessary evolution of conceptual frameworks within the field.

This is profoundly pertinent timing as well. With the global population at its highest levels and significant disparities persisting within and among nations, we face pressing challenges. The imminent threat of surpassing the 1.5°C target for global warming looms, with projections indicating a rise of 2°C to 3°C by the century's end, accompanied by irreversible and unpredictable consequences. Emerging from the grip of the most significant pandemic in a century underscores not only remarkable strides in medical science but also underscores a trend toward nations turning inward, indicative of heightened global competition in a year witnessing record-breaking participation in democratic elections. Moreover, while the likelihood remains low, the looming specter of war casts a shadow, with the potential for conflict more devastating than ever, propelled by the very technologies of the 4th Industrial Revolution.

4IR presents opportunities to address numerous longstanding challenges by leveraging enhanced knowledge. In developing countries, smartphones and their associated applications are empowering autorickshaw drivers to boost their earnings, while also assisting small-scale farmers in augmenting production and streamlining distribution channels. Geospatial knowledge is playing a pivotal role in diminishing investment needs for renewable energy initiatives and facilitating efficient power distribution for electric vehicle (EV) transport systems. Drones are poised to emerge as the forefront leaders in autonomy, surpassing cars, where human trust in technology still lags. Through the provision of reliable, contextually relevant on-demand knowledge, both people and machines can achieve levels of independence and interdependence previously unimaginable.

1.2 Evolving Geospatial Ecosystem

The geospatial landscape has witnessed rapid expansion over the past two decades. Companies specializing in remote sensing satellites now offer near-continuous surveillance capabilities, catalyzing the emergence of numerous startups that provide knowledge-centric services tailored to specific sectors, leveraging integrated data and robust tools. While Google remains the most widely recognized geospatial entity, a plethora of other companies have flourished in this domain. These entities transcend mere consumption of geospatial information and technology; they actively contribute knowledge and societal value. For instance, companies like TomTom utilize location data to aid human decision-making, while others such as Starship.com offer innovative services like autonomous robot-driven shopping delivery. What was once a traditional consumer of geospatial services has now transitioned into a fully integrated member of the geospatial ecosystem.

Simultaneously, the emergence of Generative AI and the abundance of data from various sources—land, sea, air, and space—across diverse spectra have disrupted the traditional role of government geospatial agencies, opening opportunities for companies not traditionally associated with the 20th-century geospatial ecosystem. While AI holds promise in generating high-quality geospatial data for various applications, it has yet to fully tap into geospatial data to provide instantaneous 'knowledge on demand'. Imagine the potential of such a tool if a citizen could simply inquire about the optimal property for their family, assess the insurance risk associated with it, and forecast the property's future value—all through a seamless interaction with AI.

These examples all demonstrate the evolving geospatial ecosystem. GKI is perhaps the framework to foster this ecosystem from a knowledge perspective.

National geospatial agencies have evolved as providers of geospatial information. The original output, maps; remain powerful visualisations of the real world, support development themselves and can be created far more effectively using 4IR technologies and GKI considerations. Spatial Data Infrastructures (SDIs) are established in some countries and communities. They are needed to make spatial data accessible, but they are inherently a data service in a knowledge world. GKI, though, is a direction of travel, not a "take-it or leave-it" infrastructure, with the vision to place **geospatial knowledge at the heart of tomorrow's global digital society**. It is written to help organisations, whether government, academic, industry or civil society, consider how they can generate exponential benefits by moving from data concepts to knowledge concepts and ensure geospatial information, technologies and people contribute to, and benefit from, the growing 4IR enabled knowledge society.

1.3 Geospatial Knowledge Infrastructure (GKI) concept and key elements

Geospatial Knowledge Infrastructure provides **a blueprint to integrate digital economies, societies and citizens with geospatial approaches, data and technologies and, in so doing, deliver the location-based knowledge, services and automation expected in this 4IR digital age**. It is strategic, supporting Bill Gates' statement by moving today's geospatial concepts along the value chain towards knowledge, the answers to user challenges. It breaks out from the current focus on data infrastructures to knowledge infrastructures, from production to impact. GKI is not just about geospatial knowledge today, it is about preparation for the future, whatever that future may look like.

GKI embraces **8 principles** to support the enablement of knowledge through future geospatial infrastructures:

- **Knowledge is the focus.** Humans and machines seek knowledge to make decisions.
- Geospatial technology and data organisations **participate in the wider digital ecosystem utilising the wider digital infrastructure.**
- **Predictive.** Be it in milliseconds or years, outcomes move from insight to intelligent foresight. Time and place are powerful contributions to prediction, integrated with many non-geospatial inputs.
- Normally **led by users** knowledge needs, not data and applications available.
- **Decentralized.** The web is naturally decentralized, with data and application value-chains becoming more complex across wider ranges of partnerships. Findable, Accessible, Interoperable and Reusable applications and data of known quality are critical to this principle
- **Collaborative solutions.** To more quickly meet societies' global sustainability challenges, including the digital divide.
- **Agility** is inbuilt. Innovative, dynamic and agile solutions are far better able to respond to global challenges, technology change and people's expectations in a rapidly developing world.
- Take **achievable actions now and then scale**; success cannot be achieved overnight.

To support nations and organisations apply these principles to their digital and geospatial infrastructures, GKI considers an integrated set of six elements and associated guidelines to support growing maturity.

Figure 1. The six elements of a geospatial knowledge infrastructure (GKI)

1.3.1 Integrated Policy Framework

Governments will increasingly enable and regulate the development of a knowledge economy and society, often through lenses of revenue, privacy and ethics. Geospatial Knowledge sits primarily across digital/data, AI, technology, space, infrastructure, science and privacy policy domains. Policies that are integrated and aligned assist governments to establish an improved location approach for government policy making, services and to release industry. For example, government remote sensing and geospatial data can be opened in parallel whilst ensuring that the business innovation ecosystem helps turn these policies into impact. But no domain is excluded; security, cultural, social, environment, educational, energy and business domain policies should be geared to increase the national geospatial knowledge resource. Alignment and integration of data and knowledge policies across government departments is both a challenge and an opportunity; it sets the base upon which a GKI can be built.

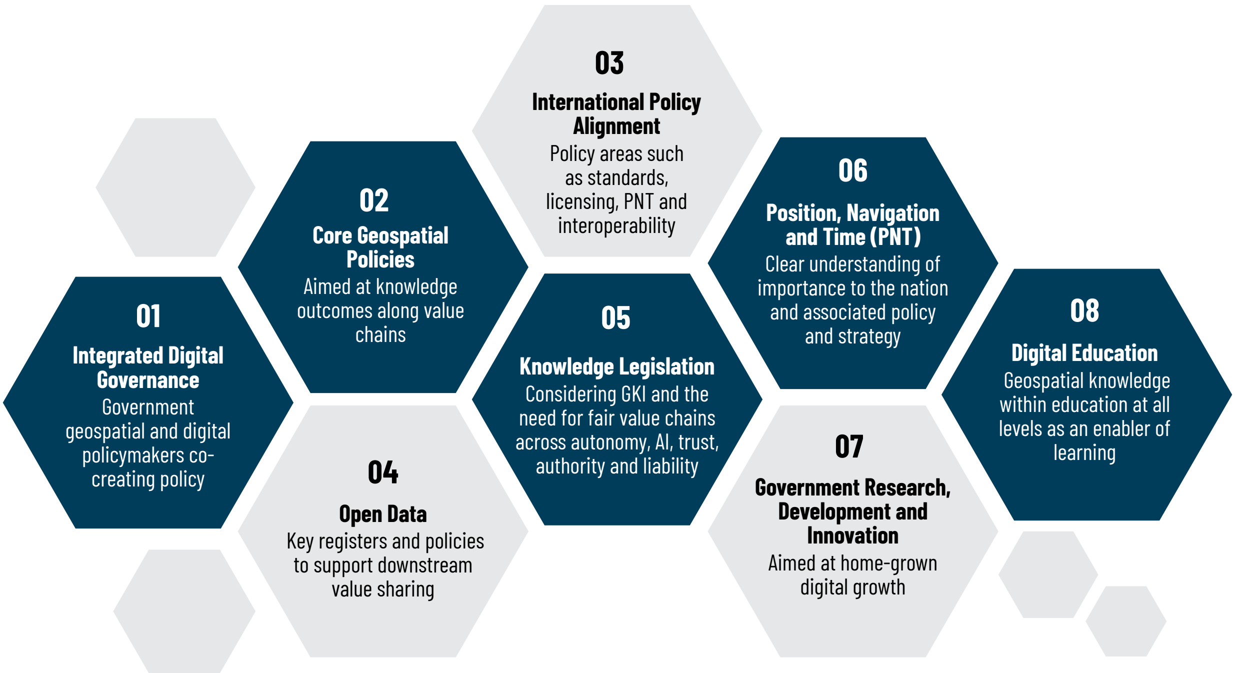


Figure 2: Initiatives of Element 1 – Integrated Policy Framework

In addition, moving up the value chain towards knowledge, integrated policy will see countries consider:

- A national geospatial knowledge strategy, which could be part of a wider geo-digital strategy.
- Policies and standards that enable Known Quality, Findable, Accessible, Interoperable and Re-useable (QFAIR) foundation data and applications. Associated geospatial-related policies on security, data protection and privacy are based on outcomes not inputs. The policies and standards are for machines and humans knowledge generation, including standards, accessibility and open data.
- Aligned and coherent national digital infrastructure, AI, space, EO, science, environmental and business policies and strategies (and in some cases, laws). Associated policies on data security, data protection and privacy are based on outcomes not inputs.
- Government departments that ensure wider digital and geospatial policies are considered when preparing sector specific policies, maximizing existing national resources. Topically, this includes emerging policies that cover safe air/land autonomous vehicle operations.
- A national positioning, navigation and time (PNT) policy and/or a PNT resilience strategy.
- Based on user needs and resilience, and collaborating with industry and academia, government national digital policies enable the establishment of interoperable and interactive knowledge-centric digital twins (in other words more than 3D geospatial data), enable digital innovation and increase the economic, social and environmental benefits of geospatially enabled knowledge.
- Digital governance, including geospatial governance, is considered holistically in government, with data and infrastructure in the same or co-operating governance structure.
- Education policies bring geospatial understanding into digital literacy in schools and into AI, data, digital and sector-specific courses – from school through to post-graduate.

- Nations liberalize partnership policies better to enable industry to partner with government to deliver government digital services.
- Ultimately, Governments create an integrated knowledge policy framework that establishes the sustainable infrastructure for a digital economy and society that utilizes location.

Together, these improve integrated, evidence-based government policies and services, establish frameworks that maximise the benefits of geospatial knowledge to nations, maximise the value of a digital infrastructure and economy and create the resilience necessary for crisis.

- Much data policy and legislation is driven by Europe: the 2019 EU Open Data directive is prescriptive, classifying Earth Observation as a high value dataset to be available for re-use; the EU's restrictive, but world-leading, approach to data protection for individual privacy impacts integrative value added services and is being emulated elsewhere in the World; the EU is now regulating AI.
- In contrast India's 2022 geospatial policy aims to grow the geospatial sector's ability to support the nation's social and economic development by liberalising data collection and management. The policy aligns with the country's digital vision and is one of several aligned and related policies including the 2023 space policy. By 2030 the policy states that India will "Develop a Geospatial Knowledge Infrastructure (GKI) underpinned by Integrated Data and Information Framework."
- In forming GEOSA, the Kingdom of Saudi Arabia is putting in place the policies necessary to grow the geospatial sector and grow the benefits of geospatial technologies and data.

1.3.2 Foundation Data

New data-hungry and digitally driven technologies operate in a complex world. There is just too much information for humans to process nor does any data owner understand all the benefits that the data can deliver; technologies that manage data and make it accessible are vital. Trusted and authoritative foundation data is an essential element of national infrastructure. Firstly, because it is a key digital representation of the physical world. This includes (1) the geodetic reference frame upon which position and height are referenced, (2) continuously updated QFAIR fundamental geospatial data, including High-Definition mapping and access to earth observation and (3) allied spatial data sets. These are national or city spatial data sets that are not considered fundamental geospatial data by the United Nations such as and

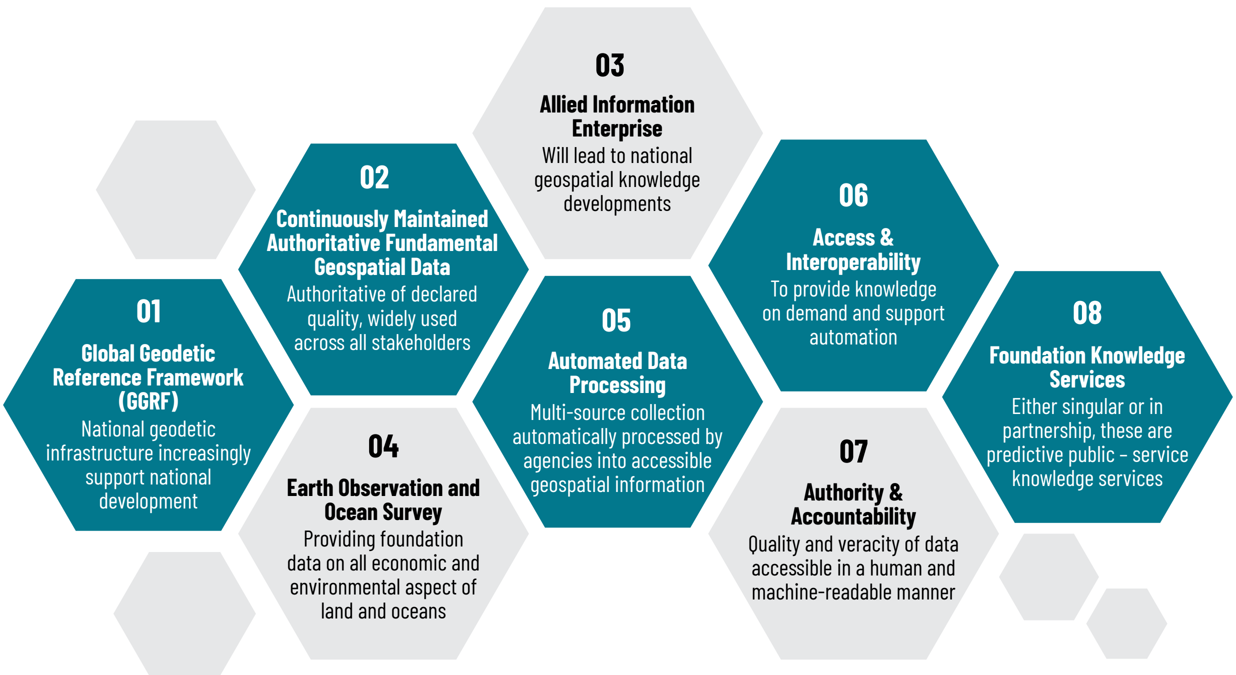


Figure 3: Initiatives of Element 2 – Foundation Data

statistics beyond demography, aeronautical information, climatology, traffic flows and people movement and, critically across all most SDGs and business sectors, meteorology. Together with analytics, these enable nations, businesses, citizens and machines to derive the knowledge they seek to solve problems, exploit opportunities, and deliver new value. Secondly, assured or authoritative foundation data provides greater assurance of, and trust in, knowledge-based machine-derived outcomes. The term foundation data is used to reflect the far greater breadth of data considered in GKI than within the UN-GGIM frameworks, where the term fundamental geospatial data is used.

In addition, moving up the value chain towards knowledge will see:

- Foundation data agencies that reach into client sectors to understand knowledge use cases and develop base data, service and knowledge services based on current and emerging requirements.
- Government designated authoritative data, with known provenance and quality that help overcome uncertainty and risk, as part of the national infrastructure.
- Foundation geospatial information carries authority but currently often not accountability. The gap will narrow.
- Government mandated system of persistent, consistent, and unique identifiers across registries and government databases to enable integration.
- Based on current and emerging requirements, holistic nationally defined geospatial information management arrangements and sensor mix, drawing upon global, government and industry systems, and where necessary national systems. Government does not replicate when industry is sustainably meeting use cases.
- Agencies that develop through product-centric, data-centric, solution-centric to knowledge-centric outputs. They capitalise on AI, EO and other sensors, cloud and high-performance computing to facilitate autonomously delivered continuously maintained and integrated data from many sources and offer knowledge services where there is a public task requirement.
- Efficient access to earth observation data services in time of peace and guaranteed in times of trouble.
- Improved understanding of our oceans through expanding many land based fundamental data themes to the oceans given their economic and environmental value.
- Agencies will adopt QFAIR principles.
- Foundation data is only as good as the people in charge of it; greater emphasis will be placed on enhancing geospatial and data skills in geospatial agencies.

Together these enable governments to deliver foundation data as part of national digital infrastructure and geospatial agencies, utilising 4IR technologies, to meet user requirements including for QFAIR continuously updated authoritative data.

- Ordnance Survey in the UK has worked with partners to determine the feasibility of incorporating sensor data from private sector utility vehicles, with automated processing, into its MasterMap database, seeking to provide better data content and currency. It continues to innovate, working with the energy sector to try and solve the problem of a single national power supply database that will accelerate UK's net zero target.

- Estonia's Land Board not only provides individual data types (such as trees) as a service but it is developing national geospatial knowledge services to support policy development, essentially 'digital recce' that allows understanding of roof insulation and residential solar energy potential. It is linking data to support better knowledge generation.

1.3.3 Partnerships and Collaborations

Given rapid technological change and the increasing sources of data and analytics across the private sector and academia, partnerships and collaborations will be increasingly common in co-creating value in data and knowledge. Costs and benefits will be shared across the value-chain, which crosses government, industry and citizens. New partnerships will emerge – be they ad hoc collaborations to solve singular challenges, long-term business partnerships servicing a user sector or public-private knowledge co-creation partnerships. Partnerships and collaboration are already happening between companies within sectors and between businesses and government innovation programs, but not typically between sectors. This element includes the important partnership with citizens, both as users and data providers, especially as citizens gain more rights to data about them and their location.

Whilst our default setting is Public private partnerships (PPPs), they have barely touched the geospatial ecosystem. Co-creation is therefore an approach strongly supported under GKI.

In moving up the value chain towards knowledge, Partnerships and Collaborations also includes:

- A national geo-digital knowledge experimentation enterprise. This derives cross-government user-focused future geospatial knowledge requirements arrangements and includes government and industry players across the future geospatial and digital ecosystems. It funds pilots and research and includes geospatial accelerators and innovation.
- The refocus of academic, professional and industry bodies towards knowledge creation and management professional development.
- Clear policy on citizens' voluntary contribution to national geospatial data, including industry-held citizen data.
- Laws that enable collaboration and partnerships in a digital world where industry is leading much innovation and disruption and can help governments achieve national objectives at a reduced cost.
- Clear relationships established between the research community and industry to monetise geospatial knowledge research.
- National partnership with organisations such GEO, W3C and OGC as well as national participation in UN GGIM global and regional committees of experts.
- The development of data trusts and similar mechanisms to help ensure fair reward across value chains.
- Geospatial innovation and accelerator programmes will be created in nations or in partnerships between nations to support the development of geospatial start-ups, including the incorporation of geospatial expertise in wider digital innovation centres.

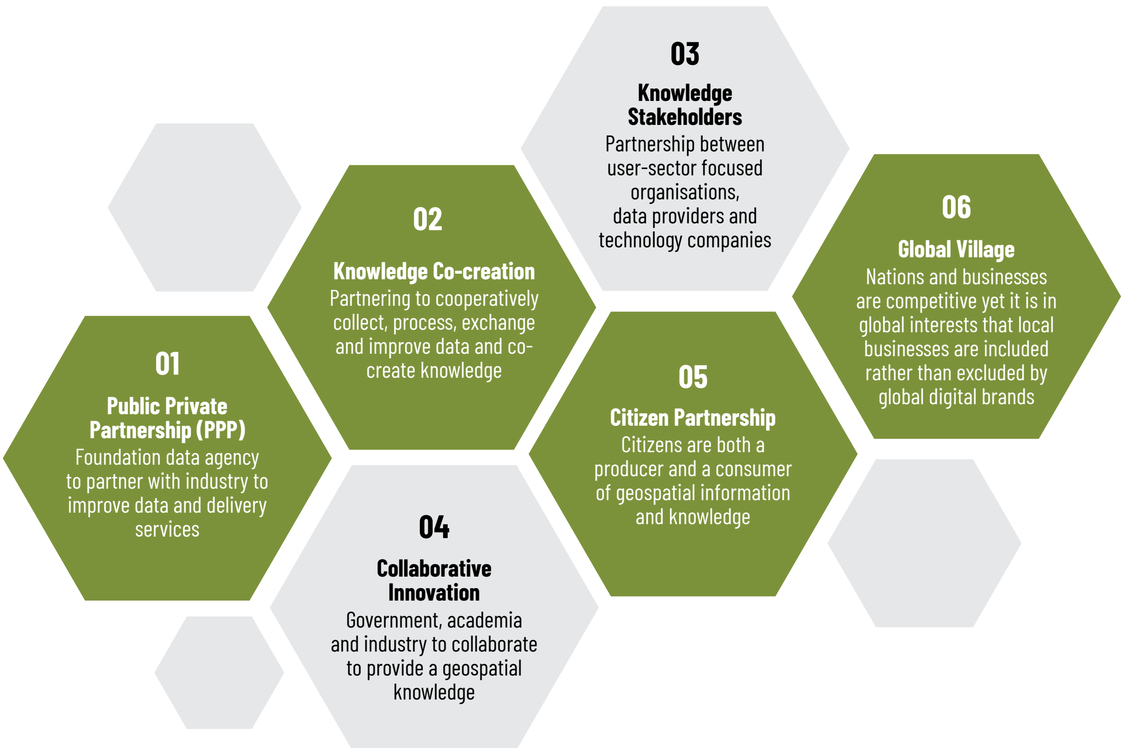


Figure 4: Initiatives of Element 3 – Partnerships and Collaborations

Together these enable widespread cross-sector geospatial knowledge engagement, industry-government partnerships and an innovative, collaborative, industry.

- As part of its wider digital agenda, the Government of Norway seeks an open national geospatial knowledge base and shared solutions and technology that support effective problem-solving and enable new applications. Norway's national mapping agency, Statens Kartverk, has focussed heavily on partnerships. It's partnerships with municipalities has led to improved national geospatial databases and it's long-term partnership with a number of eastern European nations, such as Georgia, not only helped nations create data but developed capacity and a long-term UN-IGIF country-level action plans.
- Ireland's SDG Data Hub was developed by Tailte Éireann, the national geospatial agency, in partnership with the Department of the Environment, Climate and Communications and the Central Statistics Office. It provides clear knowledge to anyone and everyone on Ireland's progress against SDG targets.

1.3.4 Industry Leadership

Even if based on government-funded research, most innovation, autonomy and knowledge services will be delivered by industry faster than Governments can understand implications or act, therefore giving industry an increasing leadership role and growth potential.

Most major businesses sectors use geospatial technologies in work processes. Covid has chased the laggards (the construction industry, for example) into new thinking, although land administration technology opportunities are often hampered by policy and legal constraints. Industry is opening many other new opportunities: collecting and managing geospatial data fit for many uses at a fraction of the price, reducing emissions and spotting abnormal emissions as we tackle climate change, dating applications, reducing banking fraud, the ability to tackle tough multi-facetted challenges, deliver eGovernment services, manage digital twins.

However, with this leadership comes social responsibility, a need to work with governments, institutions and citizens to help policy development and the development of appropriate skills. For global businesses these social responsibilities are global. Companies offering knowledge services to farmers, for example,

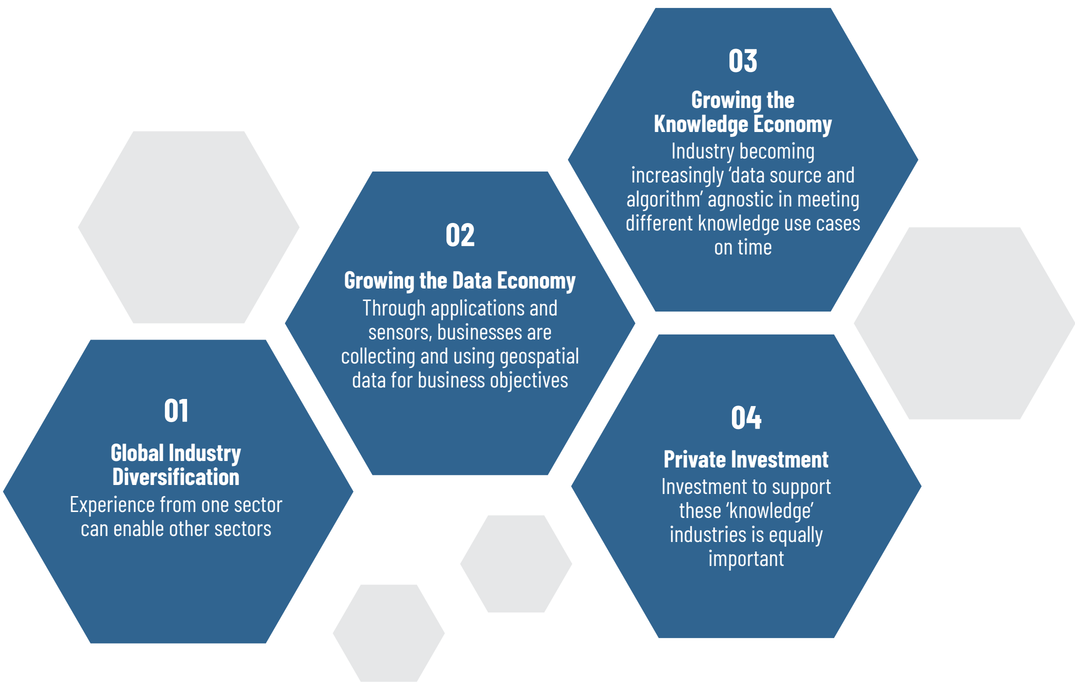


Figure 5: Initiatives of Element 3 – Industry Leadership

have the potential to transform agricultural yields but social responsibility obliges that knowledge pricing in developing nations needs to be very different to rich nations.

On investment, it only takes a glance at any article on the direction of venture capital to see that geospatial knowledge is important to the workflows or outputs of the majority of supported digital start-ups. Yet many pure geospatial-knowledge start-ups find it difficult to gain investment.

In moving up the value chain towards knowledge effective industry growth also includes:

- Growing user sector industry associations understanding of geospatial knowledge and vice-versa (in particular the World Geospatial Industry Council). This is both to increase the use of geospatial knowledge in economic sectors and for the core geospatial sector to embrace the growing geospatial ecosystems in each of those sectors.
- Digital industry involvement in policy-making from a social good, as well as profits motivated, perspective and declaring this publicly.
- Investors in upstream Space sector to further consider investment in downstream knowledge solutions companies. It's not 'one or the other' but both that deliver value.
- Differential pricing of technologies, knowledge services and data to help overcome the geo-digital divide.
- Increasing investment in geospatial research within sectors, by both government and industry.

Together these lead to increased social and economic benefits of geospatial knowledge capabilities across all industry sectors, with user-sector business processes adjusting to gain maximum value from geospatial knowledge.

- USA has the world's leading digital businesses, mostly growing without any government influence and well supported by investment and a sense of entrepreneurship. Many household digital names have major geospatial interests and integration with workflows. However, it has managed it's remote sensing industry more judicially, balancing potential capture capability with security concerns.

- Finland's national space strategy reflects the major transformation taking place in the space industry. This transformation fundamentally changes the space industry's roles and practices. New private service providers complement the satellite systems funded and operated by public organizations, for example through the provision of private launching services that enable more affordable and easier access to space.

- The European Union has dedicated incubation centres for space, that is, the European Space Agency's (ESA's) Business Incubation Centres (BICs). In the USA, the National GEOINT Agency (NGA) has clear paths for industry developments to be considered, and potentially incorporated, into workflows.

1.3.5 Users Applications, Analytics and Modelling

Users are both human and machine and across most global social, economic and environmental challenges and industry sectors.

Technologies that make sense of the world and predict changes, are vital. In the same sense that foundation data is fundamental to geospatial knowledge, so too are the applications, analytics and modelling capabilities that consume data to provide users with information and knowledge, naturally incorporating location. This knowledge is essential for human and machine decision-making, including all automation. AI/ML is allowing vast swathes of remotely-sensed data from a multitude of sensors to be processed into information, creating insights (it is estimated that 100Tb of meteorological data is collected a day, for example).

Recognising the next dominant geospatial technology is difficult amongst so much change and innovation, including AI, and thus opportunity abounds for revolutionary change although most will be evolutionary. Machine-readable foundation data could be a trigger for revolutionary change in knowledge-on-demand and autonomous worlds. Viscount Camrose, the UK Minister for AI and Intellectual Property, states in the UK's second geospatial strategy (2023) that "The geospatial ecosystem is being transformed and disrupted by advances in technology and there remains limited awareness of the full value of geospatial applications".

Algorithms can deliver elusive foresight based on the context of the question or use case, but this is a work in progress as few applications really enable prediction. When they do, they are often constrained on available data, including foundation geospatial data, and greater confidence is often sought in the derived predictive knowledge, especially in autonomy. But such algorithms and models do predict weather and flooding, they ensure automated vehicles avoid pedestrians and target consumer marketing. Many open algorithms will sit alongside open data as part of the digital infrastructure, all of which has to be of known quality and be QFAIR.

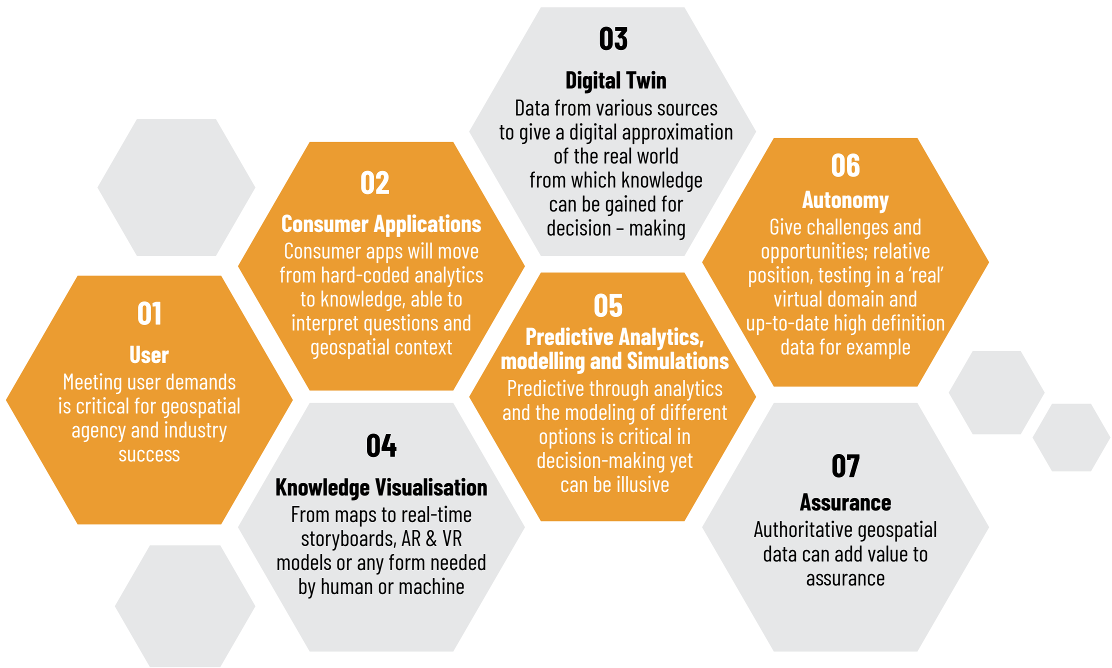


Figure 6: Initiatives of Element 5 – Applications, Analytics and Modelling

In moving towards geospatial knowledge, we will also see the following developments being considered withing user applications, analytics and modelling:

- Across government and industry, workflows and analytical teams will diversify to incorporate a variety of analytical and predictive techniques and multiple data sources, to provide better outputs and assurance.
- Guidelines and standards for the use of location in applications will be developed to enable the veracity of derived knowledge to be expressed consistently.
- Virtual testing systems for safety critical autonomous systems fully incorporate digital geospatial models from around the world.
- Open standards are developed to allow visualization, augmented reality and virtual reality systems to be data agnostic.
- Industry (and whole geo-digital ecosystem) is open with customers about the assumptions and bias within models, applications and analytics.
- Guidelines and standards for the use of location in applications will be developed to enable the veracity of derived knowledge to be expressed consistently.

Together these support the development of applications and models for knowledge economies.

- In the USA, the National GEOINT Agency (NGA) has clear paths for industry developments to be considered, and potentially incorporated, into workflows. As a result, St Louis is becoming a geospatial industry hub but with a major customer on the doorstep a different proposition to China's Deqing County geospatial industry hub.
- Starship Technologies provides a convenient last mile delivery service with fleets of sidewalk delivery robots navigating the world each day. In most places, existing data is not good enough for robots on sidewalks. The system, including the robots, therefore collects, integrates and uses geospatial data. It does not take much imagination to see how such robots could expand into other uses, for example supporting blind people.

1.3.6 Geospatial Dimension to the Wider Digital Ecosystem

Location is increasingly ubiquitous as a data attribute and the use of the web and cloud is increasing, yet much geospatial information is not accessible from geospatial data owner in a machine-to-machine transaction. GKI integrates with and supports the geospatial dimension to the wider digital ecosystem through an open positioning infrastructure, machine-accessible geospatial information that can be seamlessly integrated into sector workflows, standards, expertise, technologies, and analytics. The digital infrastructure should develop with location, geospatial data and geospatially enabled analytics at its core, a true geo-digital infrastructure. If the geospatial ecosystem does not involve itself in this, it will be left out as others develop the answers. Web 3.0 will improve the environment that enables this.

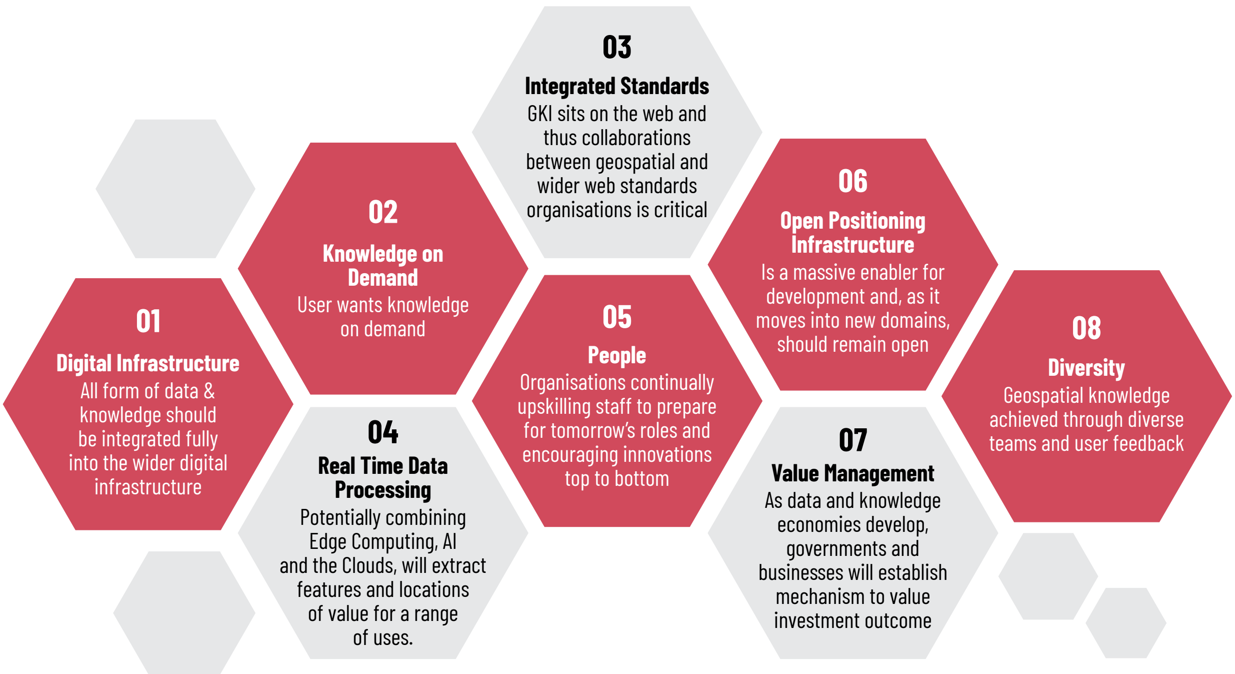


Figure 7: Initiatives of Element 6 – Geospatial Dimension to the Wider Digital Ecosystem

In moving up the value chain towards knowledge nations and organisations will consider:

- Enact knowledge on demand through globally persistent URIs for Spatial Things, spatial data indexable by search engines, API and linked data.
- Data providers use OGC Tier 3 and 4 standards and geospatial standards organisations fully engage with W3C and similar organisations to develop Tier 4 standards that support machine-to-machine transactions, help give assurance in derived knowledge (needing consideration of data and user technologies) and supports industry growth.

- Geospatial education and professional development adapt for the broader geo-digital ecosystem.
- Employers recruit and develop talent from across society to bring diversity in thinking, bias reduction and to make most of scarce talent.
- Governments and businesses better understand the economics of geospatial knowledge within the digital ecosystem.
- Data literacy, as part of education, helps address fundamental questions on why geospatial data is valuable, how it relates to the location knowledge they consume and the privacy aspects of such data.

Together these aim to embed geospatial technologies and data within digital infrastructures, including the internet.

- The common standards that are followed and implemented in almost all countries are those of the ISO and OGC. However, in addition to following the common standards, some countries like China have developed and adopted indigenous data standards of their own. OGC has worked with W3C too, to consider greater integration of geospatial data and technologies on the web.
- Within the United Nations Future Geospatial Ecosystem program, Australia's Dr Lesley Arnold has further developed the concept of Knowledge on Demand, the steps to make geospatial information findable and accessible on the web.

1.4 Role and relevance of GKI in National Development

The future lies in adopting geospatial way, and it is the most efficient and effective way to collect, process, integrate, and utilize information for overall global, national, and individual development. But, simply concentrating on data and information alone does not achieve this, so it's time to move up from the data value chain to the knowledge value chain. GKI embodies a concept that champions the evolution of a nation's geospatial ecosystem along the value chain towards knowledge, diverging from the current emphasis on data infrastructures towards the establishment of knowledge infrastructures. National geospatial infrastructures, oriented towards knowledge, serve as structures empowering countries to ascend the knowledge value chain. These infrastructures are essential for national development, and they require partnerships, knowledge sharing, and capacity building to be successful. Countries can adopt it by investing in developing a positive and collaborative approach, including government to government, government to private, government to funding agencies, academia, and research institutes.

The adoption and integration of geospatial knowledge in critical economic sectors is crucial for national development. GKI plays a significant role in this regard by integrating geospatial approaches, data, and technologies to deliver location-based knowledge, services, and solutions. By incorporating geospatial knowledge into user workflows, the potential impact on sectoral growth and development becomes evident.

1.5 Evolving roles of National Geospatial Agencies (NGAs) in Digital Age

Trusted and authoritative foundation geospatial data refers to information that is highly reliable, accurate, and accepted as a standard reference within the geospatial community. This data serves as the bedrock upon which various geospatial knowledge products and services are built and act as a foundation of knowledge – focused geospatial infrastructure. National Geospatial Agencies (NGAs) play a major part in delivering critical, hopefully trusted, authoritative data into the digital ecosystem and onto the digital infrastructure.

It's not that long ago when the role of NGAs was to solely focus on producing high quality datasets to underpin their national geospatial data infrastructures. Many NGAs around the world are doing far more than just creating data, they are providing the backbone for core decision making which is being driven by a marketplace who understand and recognize the importance of authoritative and trusted data.

Further, NGAs also play a crucial role in the development and maintenance of geodetic infrastructure. They are responsible for managing and making available national reference systems, which are fundamental to spatial data integration and autonomy. Geodetic data, serve as one of the primary components of the National Spatial Data Infrastructure (NSDI). The geodetic infrastructure provides accurate information about fundamental properties of the Earth and has numerous scientific, civil, military, and commercial applications.

Geospatial Knowledge at the heart of tomorrow's global digital society

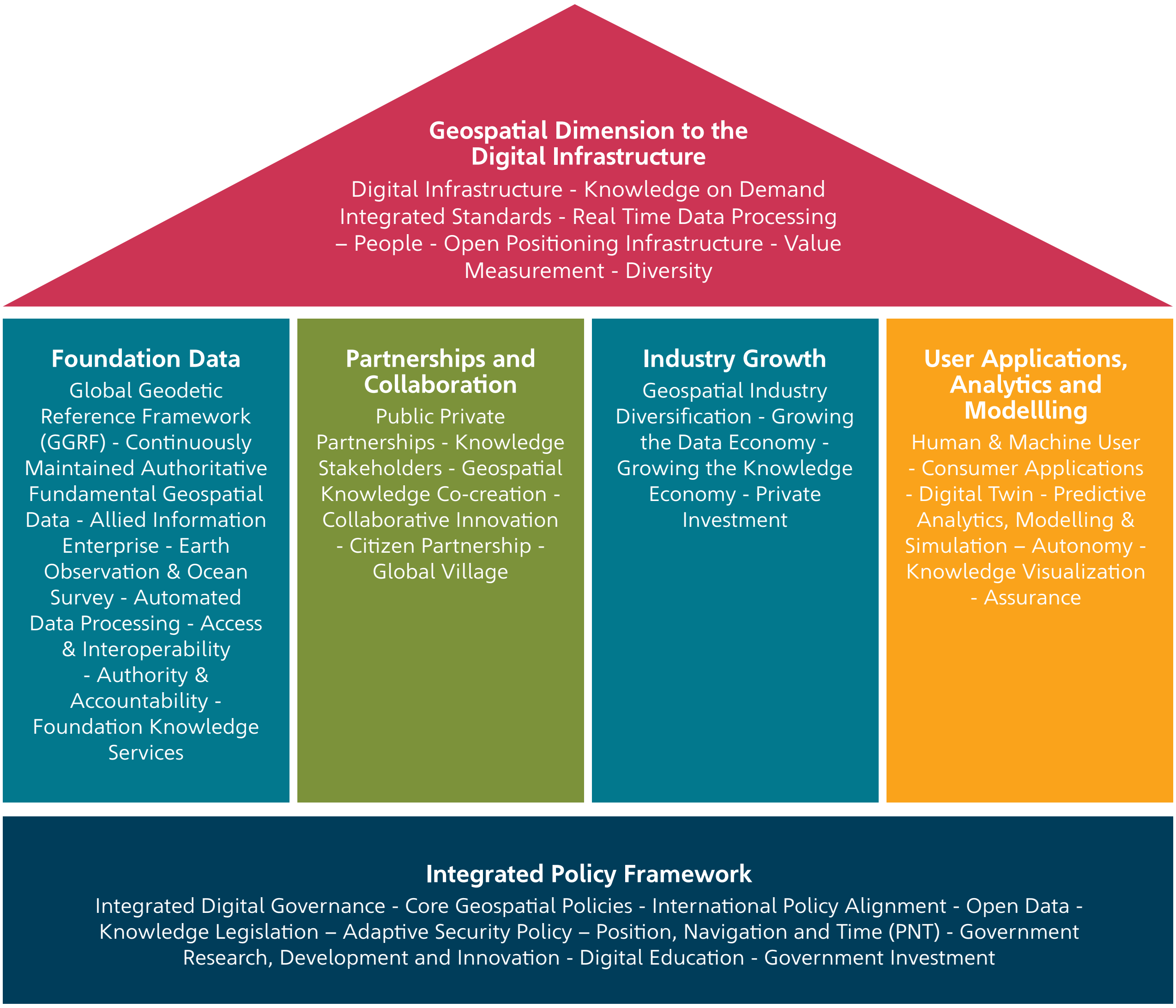


Figure 8: Pillars of Geospatial Knowledge Infrastructure (GKI) Framework

NGAs are involved in maintaining and modernizing the geodetic infrastructure to ensure its accuracy and relevance for diverse applications.

GKI looks more broadly at an ecosystem where data and applications combine to deliver user-led knowledge. The development of a future-ready geodetic infrastructure is essential to meet diverse user needs. Therefore, NGAs need to redefine their roles from a basic mapping mandate to providing digital services and solutions to all stakeholders. Renewed investment in the geodetic infrastructure is needed to maintain and modernize existing national security, scientific, and commercial applications. The Road Map for the Global Geodetic Reference Frame for Sustainable Development also recognizes the global distribution of geodetic infrastructure as far from optimum and emphasizes the need for a global investment plan to enhance geodetic infrastructure.

User demand identification and development of demand-driven knowledge are essential aspects of the GKI by NGAs. This involves understanding the evolving needs of various user groups and developing geospatial knowledge products and services tailored to their requirements.

The digital transformation and modernization of NGAs are essential to keep pace with the evolving geospatial landscape and to meet the growing demand for geospatial data and services. This transformation also involves the need for IT modernization to support GKI and to bridge the geospatial digital divide. NGAs are increasingly recognizing the importance of adopting emerging technologies and innovative practices to leapfrog to state-of-the-art geospatial information management systems. Geospatial technology is being leveraged as a powerful enabler of digital transformation in government, with NGAs persisting in updating their geospatial strategies to accommodate the growing demand and ensure effectiveness. The evolution of NGAs to support the digital and autonomous society involves collaborative efforts between geospatial, digital, and user industries to co-create enhanced geospatial knowledge services tailored to user demands. Therefore, the digital transformation and modernization of NGAs are pivotal in enabling them to provide efficient and effective geospatial services in the digital era.

There is no one model of evolution of national geospatial agencies, as circumstances are very different across a very digitally divided world. However, Figure below gives a perspective on the direction of travel. It elaborates the evolving role of national geospatial agencies in the light of development.

Maturity Stages of NGAs towards high socio-economic impact

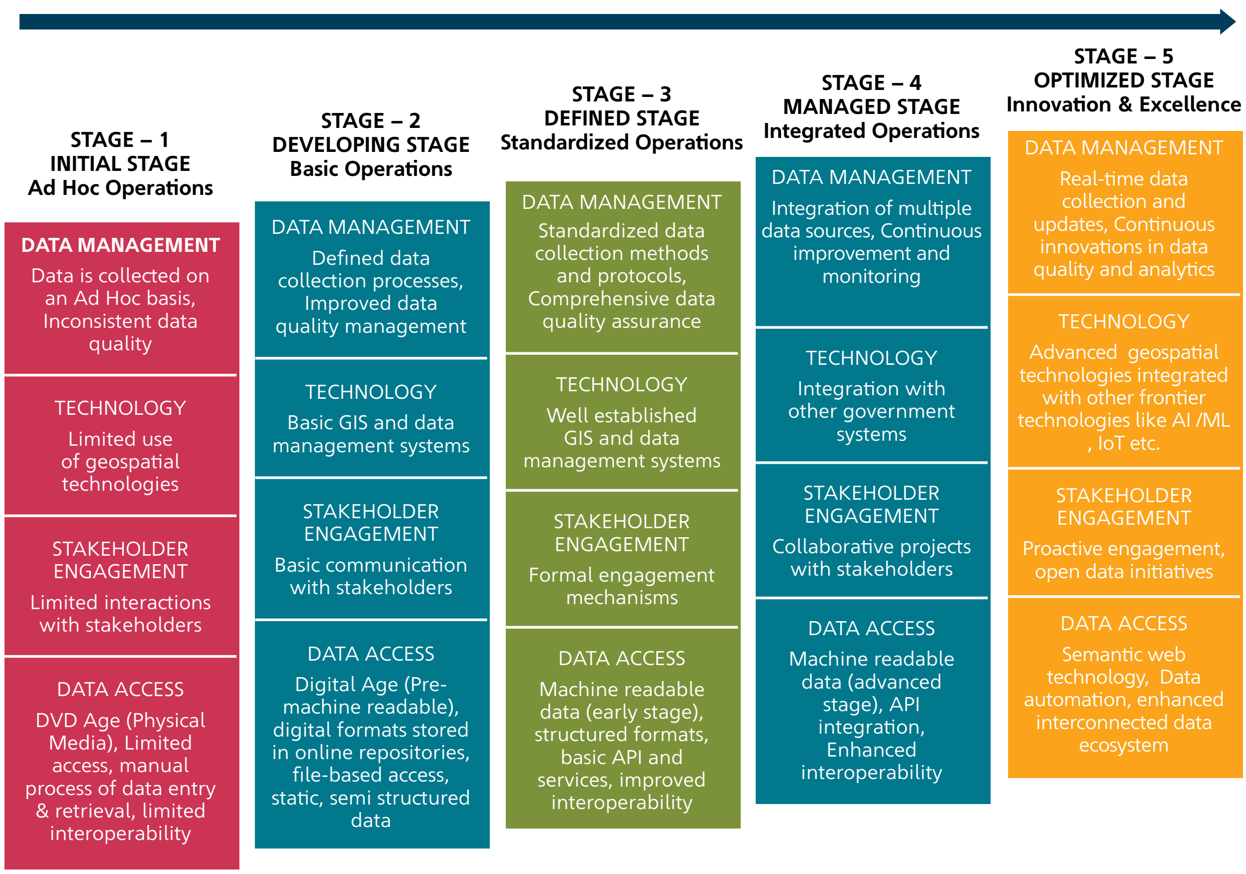


Figure 9: Maturity Stages of NGAs towards high socio-economic impact

1.6 References

1. "Climate in Asia and the Pacific", (2014) Springer Science and Business Media LLC.
2. Bibri, S.E.; Krogstie, J. Smart sustainable cities of the future: An extensive interdisciplinary literature review. Sustain. Cities Soc. (2017), 31, 183–212.
3. Chipofya, M. C., S. Jan, and A. Schwering (2021), SmartSkeMa: Scalable Documentation for Community and Customary Land Tenure. Land 10 (7):662.
4. Coetzee, Serena & Gould, Michael & McCormack, Bruce & Mohamed Ghouse, Zaffar & Scott, Greg & Kmoch, Alexander & Alameh, Nadine & Strobl, Josef & Wytzisk, Andreas & Devarajan, Thirumalaivasan. (2021). Towards a sustainable geospatial ecosystem beyond SDIs. 10.13140/RG.2.2.22555.39203.
5. COP28 Delivers Historic Consensus in Dubai To Accelerate Climate Action. COP28 press release (2023).
6. Craglia, M. (2015). Spatial Data Infrastructures. In: Wright, J.D., International Encyclopedia of the Social & Behavioral Sciences (Second Edition), Elsevier, pp130-135, https://doi-org.uplib.idm.oclc.org/10.1016/B978-0-08-097086-8.72058-7.
7. Eva-Maria Unger, Rohan Bennett, Christiaan Lemmen, Kees de Zeeuw, Jaap Zevenbergen, CheeHai Teo, Joep Crompvoets. (2020), "Global policy transfer for land administration and disaster risk management", Land Use Policy.
8. Food and Agriculture Organization (FAO) (2023). COP28: Global agrifood systems are the climate solution, FAO Director-General tells world leaders. https://www.fao.org/newsroom/detail/cop28-global-agrifood-systems-are-the-climate-solution--fao-director-general-tells-world-leaders/en (accessed 20-Dec 2023)
9. Fosu, R.; Suprabhas, K.; Rathore, Z.; Cory, C. Integration of Building Information Modeling (BIM) and Geographic Information Systems (GIS)—A literature review and future needs. In Proceedings of the 32nd CIB W78 Conference, Eindhoven, The Netherlands, 27–29 October (2015).
10. Geospatial Knowledge Infrastructure Readiness Index and value proposition in world economy, society and environment (Report) (2022), Geospatial World.
11. Geospatial World and UNSD (2021). The power of where? A geospatial knowledge infrastructure to enhance the world economy, society and environment. GKI White Paper. https://geospatialmedia.net/pdf/GKI-White-Paper.pdf
12. Griffinger, K.; Gertner, C.; Kramar, H.; Kalasek, R.; Pichler-Milanovic, N.; Meijers, E. Smart Cities: Ranking of European Medium-Sized Cities; Vienna University of Technology: Wien, Austria, (2007).
13. Grimm, N.B.; Faeth, S.H.; Golubiewski, N.E.; Redman, C.L.; Wu, J.; Bai, X.; Briggs, J.M. Global change and the ecology of cities. Science (2008), 319, 756–760.
14. Höjer, M.; Wangel, J. Smart sustainable cities: Definition and challenges. In ICT Innovations for Sustainability; Springer: Cham, Switzerland, (2015); pp. 333–349.
15. https://eurogi.org/beyond-sdi-zaffar-mohamed-ghouse-president-sssi/.
16. https://ggim.un.org/IGIF/
17. https://ggim.un.org/UN-IGIF/documents/SP5%20-%20Innovation%204Jul2020%20GLOBAL%20CONSULTATION.pdf
18. J. Oldfield, P. van Oosterom, J. Beetz, T.F. Krijnen (2017).Working with open BIM standards to source legal spaces for a 3D cadastre, ISPRS Int. J. Geo-Inf. Nov. 2017, 6 (11) (2017), pp. 1-19.
19. Kramers, A.; Höjer, M.; Lövehagen, N.; Wangel, J. Smart sustainable cities—Exploring ICT solutions for reduced energy use in cities. Environ. Model. Softw. (2014), 56, 52–62.
20. Li, D., Wang, S. and Li, D., Spatial Data Mining: Theory and Application. http://www.springer.com/us/book/9783662485361
21. Ma, Z.; Ren, Y. Integrated Application of BIM and GIS: An Overview. Procedia Eng. (2017), 196, 1072–1079.
22. Masser, I. (2005). The future of spatial data infrastructures. Proceedings of the ISPRS Hangzhou Workshop on Service and Application of Spatial Data Infrastructure, Volume XXXVI-4/W6, 14-16 Oct 2005, Hangzhou, China.
23. McDonald, R.I.; Green, P.; Balk, D.; Fekete, B.M.; Revenga, C.; Todd, M.; Montgomery, M. Urban growth, climate change, and freshwater availability. Proc. Natl. Acad. Sci. USA (2011), 108, 6312–6317.

24. Muhammad, A. N., A. M. Aseere, H. Chiroma, H. Shah, A.Y. Gital, and I. A. T. Hashem (2021). Deep learning application in smart cities: recent development, taxonomy, challenges and research prospects, Neural Computing and Applications 33 (7): 2973-3009
25. Offei, E., M. Lengoiboni, and M. Koeva. (2018), Compliance with residential building standards in the context of customary land tenure system in Ghana. PlaNext.
26. Satterthwaite, D. The implications of population growth and urbanization for climate change. Environ. Urban. (2009), 21, 545–567.
27. Schwab, K. (2016). The Fourth Industrial Revolution. Geneva: World Economic Forum.
28. Scott, G. & Rajabifard, A. 2017, Sustainable development and geospatial information: a strategic framework for integrating a global policy agenda into national geospatial capabilities, Geo-spatial Information Science, vol. 20, no. 2, pp. 59-76.
29. Stephanie M. Noble, Martin Mende, Dhruv Grewal, A. Parasuraman,The Fifth Industrial Revolution: How Harmonious Human–Machine Collaboration is Triggering a Retail and Service [R]evolution, Journal of Retailing, Volume 98, Issue 2, 2022, Pages 199-208, ISSN 0022-4359, https://doi.org/10.1016/j.jretai.2022.04.003.
30. Suhari, K.T., Leksono, B.E., Meilano, I., Gunawan, P.H., Saputra, N.R.M. , (2020). Implementation of 4D Cadastre Concept for Land Dispute Potential and Solution of Post Natural Disaster in Palu, Indonesia. Proceedings FIG Working Week 2020. International Federation of Surveyors (FIG), Copenhagen, Denmark.
31. Tomlinson, R. F. 1962. "Computer Mapping: An Introduction to the Use of Electronic Computers in the Storage, Compilation and Assessment of Natural and Economic Data for the Evaluation of Marginal Lands." Report presented to the National Land Capability Inventory Seminar held under the direction of the Agricultural Rehabilitation and Development Administration of the Canada Department of Agriculture, Ottawa, Canada, November 29-30. https://gisandscience.files.wordpress.com/2012/08/4-computermapping.pdf
32. Tomlinson, R. F. 1963. "Feasibility Report of Computer Mapping System." Prepared for the Agricultural Rehabilitation and Development Administration of the Canada Department of Agriculture, Ottawa, Canada. https://gisandscience.files.wordpress.com/2012/08/5-feasibility-report.pdf
33. Verheye, B. (2020). Land Registration in the Twenty-First Century:Blockchain Land Registers from a Civil Law Perspective. In Disruptive Technology, Legal Innovation, and the Future of Real Estate: Springer, 123-135
34. Williamson, I.P., Rajabifard, A. and Feeney, M.F (eds), 2003. Developing Spatial Data Infrastructures From Concept to Reality. CRC Press: London. https://doi.org/10.1201/9780203485774
35. Yamamura, S.; Fan, L.; Suzuki, Y. Assessment of Urban Energy Performance through Integration of BIM and GIS for Smart City Planning. In International High-Performance Built Environment Conference—A Sustainable Built Environment Conference 2016 Series (SBE16), IHBE 2016; Ding, L., Fiorito, F., Osmond, P., Eds.; Procedia Engeenering; Elsevier Science BV: Amsterdam, The Netherlands, (2017); Volume 180, pp. 1462–1472.

LESSON – 2

The Socio-economic Impact of Geospatial Knowledge Infrastructure (GKI)

Trainer's Biodata: Ananyaa Narain

Currently employed as Vice President - Consulting at Geospatial World, **Ananyaa Narain,** an economist by education, is a leading thought leader in the socio-economic and market research space for geospatial, BIM, and frontier technologies worldwide. She was recognized as a 40U40 geospatial professional for 2020 by xyHt, and a Geospatial Rising Star 2022 by Geospatial World. Over her eight years of experience in the geospatial technology domain, she is recognized for her result-oriented approach to exploring the impact of the cross-linkage of technology, economics, and policy using varied qualitative and quantitative tools and frameworks. Her professional expertise includes working with cross-sectoral and interdisciplinary stakeholders across technology, academia, and government stakeholders for market size and forecast assessment, socio-economic impact evaluation, developing national policy and governance frameworks, and initiating and contributing to global technology and policy dialogue and debates through research reports and articles, and conference convening.

2.1 Introduction

In an era marked by unprecedented technological advancement, Geospatial Knowledge Infrastructure (GKI) emerges as a pivotal catalyst reshaping the socio-economic landscape of any country, and it's economy. The Geospatial Knowledge Infrastructure (hereto referred to as GKI),is a multifaceted framework encompassing geospatial data, tools, and analytical capabilities, and is revolutionizing the manner in which societies and economies perceive, interact with, and harness spatial information. This transformative force transcends traditional boundaries, permeating industries, governance structures, and societal frameworks with its profound implications.

At its essence, geospatial knowledge infrastructure embodies the fusion of geospatial data with advanced technologies such as Geographic Information Systems (GIS), remote sensing, and spatial. This convergence empowers stakeholders across diverse sectors to derive actionable insights, make informed decisions, and drive innovation. Whether elucidating urban dynamics, optimizing supply chains, or mitigating environmental risks, GKI serves as a cornerstone for informed decision-making, underpinning sustainable development and fostering resilience in an increasingly complex world. Hence it is no surprise that the global geospatial industry is estimated to be USD 512 billion in 2023 and is forecasted to grow to USD 1.37 trillion in 2030. The industry's growth is likely to result in a direct economic impact of approximately USD 30 trillion by 2030! Undoubtedly, geospatial knowledge is a critical socio-economic tool with its ubiquitous presence across all sectors and in all technology domains today. Several mainstream IT and engineering firms have already added geospatial portfolios through strategic partnerships and acquisitions, driven by the motivation to strengthen the geospatial dimension of their existing business portfolios to stay ahead of the innovation and competitive curves.

Thus, the socio-economic ramifications of GKI, reverberate across a spectrum of domains (societal, economical and environmental), catalysing transformative change and yielding tangible benefits for individuals, communities, and nations alike. From enhancing public service delivery to facilitating precision agriculture, GKI engenders efficiencies, fosters inclusivity, and nurtures economic growth. Moreover, by democratizing access to spatial information and fostering collaboration, GKI empowers stakeholders at all levels to participate in data-driven discourse, thereby fostering transparency, accountability, and participatory governance.

However, as GKI proliferates, it also engenders nuanced challenges and considerations. Ethical dilemmas surrounding data privacy, security vulnerabilities, and digital divides underscore the imperative for responsible stewardship and equitable access. Moreover, as the boundaries of GKI extend beyond national borders, fostering international cooperation and interoperability emerges as imperative for realizing its full potential.

In light of these dynamics, it is imperative to **explore the socio-economic impact of the Geospatial Knowledge Infrastructure**, elucidating its transformative potential, discerning emerging trends, and delineating actionable insights for policymakers, industry leaders, and societal stakeholders. By interrogating the intersections of technology, policy, and society, a socio-economic impact assessment seeks to inform evidence-based decision-making, foster cross-sectoral collaboration, and pave the path towards a more resilient, inclusive, and sustainable future empowered by geospatial knowledge infrastructure.

2.2 Importance of Conducting a Socio-economic Impact Assessment:

Assessing and conducting a comprehensive socio-economic impact assessment of GKI emerges as imperative for informed decision-making, risk management, accountability, and stakeholder engagement. By elucidating the transformative potential and nuanced implications of spatial interventions, impact assessments empower decision-makers to navigate complexity, optimize resource allocation, and realize the full socio-economic benefits of geospatial knowledge infrastructure.

1. Economic Perspective:

- **Resource Allocation:** Understanding the economic impact of GKI allows policymakers to allocate resources efficiently, directing investment towards areas with the highest potential for economic growth and development. By assessing the return on investment (ROI) of GKI initiatives, decision-makers can prioritize projects that yield the greatest socio-economic benefits.
- **Market Dynamics:** GKI influences market dynamics by providing businesses with spatial insights that inform market analysis, site selection, and supply chain optimization. Assessing the economic impact of GKI enables businesses to make strategic decisions, allocate resources effectively, and capitalize on emerging opportunities, thereby enhancing competitiveness and resilience in the global marketplace.
- **Job Creation:** GKI initiatives often stimulate job creation and economic growth by fostering innovation, entrepreneurship, and knowledge-intensive industries. Assessing the socio-economic impact of GKI allows policymakers to quantify the job creation potential of spatial interventions, identify skill gaps, and design targeted workforce development programs to enhance human capital and employment opportunities.

2. Policy Formulation Perspective:

- **Evidence-Based Decision-Making:** Conducting socio-economic impact assessments of GKI provides policymakers with empirical evidence to inform policy formulation, evaluation, and revision. By quantifying the socio-economic benefits and trade-offs associated with spatial interventions, decision-makers can make informed choices, anticipate unintended consequences, and develop policies that align with societal needs and aspirations.
- **Risk Management:** GKI impact assessments enable policymakers to identify and mitigate potential risks and vulnerabilities associated with spatial interventions. By conducting scenario analysis and

sensitivity testing, decision-makers can assess the resilience of policy interventions to external shocks, anticipate future challenges, and develop adaptive strategies to enhance societal well-being and resilience.

- **Policy Integration:** GKI intersects with diverse policy domains, including urban planning, environmental management, transportation, and public health. Assessing the socio-economic impact of GKI facilitates policy integration and coherence by identifying synergies, trade-offs, and interdependencies across different sectors. By fostering cross-sectoral collaboration and coordination, impact assessments enhance the effectiveness and sustainability of policy interventions, ensuring holistic and integrated approaches to societal challenges.

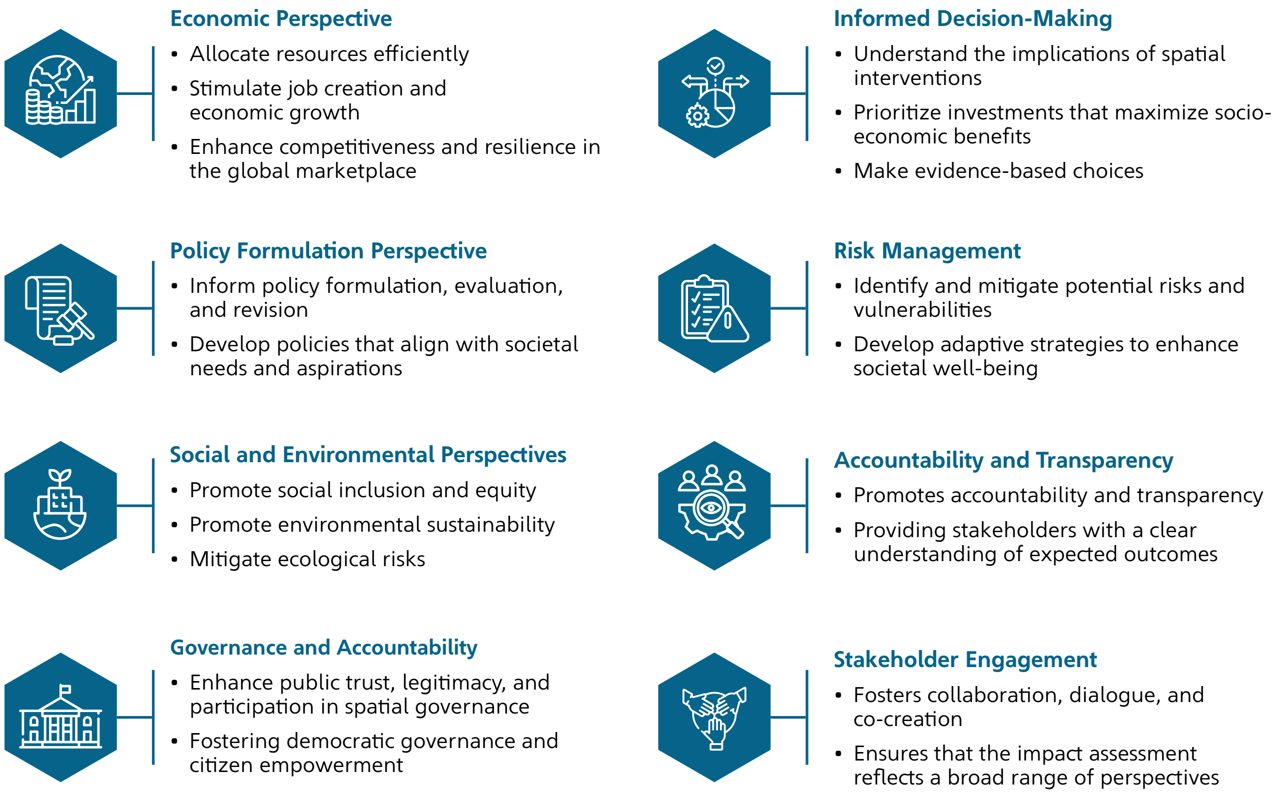
3. **Social and Environmental Perspectives:**

- **Social Inclusion:** Assessing the socio-economic impact of GKI helps identify disparities in access to spatial information and services, thereby promoting social inclusion and equity. By understanding the differential impacts of spatial interventions on marginalized communities, decision-makers can design targeted interventions to address inequities, enhance access to opportunities, and foster inclusive development.
- **Environmental Sustainability:** GKI impact assessments evaluate the environmental implications of spatial interventions, including their effects on natural resources, ecosystems, and biodiversity. By quantifying the environmental costs and benefits of GKI initiatives, decision-makers can adopt policies and strategies that promote environmental sustainability, mitigate ecological risks, and enhance resilience to climate change.

4. **Governance and Accountability:**

- Assessing the socio-economic impact of GKI promotes governance and accountability by providing stakeholders with transparency and accountability in decision-making processes. By engaging citizens, civil society organizations, and other stakeholders in the impact assessment process, decision-makers can enhance public trust, legitimacy, and participation in spatial governance, thereby fostering democratic governance and citizen empowerment.

Figure 1: Importance of Conducting a Socio-economic Impact Assessment:



Source: GW Consulting

5. **Informed Decision-Making:**

    - Conducting a socio-economic impact assessment of GKI enables stakeholders to understand the implications of spatial interventions, anticipate potential trade-offs, and prioritize investments that maximize socio-economic benefits while minimizing negative externalities.
    - By quantifying the socio-economic impacts of GKI, decision-makers can make evidence-based choices, allocate resources efficiently, and optimize policy interventions to address societal needs and aspirations.

6. **Risk Management:**

    - Assessing the socio-economic impact of GKI allows stakeholders to identify and mitigate potential risks and vulnerabilities associated with spatial interventions.
    - By conducting scenario analysis and sensitivity testing, decision-makers can anticipate the socio-economic consequences of different policy choices, assess their resilience to external shocks, and develop adaptive strategies to enhance societal well-being and resilience.

7. **Accountability and Transparency:**

    - Conducting a socio-economic impact assessment promotes accountability and transparency by providing stakeholders with a clear understanding of the expected outcomes and performance metrics associated with spatial interventions.
    - By disseminating the findings of impact assessments to the public, decision-makers can foster greater trust, engagement, and accountability, thereby enhancing the legitimacy and effectiveness of spatial governance processes.

8. **Stakeholder Engagement:**

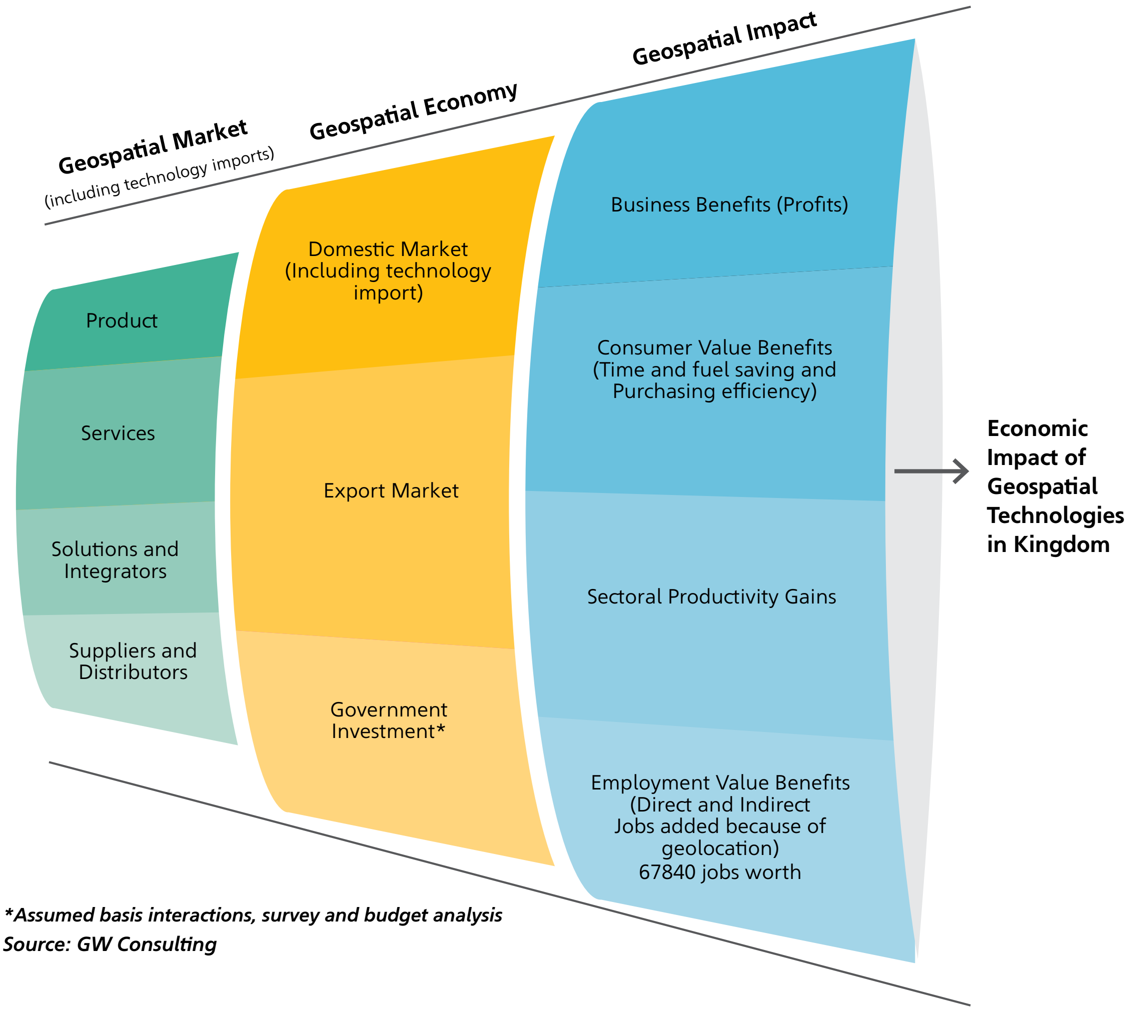
- Engaging stakeholders in the socio-economic impact assessment process fosters collaboration, dialogue, and co-creation, thereby enhancing the relevance, effectiveness, and sustainability of spatial interventions.
- By soliciting input from diverse stakeholders, including government agencies, civil society organizations, businesses, and local communities, decision-makers can ensure that the socio-economic impact assessment reflects a broad range of perspectives, priorities, and values, thereby enhancing its credibility, legitimacy, and relevance.

The integration of GKI as a concept into the socio-economic fabric of society offers a transformative potential that spans economic, governmental, and social spheres. By leveraging the power of geospatial data, societies can achieve more sustainable growth, enhanced public services, and a better quality of life for all citizens. The strategic use of GKI not only drives innovation and efficiency but also plays a crucial role in addressing some of the most pressing challenges of our time, including climate change, urbanization, and resource scarcity. As such, the development and enhancement of GKI should be a priority for policymakers, businesses, and community leaders aiming to harness the benefits of this pivotal technology.

2.3 Scope and Fundamentals of Socio-Economic Impact Assessment

Conducting a socio-economic impact assessment of geospatial technology and geospatial knowledge infrastructure on a nation's development is a multidimensional endeavour that requires a comprehensive understanding of both the technological landscape and the socio-economic dynamics at play. Quantifying the social and economic value that geospatial information contributes to modern society is a complex task. To construct reliable and consistent valuation measures, an understanding of the sequence of processes—from data acquisition to decision-makers' choices that impact society—is essential. Value is incrementally created along the chain, and pricing within a value chain is crucial for determining the distribution of this created value among the various participants (Krek, 2006). It is, therefore, important to evaluate the scope and fundamentals of socio-economic impact assessment while exploring each step of the geospatial value chain, from the technical aspects of surveying and mapping (data acquisition and management) to the delivery of services via applications, software, enterprise systems, and social media platforms, for the benefit of the global economy and society. The fundamentals of socio-economic impact assessment in the geospatial technology sector encompass the entire value chain. This assessment involves a thorough evaluation of the geospatial market, estimation of the geospatial economy, and an analysis of the overall impact of geospatial industry/technologies.

Figure 2: Fundamentals of Socio-Economic Impact Assessment



**Assumed basis interactions, survey and budget analysis*
Source: GW Consulting

To accurately assess and communicate the economic impact of the geospatial industry, it is essential to conduct a comprehensive evaluation across multiple dimensions. This evaluation involves a detailed estimation of:

Geospatial Market Size and Assessment Across Categories: This involves quantifying the scale and scope of the market, which includes Original Equipment Manufacturers (OEMs), geospatial services, solutions, system integrators, and suppliers and distributors. By measuring these segments, we gain insights into the industry's composition and the roles various entities play.

Geospatial Economy: This component encompasses the Domestic Market, the export market of geospatial technologies, and total governmental investment in this sector. Understanding these areas helps quantify the direct financial contributions of the geospatial industry to the national economy, highlighting its significance in both domestic and international contexts.

Geospatial Impact: This analysis captures the broader benefits of the geospatial industry, such as:
- Business Benefits: Profits generated by companies operating within the geospatial sector.
- Consumer Value Benefits: Enhancements in consumer services and products due to geospatial technologies, leading to improved user experiences and satisfaction.
- Sectoral Productivity Gains: Increases in efficiency and effectiveness across various sectors that adopt geospatial solutions, contributing to overall economic productivity.
- Employment Value Benefits: Job creation and enhancement, including the development of high-skill jobs and the upskilling of existing roles through technology integration.

By systematically evaluating these components, stakeholders can understand the multifaceted contributions of the geospatial industry, facilitating strategic decision-making and policy development to foster further growth and integration of geospatial technologies within the broader economic framework.

In technical sense - Socio-economic impact assessments (SEIAs) of technology are comprehensive evaluations designed to understand the broader effects of technological innovations on society and the economy. From a geospatial perspective also, these assessments help policymakers, businesses, and other stakeholders make informed decisions by providing insights into the potential impacts of deploying or developing new technologies. Here are the key scope of such assessments:

The scope of such an assessment typically encompasses various facets, including but not limited to:

- **Technological Infrastructure:** Assessing the current state of geospatial technology infrastructure within the nation, including the availability of data, satellite imagery, geospatial tools, and communication networks. This involves evaluating the extent of coverage, resolution, and accessibility of geospatial data sources.
- **Applications and Use Cases:** Examining the diverse applications of geospatial technology across sectors such as agriculture, urban planning, natural resource management, disaster response, transportation, and public health. Understanding the specific use cases and how geospatial technology is being utilized to address societal challenges and drive economic growth.
- **Impact on Governance and Policy:** Analysing the role of geospatial technology in improving governance, transparency, and decision-making processes at various levels of government. Assessing how geospatial data is integrated into policy formulation, resource allocation, and monitoring of development initiatives.
- **Economic Contribution:** Estimating the economic value generated by the geospatial industry, including revenue from geospatial products and services, job creation, and contributions to GDP growth. Evaluating the cost savings, efficiency gains, and revenue generation enabled by geospatial technology adoption across different sectors.
- **Social and Environmental Impacts:** Investigating the social and environmental implications of geospatial technology deployment, such as its influence on community resilience, environmental conservation, and socio-economic equity. Understanding how geospatial tools can facilitate inclusive development and mitigate negative externalities.
- **Capacity Building and Human Capital Development:** Assessing the capacity of individuals, organizations, and institutions to effectively leverage geospatial technology for socio-economic development. Identifying gaps in skills, infrastructure, and institutional frameworks, and recommending strategies for capacity building and knowledge transfer.
- **Policy Recommendations and Strategic Interventions:** Formulating evidence-based policy recommendations and strategic interventions to maximize the positive socio-economic impacts of geospatial technology while minimizing potential risks and challenges. This may involve promoting interoperability, fostering public-private partnerships, and investing in research and development.

Additional Considerations which should be taken into account for Socio-Economic Impact Assessment (SEIA) -

- **Evaluate Technological Advancements and Trends:** Conduct an analysis of emerging technologies and innovations within the geospatial sector. This will aid in anticipating future developments and integrating these insights into strategic planning efforts.
- **Examine Integration with Existing Technologies:** Investigate how geospatial technologies can be synergized with other advanced technologies, such as artificial intelligence (AI) and the Internet of Things (IoT). This examination should aim to uncover potential impacts and opportunities for cross-technology efficiencies.
- **Assess Impact on Privacy and Data Security:** Due to the sensitive nature of geospatial data, it is essential to evaluate the associated privacy and data security challenges. This assessment should also consider the current regulatory landscape and identify necessary compliance measures.
- **Implement Longitudinal Studies:** Initiate or advocate for longitudinal studies to track and understand the long-term effects of geospatial technologies. These studies are particularly vital in sectors experiencing rapid technological evolution.
- **Conduct International Comparisons and Benchmarks:** Engage in comparative analysis with international standards and practices in geospatial technology. This approach should aim to establish benchmarks and extract valuable insights from countries that are leaders in geospatial capabilities.

Scope of Work

Technological Infrastructure

- Data Availability: Availability of geospatial data types (satellite imagery, LiDAR, etc.).
- Coverage Extent: Geographic and thematic coverage of geospatial data.
- Data Resolution: Spatial, temporal, and spectral resolution of datasets.
- Accessibility: Ease of access to geospatial data for various user groups.
- Network Reliability: Reliability and uptime of data transmission networks.
- Integration Capability: Ease of integration with other data systems and platforms.
- Update Frequency: Frequency of data updates and refreshes.
- Technology Adoption: Rate of adoption of new geospatial tools and technologies.
- User Interface: User-friendliness of geospatial tools and platforms.
- Support Systems: Availability of technical support and user assistance.

Applications and Use Cases

- Sectoral Deployment: Usage of geospatial technologies across different sectors.
- Impact on Efficiency: Improvements in operational efficiency within sectors.
- Innovation Trigger: Role of geospatial tech in triggering sector-specific innovations.
- Problem Solving: Effectiveness in addressing specific societal challenges.
- Cost Reduction: Impact on reducing costs in critical operations.
- Decision Support: Contribution to decision-making processes.
- Public Service Enhancement: Enhancements in public services due to geospatial tech.
- Resource Management: Improvements in natural resource management.
- Emergency Response: Efficiency in disaster and emergency response.
- Health Outcomes: Contributions to public health surveillance and interventions

Impact on Governance and Policy

- Policy Formulation: Influence on the development of new public policies.
- Transparency Enhancement: Role in increasing transparency in governance.
- Public Participation: Facilitation of public involvement in governance.
- Regulatory Compliance: Assistance in monitoring regulatory compliance.
- Resource Allocation: Impact on the distribution and allocation of public resources.
- Development Monitoring: Usage in tracking development projects and initiatives.
- Data-Driven Decisions: Extent of adoption of data-driven decision-making processes.
- Cross-Agency Collaboration: Enhancement of inter-agency collaboration.
- Legislative Support: Support provided to legislative processes and law-making.
- International Cooperation: Role in fostering international governance cooperation.

Economic Contribution and Impact

- Industry Revenue: Total revenue generated by the geospatial sector.
- Job Creation: Number of jobs created directly and indirectly by the sector.
- GDP Contribution: Contribution to national GDP from geospatial activities.
- Investment Attraction: Ability to attract domestic and foreign investment.
- Innovation Rate: Rate of technological innovation driven by geospatial technologies.
- Export Capability: Impact on the export of geospatial products and services.
- Skill Development: Role in developing high-demand economic skills.
- Entrepreneurship: Stimulation of new business ventures in geospatial technologies.
- Cost Savings: Cost savings across industries due to geospatial solutions.
- Productivity Gains: Overall productivity improvements across sectors.

Social and Environmental Impact

- Community Resilience: Contributions to building resilient communities.
- Environmental Monitoring: Role in monitoring environmental changes and hazards.
- Socio-economic Equity: Impact on reducing socio-economic disparities.
- Urban Planning: Contributions to sustainable urban and regional planning.
- Cultural Heritage: Role in protecting and managing cultural heritage sites.
- Climate Action: Contributions to climate change mitigation and adaptation.
- Biodiversity Conservation: Impact on biodiversity conservation efforts.
- Water Resources Management: Improvements in managing water resources.
- Pollution Control: Role in pollution detection and control.
- Land Use Optimization: Enhancements in land use planning and management.

Source: GW Consulting

Thus in, conducting a socio-economic impact assessment, it is essential to employ a rigorous methodological framework that integrates quantitative and qualitative data analysis techniques. This may include surveys, interviews, case studies, economic modeling, spatial analysis, and stakeholder consultations. Additionally, collaboration between government agencies, academia, industry stakeholders, and civil society organizations is crucial to ensure the credibility, relevance, and legitimacy of the assessment findings and recommendations. When conducting a socio-economic impact assessment of geospatial technology on a nation's development, several key fundamentals should be considered to ensure a comprehensive and rigorous analysis.

2.4 Research Methodologies

Conducting a socio-economic impact assessment of geospatial technology on a nation's development requires employing a diverse array of research methodologies to capture the multifaceted nature of impacts. These methodologies encompass both quantitative and qualitative approaches, each offering unique insights into different aspects of the assessment. Below are several varied research methodologies commonly used in such assessments:

Case Study Analysis for Socio-Economic Impact Assessment [SEIA]

Using case study analysis in a socio-economic impact assessment (SEIA) provides a detailed and contextual approach to evaluate the effects of an intervention, such as the implementation of geospatial technology. Case studies can offer in-depth insights into complex phenomena by exploring how specific variables interact. Here's a structured approach to integrating case study analysis for socio-economic impact analysis -

1. Selection of Case Studies
Select case studies that are relevant to the technology and its application representing a variety of geographic locations, population groups, and economic contexts to capture diverse impacts including multiple case studies can provide a comparative analysis that highlights different outcomes and reasons behind those differences.

2. Assessment Framework Development
Define the specific socio-economic areas you want to assess, such as employment, productivity, access to services, inequality, and environmental sustainability. Establish clear indicators for each impact area which should be measurable and directly related to the goals of geospatial technology implementation.

3. Data Collection
Gather data on investment, operational costs, usage rates, and economic output linked to the technology in the case studies. This may include budgetary allocations, revenue generation, and cost-saving metrics. Simultaneously, conduct interviews, focus groups, and surveys with stakeholders such as local government officials, business owners, technology users, and community members. Document narratives and personal experiences that describe the perceived impacts.

4. Defining Patterns and Analysis Method
Compare findings across different case studies to identify patterns or variances in socio-economic impacts. This can help in understanding the conditions under which geospatial technology has the most beneficial impacts. Analyze how external factors such as economic conditions, policy environments, and cultural aspects influence the outcomes observed in the case studies.

5. Synthesis and Generalization
Draw together the data from various case studies to synthesize overarching insights. Highlight the key drivers of successful implementation and the barriers faced. Carefully generalize the findings to broader contexts, noting any limitations in applicability due to specific conditions in the case studies.

Incorporating case study analysis in socio-economic impact assessments not only provides detailed insights into individual and community-level impacts but also enhances the robustness of the findings by illustrating how various factors contribute to different outcomes. This approach is particularly effective in understanding new technologies like geospatial tools, where direct causal relationships may not be readily apparent.

Survey Research: Conduct surveys to gather quantitative data on the adoption, usage patterns, and perceived impacts of geospatial technology among various stakeholders, including government agencies, businesses, non-profit organizations, and individuals. Surveys can be designed to capture demographic information, technology adoption rates, user satisfaction, and behavioural changes resulting from geospatial technology interventions.

Case Studies: Undertaking in-depth case studies to explore specific geospatial technology initiatives and their socio-economic impacts in real-world contexts provides rich qualitative insights into the mechanisms, challenges, and outcomes of geospatial technology implementation across different sectors and geographical regions.

Econometric Analysis: Applying econometric techniques to analyze large-scale datasets and estimate the causal effects of geospatial technology adoption on various socio-economic indicators, such as economic growth, employment, productivity, and income distribution. Econometric models may include regression analysis, difference-in-differences estimation, propensity score matching, and instrumental variable methods to control for confounding factors and identify causal relationships.

Qualitative Interviews and Focus Groups: Conducting qualitative interviews and focus groups with key stakeholders to explore their perspectives, experiences, and perceptions regarding the socio-economic impacts of geospatial technology provide valuable insights into the contextual factors, social dynamics, and institutional arrangements that shape the adoption and diffusion of geospatial technology innovations.

Cost-Benefit Analysis (CBA): Performing cost-benefit analysis helps to assess the economic efficiency and societal welfare implications of geospatial technology investments. CBA involves quantifying the costs and benefits associated with geospatial technology projects, including both monetary and non-monetary factors, and comparing them to determine the net social value or rate of return on investment.

Cost-Benefit Analysis for Socio-Economic Impact Assessment [SEIA]

1. Define the Scope and Objectives
Define what is the aim to assess with the CBA, focusing on the economic impacts of geospatial technology. Determine the geographical and temporal scope of the analysis and whether to include all or which of direct, indirect, and induced impacts.

2. Identify Costs and Benefits
Include all costs associated with the implementation and maintenance of geospatial technologies. This includes initial capital investment, ongoing operational expenses, training costs, and any upgrades or expansions. Alternatively, measure direct benefits such as improvements in productivity, efficiency gains, reduction in costs due to optimized resource use, and increased revenue from enhanced services using case study analysis, and any information available. Account for broader economic benefits such as job creation, enhanced data accuracy for other sectors, improved environmental management, and better planning and development outcomes due to the availability of detailed spatial data.

3. Quantify Costs and Benefits
Assign monetary values to both costs and benefits using market prices where applicable. For benefits or costs where market prices are not available, use alternative valuation methods such as **shadow pricing, willingness to pay, or cost of replacement**. Estimate the expected life of the technology and the time over which the costs and benefits will accrue. This is crucial for calculating net present values.

4. Discount Costs and Benefits
Apply a discount rate to reflect the time value of money. Choosing an appropriate discount rate is critical and can be based on government economic guidelines or sector-specific standards. Alternatively, calculate the NPV of the project by discounting the benefits and costs to their present values. An NPV greater than zero indicates that the benefits outweigh the costs.

5. Sensitivity Analysis
Perform sensitivity analyses to understand how changes in key assumptions (such as discount rates, benefit estimates, or cost overruns) impact the CBA outcomes. Consider different scenarios (e.g., high adoption rates, technological advancements) to reflect varying future possibilities and their impact on costs and benefits.

Define the Scope and Objectives → Identify Cost and Benefits → Quantify Cost and Benefits → Discount Cost and Benefits → Sensitivity Analysis

Cost-Benefit Analysis provides a robust framework for decision-makers to evaluate the economic worthiness of investing in geospatial technologies, ensuring that the economic benefits derived from the technology justify the expenditures incurred.

By employing a combination of these varied research methodologies, agencies can conduct a robust and nuanced socio-economic impact assessment of geospatial technology on a nation's development, generating actionable insights to inform policy decisions, investment strategies, and technology interventions. In socio-economic impact assessment, it is import to consider and analyze the effects of interventions or technologies on society and the economy. Commonly used models include -

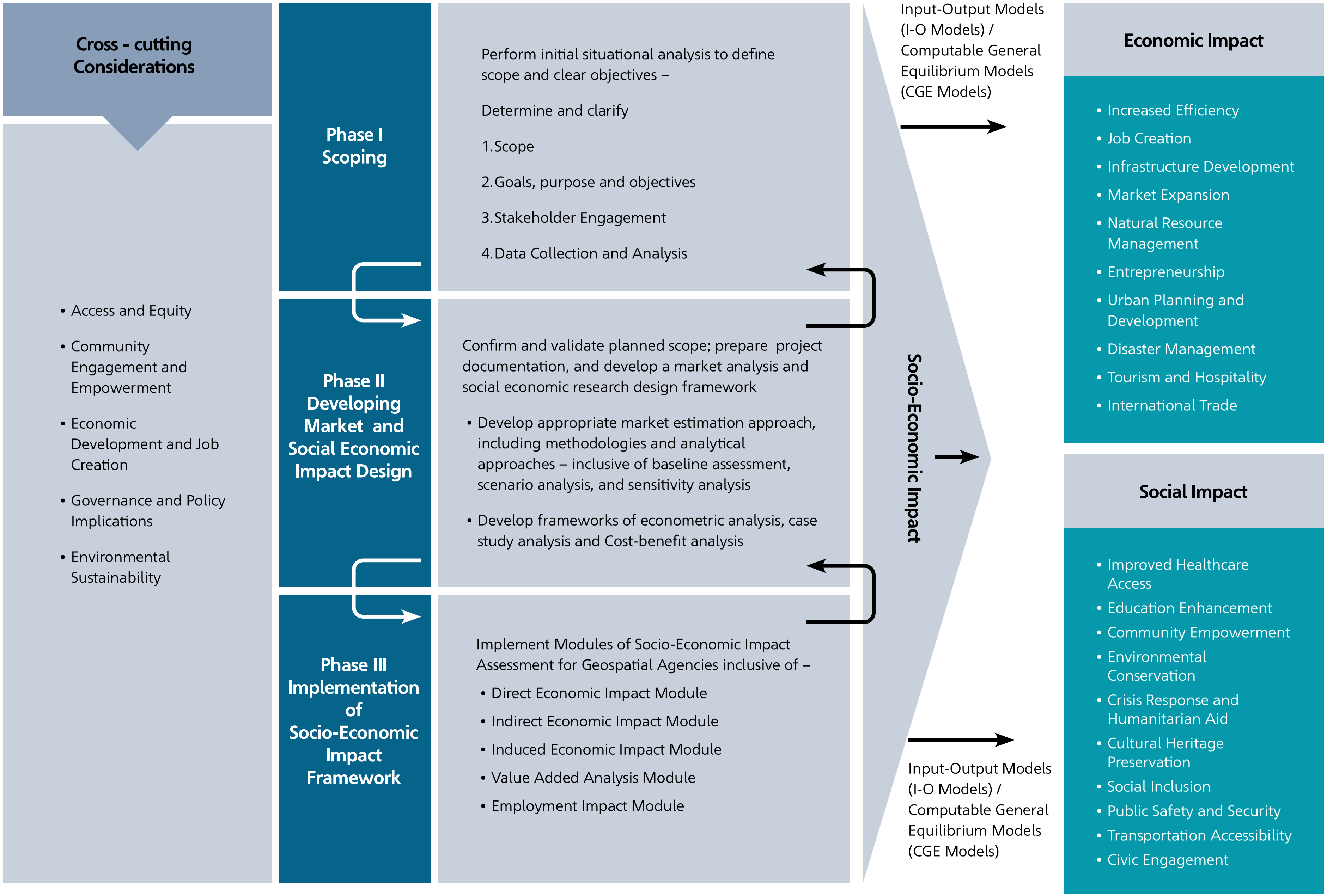
I. Input-Output Models (I-O Models):

Input-output models are economic models that capture the interdependencies between different sectors of an economy by quantifying the flows of goods and services between them. These models provide a systematic framework for assessing the direct and indirect effects of an intervention or technology on various sectors of the economy, including upstream suppliers and downstream consumers. Input-output models are often used to estimate the multiplier effects of investment or expenditure changes, which help quantify the broader economic impacts beyond the initial direct effects. In the context of geospatial technology, input-output models can be used to analyze the economic ripple effects of investments in geospatial infrastructure, software development, or technology adoption across different sectors, such as agriculture, transportation, and telecommunications.

II. Computable General Equilibrium Models (CGE Models):

Computable general equilibrium models are macroeconomic models that simulate the behaviour of an entire economy, taking into account the interactions between producers, consumers, and markets.

CGE models represent economic agents' decision-making processes, such as consumption, production, and investment, within a consistent framework of supply and demand equations. These models are particularly useful for analysing the long-term equilibrium effects of policy changes, technological innovations, or external shocks on key macroeconomic indicators, such as GDP, employment, prices, and trade balances. In the context of geospatial technology, CGE models can be employed to assess the economy-wide impacts of spatially targeted policies or investments, such as infrastructure development, land use planning, or environmental regulations.



Source: GW Consulting

Ways to Issue Challenges with Data Gaps

Dealing with data gaps is a common challenge in socio-economic assessments, including cost-benefit analyses (CBA), Input/Output Modelling, etc. These gaps can significantly impact the accuracy and credibility of the analysis. To address data gaps, especially in economic valuations where direct market prices are unavailable, methods such as shadow pricing, willingness to pay, and cost of replacement are often utilized. Here's how you can approach these strategies:

1. Shadow Pricing
Shadow pricing involves assigning a value to goods or services for which there is no direct market price, based on their implicit or 'shadow' value.

→ Identify Proxies: Look for related goods or services that have a market price and can serve as proxies. For example, if assessing the value of time saved due to faster data processing from geospatial technology, use wage rates or transportation cost savings as proxies.

→ Use Economic Models: Apply economic models to estimate the indirect market value of environmental benefits or other non-marketed goods. Models can include cost of illness for health-related benefits or avoided cost models for ecosystem services.

2. Willingness to Pay
The most used method when there is significant data gap, the Willingness to pay (WTP) measures how much individuals are willing to spend to obtain a benefit or avoid a cost.

→ Surveys and Interviews: Conduct surveys or interviews to gather data on how much users or beneficiaries are willing to pay for the benefits provided by the geospatial technology.

→ Choice Modeling: Implement choice experiments where respondents choose between different scenarios involving the technology with varying levels of benefits and costs. This helps infer their valuation of these benefits.

Contingent Valuation: Use contingent valuation methods that ask people to value an hypothetical scenario directly, such as the introduction of a new geospatial service.

3. Cost of Replacement
The cost of replacement method estimates the value of a service or good by calculating how much it would cost to replace it with an alternative.

→ Replacement Costs: Estimate the costs of replacing a service provided by the technology. For example, if geospatial technology aids in flood prevention by providing better mapping, consider the costs of alternative flood prevention measures.

→ Avoided Costs: Calculate the costs that are avoided by using the technology, such as the cost of damage from natural disasters that can be mitigated through improved planning and early warning systems enabled by geospatial data.

2.5 Modules of economic impact assessment for geospatial agencies

Expanding on the various modules outlined for the economic impact assessment of geospatial agencies provides a detailed understanding of how each contributes to a comprehensive evaluation of their economic significance. Here's a detailed look at each module:

Direct Economic Impact Module: This module demonstrates the agency's immediate economic contributions through revenue and employment, emphasizing its role as a pivotal economic entity both locally and globally. By highlighting direct financial impacts and job creation, the analysis substantiates the agency's significance in the economic landscape. However, data gaps such as the lack of comprehensive sales data segmented by customer type or geographical distribution can hinder precise assessments. Addressing these gaps involves enhancing data tracking systems, utilizing financial records, and conducting customer surveys to better capture and categorize revenue streams.

Indirect Economic Impact Module: Focusing on the broader economic effects catalysed by the agency's activities, this module elucidates how geospatial agencies enhance economic activity across various industries through their demand for goods and services like computer hardware and consultancy. It not only maps out the supply chain multipliers but also showcases the extended economic benefits resulting from agency operations, thereby justifying further public or private investment. The challenge remains in the incomplete data on supply chain impacts due to variability in supplier reporting, which can be mitigated by forging stronger partnerships for data sharing and employing robust economic modeling techniques.

Induced Economic Impact Module: This segment assesses the wider economic impacts stemming from the spending of incomes by employees associated directly or indirectly with the geospatial agencies. It highlights how these expenditures boost local economies through purchases in housing, groceries, and more. Addressing the challenge of insufficient data on spending patterns involves conducting detailed surveys of employee expenditure and applying regional economic models to estimate broader economic influences accurately.

Value Added Analysis Module: This module provides insights into the economic value added by geospatial agencies at various stages of the production process. From data collection to distribution, it assesses how these activities contribute to GDP and enhance overall economic efficiency and innovation. The main data gaps include difficulties in tracking and quantifying added value due to complex interdependencies within data flows, which can be addressed by implementing rigorous internal reporting and collaborating with economic analysts to map out these processes more precisely.

Employment Impact Module: Delving deeper into the labor market effects, this module evaluates not only the quantity but also the quality of jobs created by geospatial agencies. It underscores the agency's role in fostering substantial workforce development and providing high-quality employment opportunities. Overcoming the lack of detailed employment quality data requires regular workforce surveys, enhanced HR reporting, and longitudinal employment trend analysis.

Indicative Socio-Economic Impact Assessment Framework for National Geospatial Agencies

Module	Questions	Data Requirement
Direct Economic Impact Module	• What is the total annual revenue generated from the sale of geospatial products and services directly provided by the agency? • How much does the agency invest annually in the development and maintenance of geospatial data and technologies? • What are the direct cost savings for government agencies and private sector entities resulting from the use of our geospatial products and services? • Can we quantify the efficiency gains in sectors like transportation, logistics, and urban planning directly attributable to our geospatial services? • What percentage of our direct revenue is reinvested into research and development of new geospatial technologies?	• Revenue Data: Total annual revenue from the sale of geospatial products and services. Breakdown of revenue sources (e.g., government contracts, private sector, exports). • Employment Data: Number of employees directly employed by the geospatial agencies, categorized by full-time, part-time, and contract positions. Details on wage levels and job functions. • Operational Data: Information on major expenses including technology acquisition, research and development, and operational cost.

Module	Questions	Data Requirement
Indirect Economic Impact Module	• How do our geospatial services enhance the operations of intermediate industries, and can these enhancements be quantified in terms of output increase? • What is the total economic output generated by downstream industries as a result of integrating our geospatial data into their business processes? • Can we identify specific sectors or industries that significantly depend on our data for their operational decision-making? • What are the supply chain effects triggered by the demand for our geospatial products and services? • How do our products and services enable new market opportunities for other sectors or industries?	• Total spending on suppliers, categorized by type of goods or services (e.g., software, hardware, consultancy services). • Supplier Relationships: Information on key suppliers and the nature of their business relationships with the agency. The extent of dependency of suppliers on the agency's contracts. • Multiplier Effects: Multiplier coefficients to estimate how initial spending on suppliers generates further economic activity in the economy.
Induced Economic Impact Module	• How does the spending by employees in the geospatial sector contribute to local economic development? • What is the increase in consumer spending attributable to the wages paid by the geospatial sector? • Can we track the economic contributions of the geospatial sector to the local economy through household spending patterns? • What sectors benefit most from the induced economic activities stemming from the geospatial industry's payroll? • How does the increase in household income from geospatial industry employment affect local service demand, such as education and healthcare?	• Wage Disbursements: Total wages and salaries paid to employees, including bonuses and other compensation. • Consumer Spending Patterns: Data on how employees spend their income, potentially gathered through surveys or inferred from regional spending data. • Regional Economic Characteristics: Information on the economic characteristics of the region where employees live, which influences spending behaviors.
Value Added Analysis Module	• What is the total value added by the geospatial agency to the national GDP? • How do our products and services contribute to the enhancement of productivity across different economic sectors? • Can we measure the innovation boost provided by geospatial technologies in traditional sectors like agriculture or mining? • What portion of the national value added by our industry is attributable to advancements in geospatial technology? • How do improvements in geospatial data accuracy and availability translate into economic value for other industries?	• Production Data: Detailed data on the production processes, inputs, and outputs of the geospatial agencies. • Added Value Components: Contributions to GDP including wages, profits, taxes, and depreciation. Specific contributions from different stages of the geospatial data lifecycle. • Efficiency Metrics: Measures of productivity enhancements due to geospatial technologies, possibly comparing sectors or regions with and without such technological inputs.

^^ The above questions are indicative and require the expertise of a consultant, researcher or an economist to take forward the analysis.

Module	Questions	Data Requirement
Employment Impact Module	• How many direct jobs does the geospatial sector create annually? • What types of employment (e.g., technical, administrative) does our sector predominantly support? • Can we quantify the ripple effect of our employment on job creation in other industries? • What are the skill development and training opportunities created by the geospatial sector? • How does employment in the geospatial sector compare in stability and wages to national averages across sectors?	• Job Creation Data: Detailed counts of jobs created both directly and indirectly, including temporary and permanent roles. • Sectoral Employment: Breakdown of how employment is distributed across different sectors due to the agency's activities. • Skill Level Analysis: Data on the skill requirements and educational background of jobs created, assessing the role of the agency in promoting higher-skilled employment.

2.1 Time and Cost Savings Analysis

One of the important things for National Geospatial Agencies is to measure cost and time savings, and we can break down the process into distinct steps that each agency can follow systematically:

2.1.1 Time Savings Measurement

Phase 1: Identify Savings Categories
The first step involves categorizing the potential time savings to better understand and document where efficiencies are gained. These categories are:

- Internal Time Savings: Includes reductions in time spent on internal processes and queries, along with improved service delivery to citizens by Ministries, Municipalities, and other governmental bodies.
- Business Time Savings: Captures time savings for private sector stakeholders such as geospatial service providers, system integrators, and application developers.
- Consultant Time Savings: Accounts for the time reductions consultants achieve by utilizing more efficient geospatial data and tools.
- Citizen Time Savings: Reflects the direct time savings experienced by citizens through more efficient public services and interactions.

Phase 2: Data Collection
To quantify these savings, the data collection methodology employs both surveys and structured interviews to gather data from the identified stakeholders: ministries, businesses, consultants, and citizens. Additionally, time tracking tools and historical data comparisons can be utilized to accurately measure the actual time saved across various activities and interactions.

Phase 3: Quantitative Analysis
The third phase of time savings analysis involves converting the time data collected into monetary values, using average wage rates or assigning an economic value to time. This allows for a clearer assessment of the financial impact of time savings. The data is then aggregated to provide a summary of the total time saved annually for each category and across all stakeholders, highlighting the substantial benefits achieved.

Phase 4: Assess Impact
The methodology also examines how these time savings translate into broader benefits such as increased productivity and enhanced service delivery. It considers variables such as the organizational complexity and technological maturity of each entity to understand the varying impacts of time savings across different environments.

By following these phased steps, national geospatial agencies can systematically measure and articulate the significant time savings achieved through their initiatives, supporting continuous improvement and demonstrating the value of investments in geospatial technologies.

2.1.2 Cost Savings Measurement Methodology

To rigorously assess and report on cost savings for national geospatial agencies, a structured methodology must be implemented. This approach not only quantifies direct savings but also highlights the strategic and operational efficiencies achieved through enhanced geospatial services. The methodology is delineated into distinct phases:

Phase 1: Define Cost Components
The initial phase involves a detailed identification of the key areas where cost savings can be realized. These include:

- Data Infrastructure: Costs avoided by not having to establish and maintain separate geospatial data infrastructures.
- Process Enhancements: Savings accrued from improved processes that increase efficiency and reduce operational redundancies.
- Data Sharing Capabilities: Cost reductions resulting from enhanced data sharing among different stakeholders, which prevent duplication of efforts.

Phase 2: Estimate Baseline Costs
This phase focuses on estimating the costs organizations would incur if they had to independently establish their geospatial data infrastructure and processes. This involves calculating the investments required for building data platforms, validation methodologies, positioning networks, and other necessary components in the absence of services provided by the National Geospatial Agencies

Phase 3: Willingness to Pay
In this phase, the 'willingness to pay' method should be employed to determine the economic value of the services provided by the National Geospatial Agencies. Surveys and interviews should be conducted with organizations to capture data on what they would be willing to spend to establish and maintain their own geospatial services versus using those centrally provided. This approach helps in quantifying the perceived value of shared services and centralized data access.

Phase 4: Multiplier Effects
Using feedback from stakeholders, a multiplier effect should be applied to estimate the indirect impacts of cost savings. This includes savings from reusing data, reducing duplicative efforts, and achieving greater interoperability among different data systems. The multiplier effect quantifies how often data collected once can be reused, amplifying the cost savings across multiple applications and initiatives.

Phase 5: Synthesis and Reporting
The final phase involves synthesizing both direct and indirect cost savings into a comprehensive report. This report not only details the financial benefits but also underscores the operational improvements like reduced duplicity and the efficacy of the 'collect once, use many times' principle. The report also discusses the strategic advantages of centralized data management, highlighting its alignment with national standards and regulations, and illustrating its impact on national geospatial strategies and operations.

By meticulously following these phased steps, national geospatial agencies can effectively measure and articulate the substantial cost savings and operational efficiencies realized through their geospatial initiatives, fostering an environment of continuous improvement and strategic decision-making.

By integrating these different modules into a comprehensive economic impact assessment, geospatial agencies can articulate their value proposition to stakeholders and policymakers. These assessments are crucial for strategic planning, justifying funding allocations, and demonstrating the sector's role in economic development and public policy. Addressing data gaps is essential to ensuring the reliability of the assessments, enhancing stakeholder confidence in the results, and ultimately supporting informed decision-making.

2.5 Value of economic impact and its relationship with GKI for National Development

Geospatial Knowledge Infrastructure (GKI) plays a transformative role in national development by facilitating the integration of geospatial data and technologies into various economic sectors. The economic impacts of GKI are substantial and multifaceted, spanning from enhanced decision-making capabilities to innovation and the development of new markets. This document outlines the significant examples of GKI's impact at a national level, illustrating its value and relationship with broader economic and developmental goals.

The relationship between GKI and national development is deeply interlinked through the direct and indirect benefits that GKI delivers across multiple economic sectors. The integration of geospatial data into national strategies not only enhances operational efficiencies but also boosts strategic planning and governance, providing a robust foundation for sustainable growth. As countries increasingly recognize the importance of spatial data, investments in GKI can lead to substantial economic dividends, enhancing national competitiveness on a global scale. The ability to harness and effectively utilize geospatial data is quickly becoming a determinant of a nation's economic vitality and resilience, positioning GKI as a critical asset in the global knowledge economy.

Geospatial Knowledge Infrastructure (GKI) is, thus, increasingly recognized as a strategic asset for national development, enhancing economic performance, fostering innovation, and improving governance across various sectors. The value created by GKI extends well beyond traditional mapping functions, influencing key areas such as urban planning, disaster management, agriculture, and environmental protection. This expanded discussion further elaborates on the multifaceted impacts of GKI, with additional national-level examples to illustrate its profound influence on economic and developmental outcomes.

Enhanced Decision-Making and Resource Allocation

At the national level, Geospatial Knowledge Infrastructure (GKI) significantly enhances decision-making processes and resource allocation strategies. For instance, the United States Geological Survey (USGS) employs geospatial data to effectively manage water resources. Through this approach, the USGS can forecast floods, monitor droughts, and oversee irrigation systems, thereby preventing potential damage and preserving agricultural yields worth billions of dollars. Similarly, in countries like Singapore, urban planners leverage GIS technologies to optimize land use and infrastructure development. Given Singapore's dense population, integrating geospatial data into planning processes has enhanced urban sustainability and resource efficiency, resulting in considerable economic savings and improved living conditions.

Geospatial data and knowledge, and its mainstreaming within sectoral applications, elevates national decision-making capabilities by providing precise, real-time data to support complex planning and governance tasks. In Germany, for example, the government utilizes geospatial data to manage its energy grid, particularly in integrating renewable energy sources more efficiently. This optimization minimizes energy wastage and ensures a balanced supply and demand, significantly impacting energy economics and sustainability objectives.

Another compelling example is South Korea's use of geospatial knowledge for urban planning. The government employs advanced GIS technologies to efficiently manage land use, transportation planning, and disaster readiness, particularly in response to urban expansion and associated socio-economic challenges. This strategic utilization of geospatial knowledge has enabled South Korea to enhance infrastructural efficiencies and urban resilience, fostering economic stability and an improved quality of urban life.

Boosting Efficiency and Productivity

Geospatial knowledge significantly enhances efficiency and productivity across various sectors, including transportation, logistics, agriculture, and public services. Companies and governments leverage geospatial technologies to optimize operations and streamline processes, leading to tangible benefits in economic performance.

In transportation and logistics, organizations utilize geospatial data and knowledge to optimize delivery routes and manage fleet operations effectively. For instance, the implementation of the e-kakashi platform in Japan exemplifies the transformative impact of GKI. By integrating artificial intelligence (AI) and Geographic Information Systems (GIS), e-kakashi provides precise farming advice, enabling optimal planting, watering, and harvesting practices. This technology not only increases crop yields but also reduces resource waste, thereby directly impacting national agricultural productivity and economic returns.

Furthermore, GKI's influence on economic efficiency is evident in sectors such as logistics, agriculture, and public services. Major companies like UPS and FedEx in the United States leverage GIS for route optimization, resulting in reduced fuel costs and delivery times. This optimization directly enhances their bottom line and service quality, illustrating the economic benefits of GKI implementation in logistics operations.

In agriculture, GKI plays a pivotal role in precision farming, particularly in countries like Australia and Canada. Farmers utilize satellite imagery and GIS data to monitor crop health, optimize resource utilization, and maximize yield per hectare. This strategic use of GKI not only improves agricultural productivity but also contributes to sustainable farming practices, ensuring long-term viability and economic prosperity.

Overall, the impact of geospatial knowledge on economic efficiency and productivity is profound, spanning multiple sectors and driving tangible improvements in operational performance and resource utilization. As organizations continue to harness the power of geospatial technologies, the potential for enhancing economic returns and fostering sustainable development remains significant.

Innovation and Development of New Markets

Geospatial technologies are driving innovation and catalyzing the creation of new markets worldwide. In India, the government's Bhuvan initiative stands as a prime example. By providing satellite imagery and GIS data freely, Bhuvan has sparked innovation across various sectors, ranging from agriculture to disaster management. This initiative has not only enhanced India's technological capabilities but has also spurred the emergence of new businesses and services that leverage the freely available data. As a result, Bhuvan has contributed to economic growth and employment by fostering a vibrant ecosystem of innovation and entrepreneurship.

Similarly, in Africa, startups like Zipline are leveraging GIS and drone technology to revolutionize healthcare logistics. By using drones to deliver medical supplies to remote areas, these startups are not only addressing critical healthcare needs but also stimulating local economies and technology sectors. This innovative approach not only improves access to healthcare but also creates new economic opportunities and fosters technological advancement in the region.

Furthermore, geospatial data is driving innovation in water management in countries like the Netherlands. Given the country's unique geographical challenges with flooding and sea-level rise, geospatial technologies play a crucial role in developing innovative solutions. By harnessing geospatial data, the Netherlands has been able to implement water management innovations that mitigate the risks associated with flooding and ensure the sustainable use of water resources. This not only safeguards communities and infrastructure but also creates opportunities for economic growth and environmental sustainability.

Thus, geospatial technologies serve as catalysts for innovation and the creation of new markets. Initiatives like this as demonstrate the transformative impact of geospatial data in driving economic growth, fostering entrepreneurship, and addressing pressing societal challenges. As countries continue to embrace geospatial technologies, the potential for innovation and market creation will only continue to expand, unlocking new opportunities for sustainable development and prosperity.

Sustainable Development and Environmental Management

The role of Geospatial Knowledge Infrastructure (GKI) in sustainable development and environmental management is indispensable, with significant implications for global conservation efforts. In Brazil, GKI plays a pivotal role in the preservation of the Amazon rainforest, one of the world's most critical ecosystems. Through the utilization of satellite imagery and Geographic Information Systems (GIS), Brazil monitors the Amazon, combating illegal deforestation and managing this vital global resource. The economic ramifications of this endeavour are profound, as the health of the Amazon directly impacts multiple sectors crucial to Brazil's economic stability and development, including agriculture, biodiversity, tourism, and carbon sequestration.

Similarly, Canada leverages GKI to monitor forest fires and manage its extensive natural resources, contributing to sustainable forestry practices and disaster mitigation efforts. By employing geospatial technologies, Canada enhances environmental preservation and fosters economic sustainability, ensuring the long-term viability of its natural ecosystems and resource-dependent industries.

In the Philippines, Geospatial Knowledge plays a crucial role in coastal management and protection against natural disasters, particularly in a nation prone to typhoons and coastal erosion. Through the use of geospatial data, the Philippines implements effective strategies to safeguard lives, property, and coastal ecosystems. Additionally, Geospatial knowledge supports the tourism industry, a major economic driver in the Philippines, by preserving coastal areas and enhancing their resilience to natural hazards.

GKI serves as a powerful tool for sustainable development and environmental management, offering invaluable insights and capabilities to protect and preserve the planet's natural resources. Through initiatives in Brazil, Canada, the Philippines, and beyond, GKI contributes to global efforts to address environmental challenges and promote economic sustainability, ensuring a prosperous and resilient future for all.

Thus, the diverse and profound impacts of Geospatial Knowledge Infrastructure (GKI) on national development are vividly demonstrated through expanded examples spanning countries such as Germany, South Korea, the United States, Australia, Canada, Africa, the Netherlands, and the Philippines. By enhancing decision-making processes, improving economic efficiency, fostering innovation, and advancing sustainable development initiatives, GKI emerges as a cornerstone of modern national infrastructure.

Strategic implementation of GKI is paramount for nations aiming to bolster economic resilience, compete effectively on a global scale, and ensure the well-being of their citizens. By enabling more informed decision-making, enhancing operational efficiency, driving innovation, and promoting sustainable practices, geospatial knowledge serves as a foundational technology that not only supports economic growth but also enhances the overall quality of life for citizens.

As nations navigate evolving challenges and opportunities, the strategic integration of GKI will be indispensable for achieving sustainable economic resilience and maintaining competitiveness in the global arena. By leveraging the power of geospatial technologies, countries can navigate complex socio-economic landscapes with agility and foresight, ensuring continued progress and prosperity for generations to come.

2.6 References

1. World Bank Group & Global Geospatial Information Management. (2022, April). IGIF Socio-economic Impact Assessment Template [PDF]. Retrieved from https://www.wbgkggtf.org/sites/kggtf/files/2022-04/3.%20IGIF%20Socio-economic%20Impact%20Assessment%20Template%20Final.pdf
2. Oxera. (2018). What is the economic impact of geospatial services [PDF]. Retrieved from https://www.oxera.com/wp-content/uploads/2018/03/What-is-the-economic-impact-of-Geo-services_1-1.pdf
3. Oxera. (2018, March). Advancing economics in geospatial business [PDF]. Retrieved from https://www.oxera.com/wp-content/uploads/2018/03/Geo-services.pdf
4. Crompvoets, J., Rajabifard, A., van Loenen, B., & Delgado, T. (2015). Economic Assessment of the Use Value of Geospatial Information. ISPRS International Journal of Geo-Information, 4(3), 1142. https://doi.org/10.3390/ijgi4031142
5. Access Partnership. (2017, September). Geospatial Report [PDF]. Retrieved from https://accesspartnership.com/wp-content/uploads/2023/03/GeoSpatial-Report_Sept-2017.pdf

6. United Nations Economic Commission for Europe. (2022). WPLA webinar: Socio-economic benefits of geospatial [PDF]. Retrieved from https://unece.org/sites/default/files/2022-12/WPLA%20webinar%20Socio-economic%20benefits%20geospatial%20Report.pdf
7. OPTICS Project. (2017). Societal Impact Assessment Methodology (Version 2.0) [PDF]. Retrieved from https://www.optics-project.eu/optics1/wp-content/uploads/2017/07/OPTICS-D1.4-Societal-Impact-Assessment-Methodology_2.0.pdf
8. Geospatial World. (2018). GeoBuiz Report 2018. Retrieved from https://geobuiz.com/geobuiz-2018-report.html
9. Geospatial World's GeoBuiz. (2019). GeoBuiz Report 2019. Retrieved from https://geobuiz.com/geobuiz-report-2019/
10. Geospatial World. (2023). GeoBuiz 2023 Report. Retrieved from https://geospatialworld.net/consulting/reports/geobuiz/2023/index.html
11. Mahara, Gehendra, Kun Yang, Sipeng Chen, Wei Wang, and Xiuhua Guo. 2018. "Socio-Economic Predictors and Distribution of Tuberculosis Incidence in Beijing, China: A Study Using a Combination of Spatial Statistics and GIS Technology" Medical Sciences 6, no. 2: 26. https://doi.org/10.3390/medsci6020026
12. Igor Musikhin & Alexander Karpik (2023) Use of GIS technology and cellular automata for modeling multiple socio-economic scenarios of regional spatial development and inter-regional cooperation, Geo-spatial Information Science, 26:1, 71-93, DOI: 10.1080/10095020.2023.2182237
13. Amoako Johnson, F., Hutton, C.W. (2018). A Geospatial Analysis of the Social, Economic and Environmental Dimensions and Drivers of Poverty in South-West Coastal Bangladesh. In: Nicholls, R., Hutton, C., Adger, W., Hanson, S., Rahman, M., Salehin, M. (eds) Ecosystem Services for Well-Being in Deltas. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-71093-8_21
14. Wang, Fahui. (2014). Quantitative Methods and Socio-Economic Applications in GIS. 10.1201/b17967.
15. Bernknopf, Richard, and Carl Shapiro. 2015. "Economic Assessment of the Use Value of Geospatial Information" ISPRS International Journal of Geo-Information 4, no. 3: 1142-1165. https://doi.org/10.3390/ijgi4031142
16. Bose, P., Pattnaik, B. K., & Mittal, M. (2001). Development of socio-economic impact assessment methodology applicable to large water resource projects in India. International Journal of Sustainable Development & World Ecology, 8(2), 167–180. https://doi.org/10.1080/13504500109470073
17. Ladykova, Tatyana, Tatyana Ivanova, Viktor Streltsov, Anatoly Sayganov, and Ekaterina Ladykova. 2023. "Methodological Approach to Assessment Socio-Economic Development of the Chuvash Republic (Subject of the Russian Federation)" Economies 11, no. 7: 176. https://doi.org/10.3390/economies11070176
18. Balzekiene, Aiste & Butkeviciene, Egle & Telesiene, Audrone. (2008). Methodological framework for analyzing social impact of technological innovations. Socialiniai mokslai = Social Sciences. ISSN 1392-0758. 1. 71-80.
19. Geospatial Knowledge Infrastructure Readiness Index and value proposition in world economy, society and environment (Report) (2022), Geospatial World.
20. Geospatial World and UNSD (2021). The power of where? A geospatial knowledge infrastructure to enhance the world economy, society and environment. GKI White Paper. https://geospatialmedia.net/pdf/GKI-White-Paper.pdf
21. IRCC, IIT Bombay. (2002, September). Analytics. Retrieved from https://www.ircc.iitb.ac.in/IRCC-Webpage/PDF/update/archives/sept_2002/analytics.html
22. Zimmermannova, Jarmila, Andreas P. Redecker, Michal Mensik, and Carsten Juergens. 2021. "Geospatial Data Analysis and Economic Evaluation of Companies for Sustainable Business Development—An Interdisciplinary Teaching Approach" Sustainability 13, no. 20: 11245. https://doi.org/10.3390/su132011245

LESSON – 3

Live data streams in Geospatial Knowledge Infrastructure (GKI)

Trainer's Biodata: Prof Josef Strobl

Prof Josef Strobl is the founder and former head of University of Salzburg's Department of Geoinformatics. He holds degrees in Geography from Vienna University and has been teaching GIScience and related subjects at universities worldwide. He is a full member of the Austrian Academy of Sciences and lead of its Geographic Information Science initiatives. Josef Strobl continues to be involved in geospatial curriculum development projects and the practice of teaching and learning in online and blended environments, contributing to the capacity building 'brainware' component of a Digital Earth.

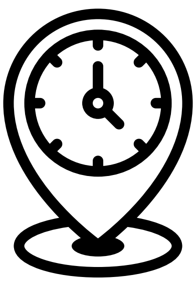
3.1 Introduction

Live data streams in the geospatial domain refer to the continuous flow of geospatial data generated from various sources, such as sensors, GPS devices, and social media feeds. These data streams are often massive in volume and are generated in real time, requiring specialized tools and techniques for processing, management, and analysis. There is a growing need for geospatial data stream management, which includes real-time monitoring, alerting, and long-term analysis of processed data. This is particularly important in applications such as geo-sensor networks, moving objects tracking, real-time traffic analysis, and environmental data management.

Figure 1

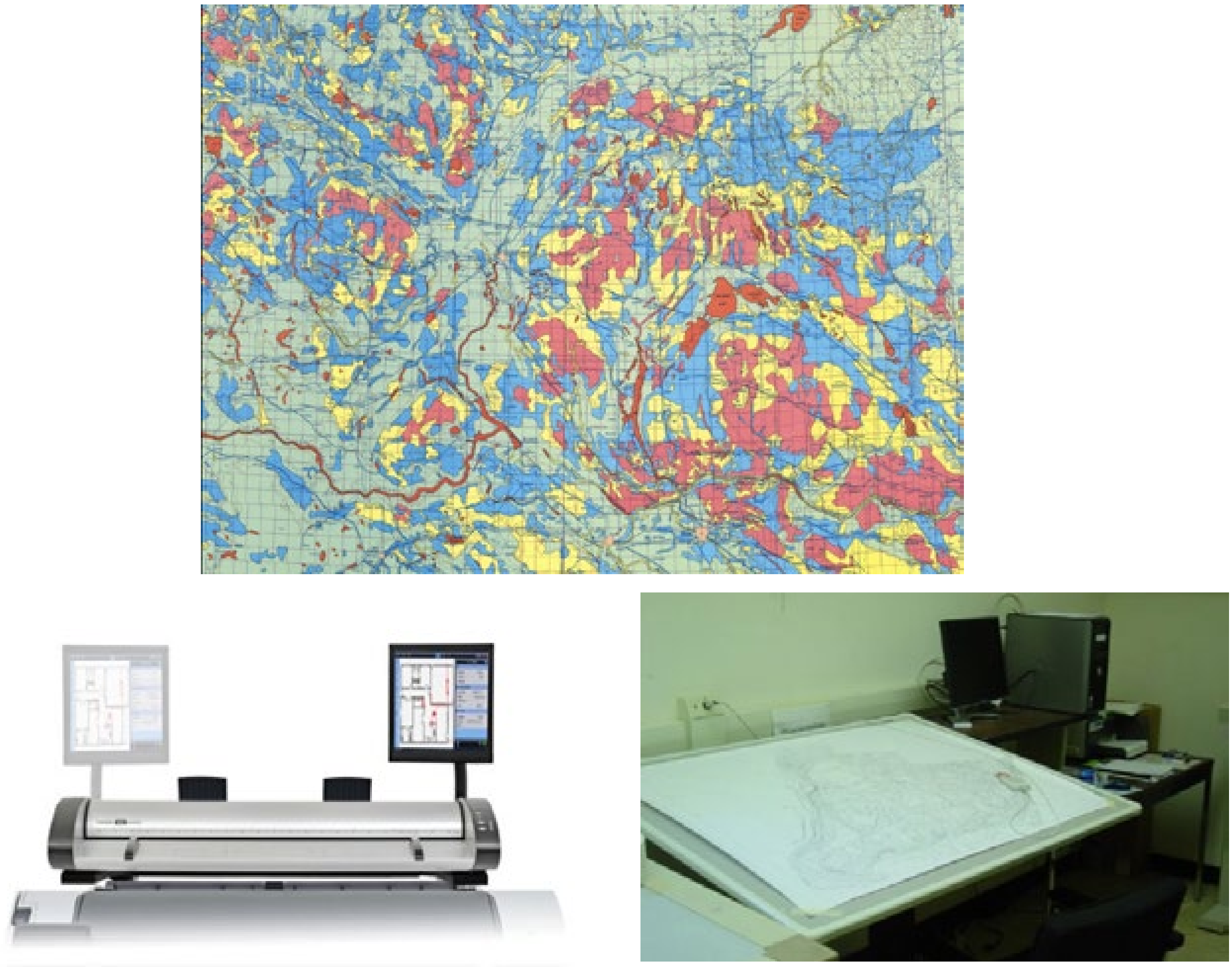
Science of Where

Awhereness



Real-time GIS data models and sensor web service platforms are essential for managing real-time geospatial data streams. Geospatial data stream management systems often rely on integrating and consolidating heterogeneous sensor data streams, which can be challenging due to the diversity of data sources, formats, and protocols.

Figure 2



3.2 Evolution from static data to data streams to knowledge

The evolution from static geospatial data to data streams to knowledge represents a significant progression in how we collect, analyze, and derive insights from geospatial information.

Figure 3

- **Static Geospatial Data**: Initially, geospatial data was primarily static, meaning it was collected at a specific point in time and represented a snapshot of a particular geographic area. This data includes maps, satellite images, aerial photographs, and other forms of spatial information. While valuable, static geospatial data has limitations in providing real-time insights and responding to dynamic changes in the environment.
- **Data Streams**: With advancements in technology and the proliferation of sensors and IoT devices, we entered the era of data streams. Data streams consist of continuous, real-time data generated by sensors, GPS devices, satellites, social media, and other sources. These streams provide a constant flow of information about various phenomena such as weather patterns, traffic conditions, population movements, and environmental changes. Analyzing data streams allows for the detection of patterns, anomalies, and trends as they occur, enabling more timely decision-making and response to events.
- **Knowledge Generation:** The ultimate goal of processing geospatial data streams is to derive actionable knowledge. By applying advanced analytics techniques such as machine learning, artificial intelligence, and spatial modeling to continuously streaming geospatial data, we can extract meaningful insights, identify patterns, predict future events, and make informed decisions. This knowledge can be utilized in a variety of domains including urban planning, disaster management, transportation, agriculture, and environmental conservation.

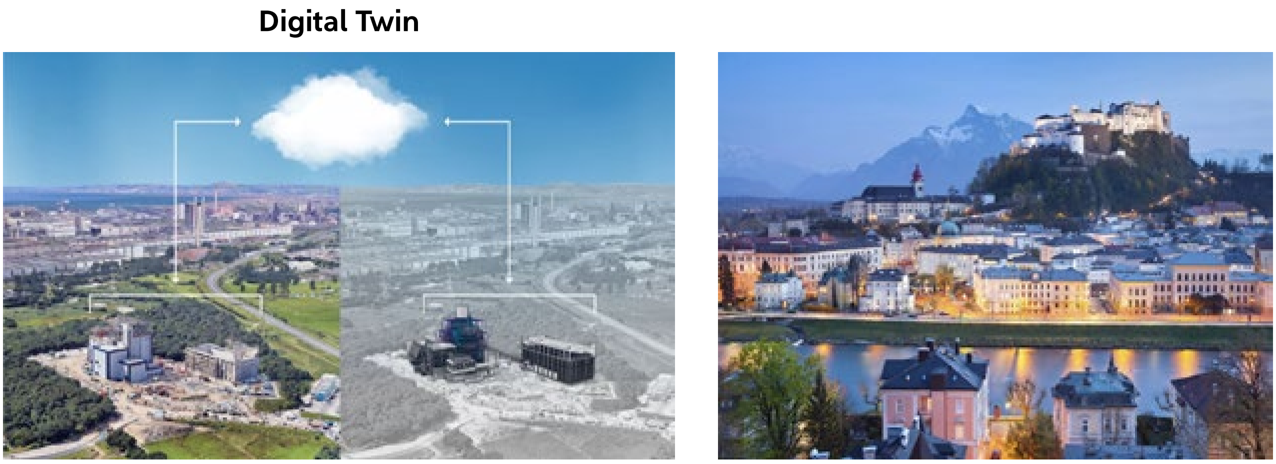
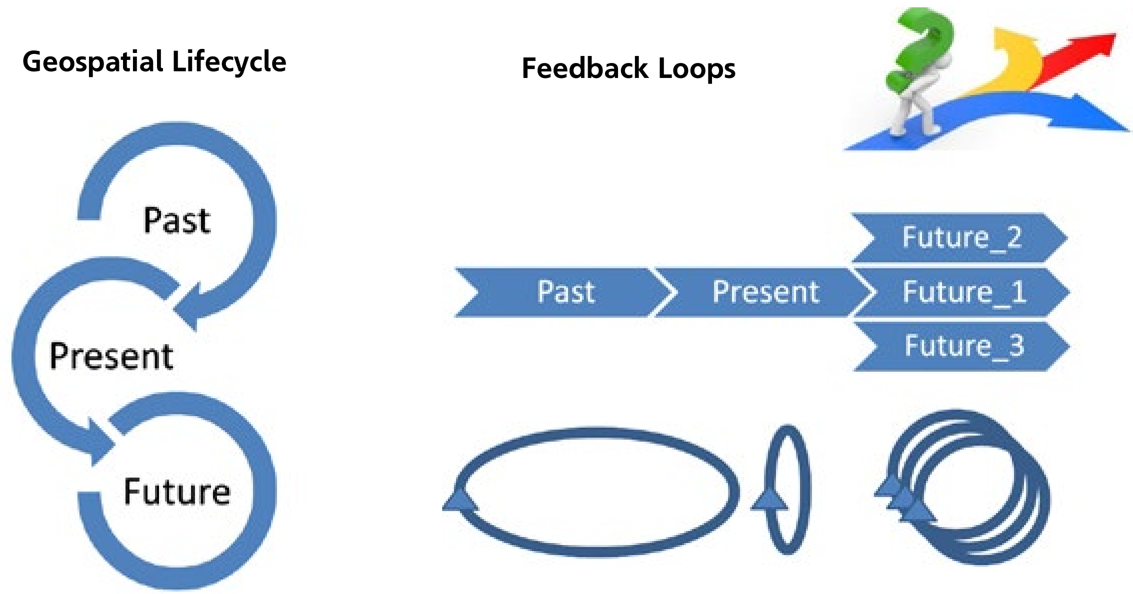


Figure 4

This evolution represents a shift from passive observation to active engagement with geospatial data, empowering organisations and individuals to harness the power of real-time information for improved planning, decision-making and resource allocation.



3.3 Role of real-time data in providing actionable context

Real-time data plays a crucial role in providing actionable context by offering up-to-date information about various phenomena, enabling organizations and individuals to make timely and informed decisions. Here's how real-time data contributes to actionable context:

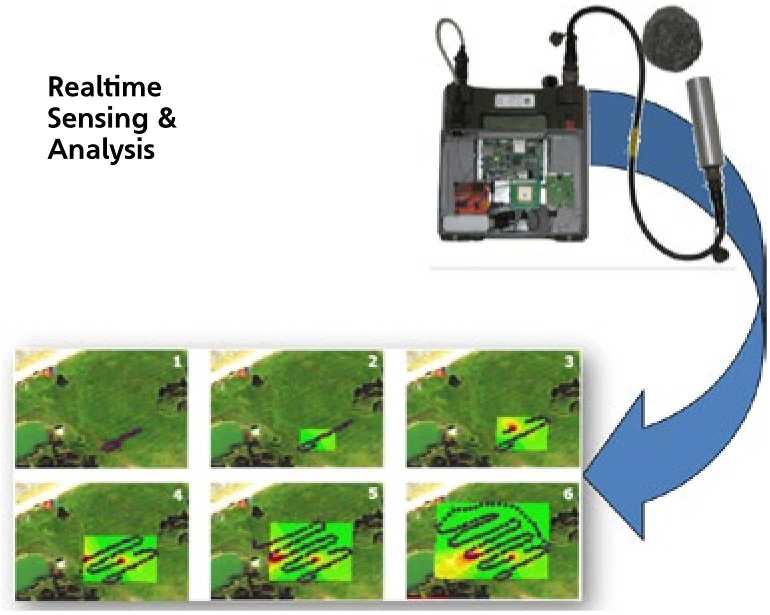


Figure 6

- **Timeliness:** Real-time data provides information as events unfold, allowing decision-makers to respond promptly to changes in the environment. Whether it's traffic conditions, weather patterns, or market trends, having access to the most recent data ensures that decisions are based on the latest information available.
- **Accuracy:** Real-time data offers a more accurate representation of current conditions compared to historical or static data. By capturing information as it happens, real-time data minimizes the risk of relying on outdated or incomplete information, thus enhancing the accuracy of decision-making.
- **Contextual Awareness:** Real-time data provides context by capturing the current state of affairs in a specific geographic area or domain. This contextual awareness helps decision-makers understand the situation comprehensively and make decisions that are tailored to the specific circumstances at hand.
- **Predictive Insights:** By analyzing real-time data streams over time, patterns and trends can be identified, enabling predictive insights. These insights allow organizations to anticipate future events and trends, enabling proactive decision-making rather than reactive responses.

Figure 7

- **Operational Efficiency:** Real-time data enables organizations to optimize their operations by identifying bottlenecks, inefficiencies, and opportunities for improvement in real-time. This leads to more efficient resource allocation, better utilization of assets, and overall improved performance.
- **Risk Management:** Real-time data allows for the timely identification and mitigation of risks. Whether it's detecting anomalies in financial transactions, monitoring changes in environmental conditions, or tracking the spread of diseases, real-time data enables proactive risk management strategies to be implemented.

In essence, real-time data transforms information into actionable insights by providing timely, accurate, and contextually relevant information that enables decision-makers to respond effectively to changing conditions and make informed choices.

3.4 Technology enabling real-time observations

Several technologies enable real-time observations across various domains. These technologies utilize sensors, data processing algorithms, communication networks, and visualization tools to capture, analyze, and present real-time data. Here are some key technologies enabling real-time observations:

- **Internet of Things (IoT):** IoT devices are equipped with sensors that collect data from the physical environment and transmit it over the internet in real-time. These devices can monitor various parameters such as temperature, humidity, air quality, motion, and more. IoT enables real-time monitoring and management of infrastructure, buildings, transportation systems, and environmental conditions.

Figure 8

- **Satellite and Remote Sensing:** Satellites equipped with remote sensing instruments capture high-resolution images and data of the Earth's surface in real-time. Remote sensing technologies, including radar, LiDAR, and multispectral imaging, provide valuable information about land use, vegetation, water bodies, weather patterns, and natural disasters. These observations are crucial for environmental monitoring, agriculture, disaster response, and urban planning.

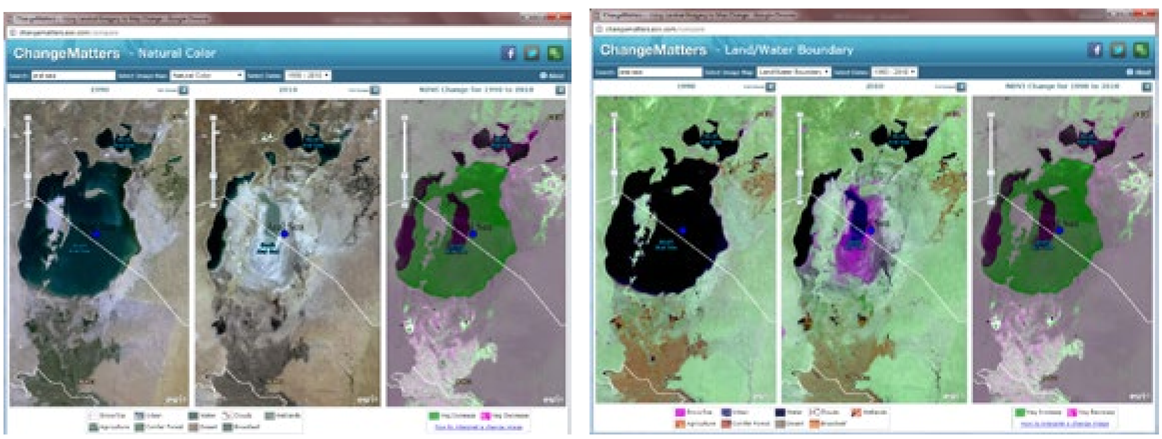


Figure 9

- **Wireless Sensor Networks (WSN):** WSN consists of spatially distributed sensors that communicate wirelessly to collect and transmit data in real-time. These networks are deployed in various applications such as environmental monitoring, structural health monitoring, healthcare, and smart agriculture. WSN enables continuous monitoring of physical phenomena and facilitates real-time decision-making based on the collected data.

Sensors, Sensors Everywhere

Figure 10

- **Geographic information Systems (GIS):** GIS technology integrates spatial data from multiple sources and allows users to visualize, analyze, and interpret real-time geospatial information. GIS platforms enable the creation of interactive maps, dashboards, and spatial analytics tools for monitoring and managing assets, infrastructure, and natural resources in real-time.

Figure 11

- **Machine Learning and Artificial Intelligence:** Machine learning and AI algorithms are utilized to process and analyze large volumes of real-time data streams, identify patterns, detect anomalies, and make predictions. These technologies enable automated decision-making, adaptive control systems, and intelligent data-driven insights in various domains, including finance, healthcare, manufacturing, and transportation.

Social Sensing - "Urban Emotions"

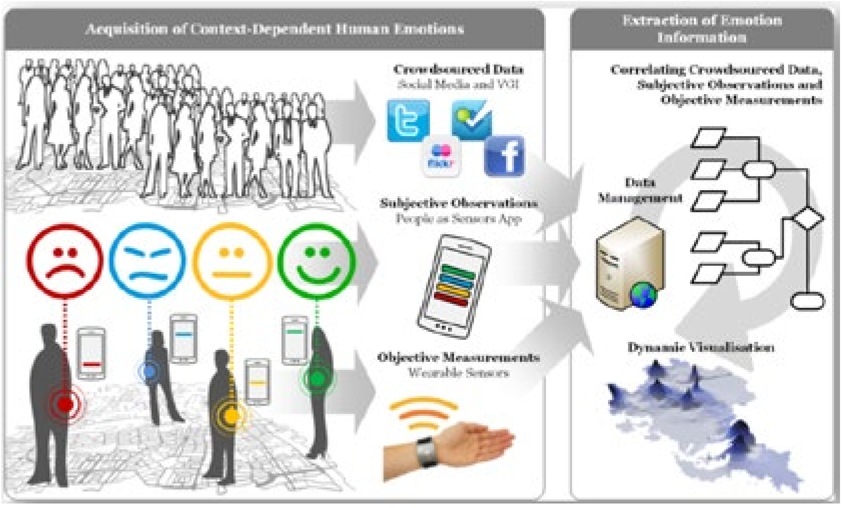


Figure 12

- **Cloud Computing:** Cloud computing platforms provide scalable infrastructure and storage solutions for processing and analyzing real-time data streams. Cloud-based services offer real-time data processing capabilities, enabling organizations to deploy distributed computing resources and perform complex analytics tasks on streaming data in a cost-effective manner.

Cloud Computing

Figure 13

These technologies collectively empower organizations and individuals to observe, monitor, and analyze the world in real-time, leading to enhanced situational awareness, improved decision-making, and better outcomes across diverse fields and industries.

3.5 Integration of real-time data for knowledge creation to support decision-making

The integration of real-time data for geospatial knowledge creation is instrumental in supporting decision-making across various domains. Here's how the process unfolds:

Figure 14

- **Data Acquisition:** Real-time geospatial data is collected from sources such as IoT devices, satellites, sensors, and social media feeds. These data sources capture information about the physical environment, including weather conditions, traffic patterns, infrastructure status, natural phenomena, and human activities.

Figure 15

- **Data Processing and Analysis:** Real-time data streams are processed and analyzed using advanced geospatial analytics techniques. Machine learning algorithms, spatial analysis tools, and AI models are applied to identify patterns, trends, and anomalies in the data. This analysis helps extract meaningful insights and derive actionable knowledge from the real-time geospatial data.

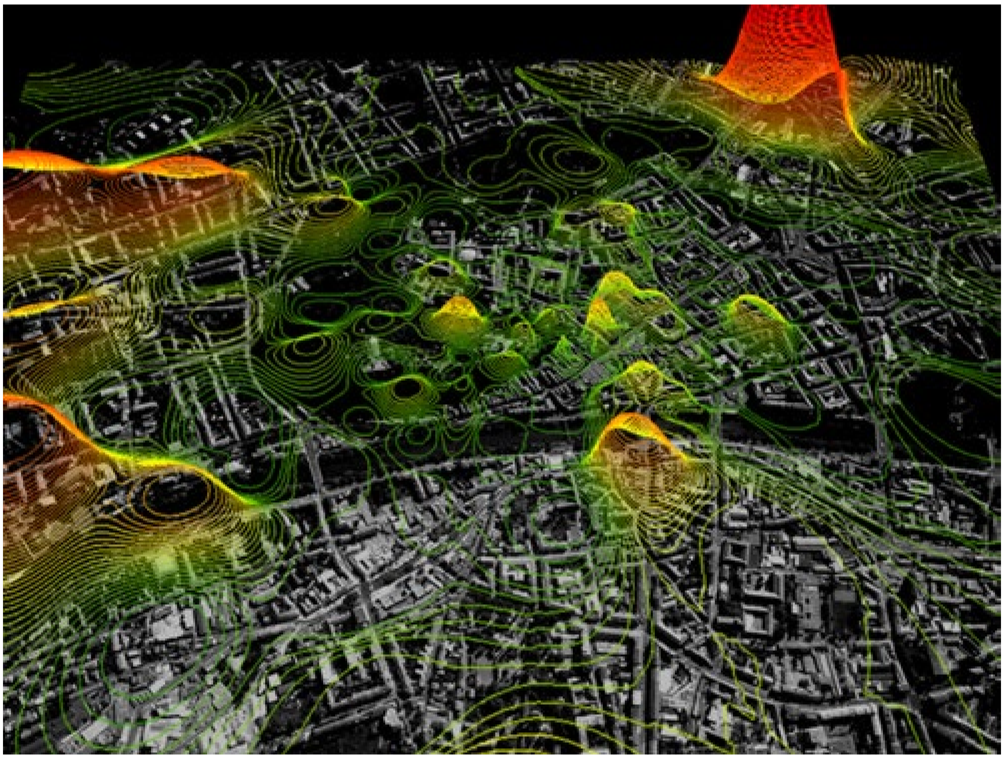


Figure 16

- **Contextualization:** Real-time data is contextualized by integrating it with historical data, static datasets, and domain knowledge. This contextualization helps to understand the significance of real-time observations in relation to past trends, existing conditions, and expected behaviors.
- **Knowledge Creation:** Real-time data is used to create actionable knowledge by identifying patterns, trends, correlations, and causal relationships. This knowledge may take the form of spatial analysis, temporal modeling, predictive analytics, or geospatial visualization.

Figure 17

- **Decision Support Systems (DSS):** The knowledge derived from real-time data is integrated into decision support systems to assist decision-makers in making informed and timely decisions. DSS may include dashboards, alerts, reports, and visualization tools that provide stakeholders with actionable insights based on real-time observations.

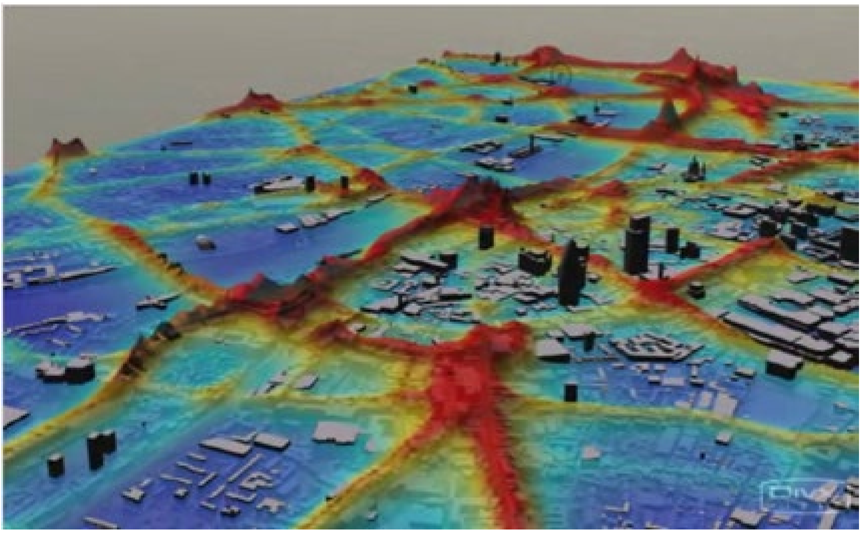


Figure 18

- **Scenario Analysis and Simulations:** Real-time data can be used to perform scenario analysis and simulation to evaluate different decision options and their potential outcomes. This allows decision-makers to assess the impact of their decisions in real-time and choose the most favorable course of action.

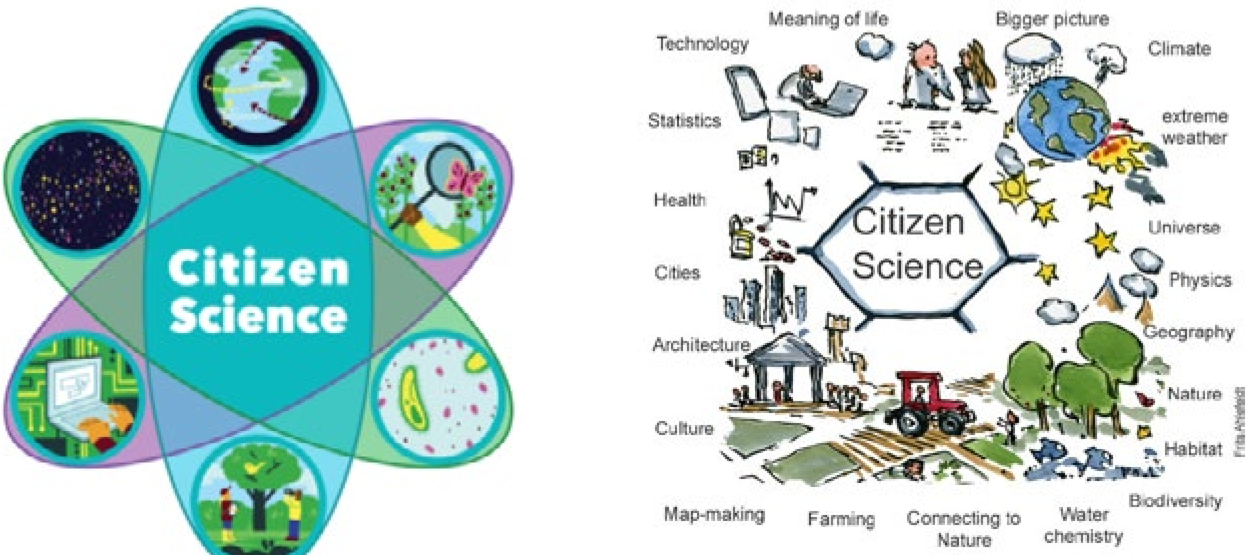


Figure 19

- **Feedback Loop and Continuous Improvement:** Decision-making based on real-time data generates feedback that can be used to refine models, update algorithms, and improve decision support systems over time. This iterative process ensures that decision-making capabilities evolve in response to changing conditions and new insights.

Figure 20

By integrating real-time data for knowledge creation, organizations can gain a deeper understanding of their operating environment, anticipate future trends, and make more

LESSON – 4

National Geospatial Portal

Trainer's Biodata: Alan Devenish

Alan Devenish is Esri Inc.'s technical advisor to the EU institutions, supporting agencies working on complex problems at the intersection of geography and policy. He has wide-ranging experience as a geographer, including research in cultural geography and expertise in earth observation. A certified technical trainer, he has helped hundreds of people build their geospatial skills. Currently based in Amsterdam, he supports a community of Esri Distributors across Europe in domains such as climate resilience and sustainable agriculture.

4.1 Introduction to National Geospatial Portals (NGPs) and their objectives

The mission of National Geospatial Portals (NGPs) is to connect people with easy-to-use tools and the curated data they need to solve problems. National Geospatial Portals are inherently works in progress, assimilating and distributing new sources of data, reaching an ever-growing community of users, spatially enabling new topical domains, and navigating evolving relationships with private industry, public agencies, and academia. However strong the strategic vision motivating a particular NGP, or clear its conceptual framing, users and stakeholders will judge each portal on their experience of it, in its current implementation.

Figure 1

Consumer expectations of web experiences in general have risen, as have map literacy and familiarity with information products that incorporate geographic information, whether those products be interactive web tools like dashboards or static infographics. To meet these expanding audiences and expectations, the potential of mapping and location data for enhancing storytelling is increasingly being realized. For example, major publications now commonly developing vertical storytelling experiences that incorporate mapping applications, video, and narrative text.

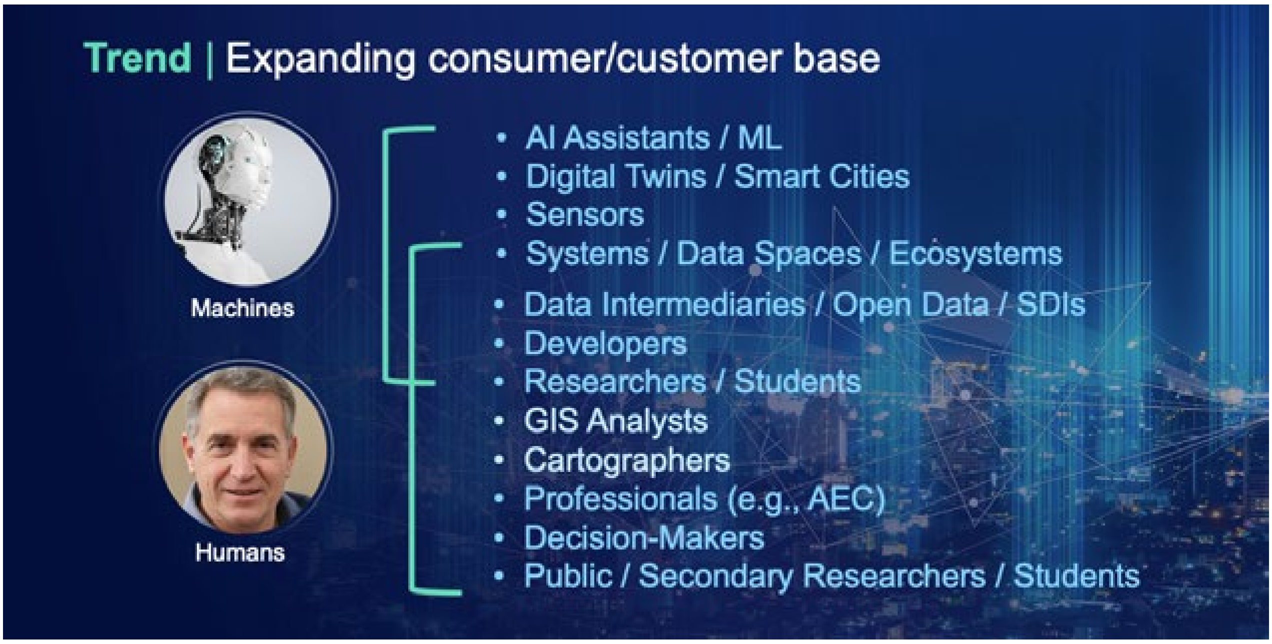


Figure 2

At the same time, massive datasets that are technically difficult to collect, maintain, and serve are increasingly available through a commercial entities, across an expanding range of sectors. Earth observation, demographic, and real-time/IoT data are driving novel geospatial applications and facilitating innovative research agendas.

To maintain audience-share across topics important to the public, and to meet requirements of the community of analysts, decision makers, and system architects that rely on current, high-quality geospatial data, successful national geospatial portals are embracing customer-centric thinking and positioning

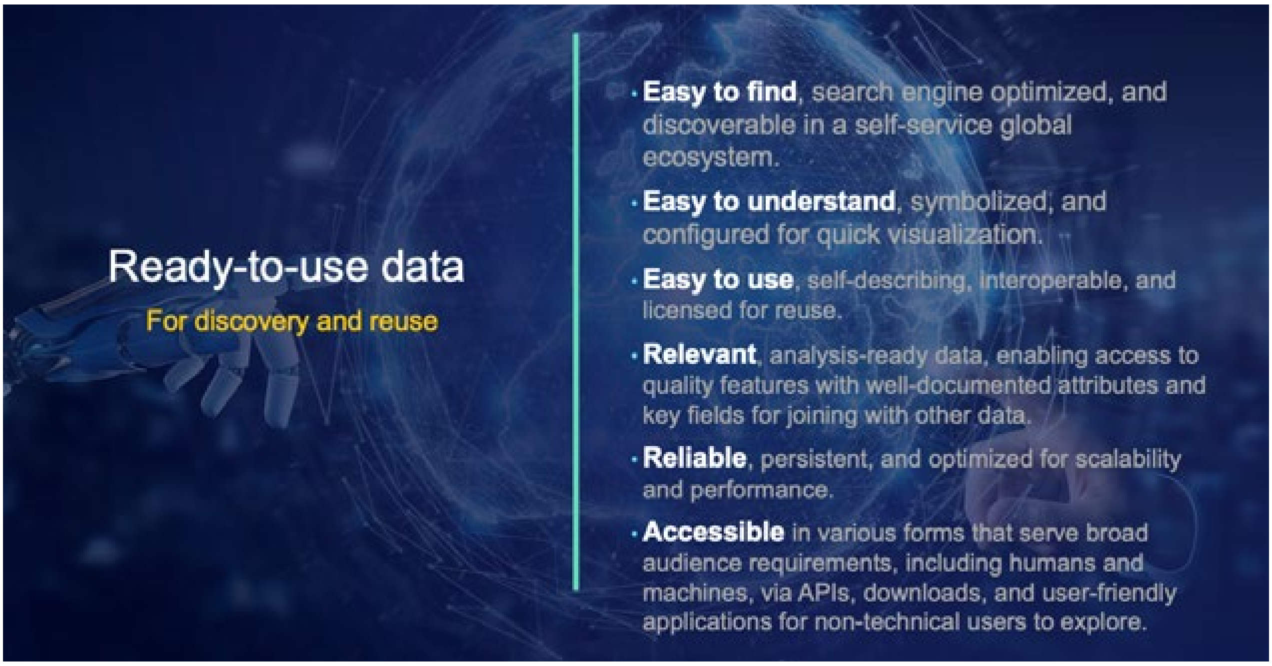


Figure 3

Once the domain of national mapping and statistics agencies, national geospatial portals now meet public expectations for in-depth treatment of critical priority topics across the sciences, public policy, and administration, providing knowledge infrastructure and content for better understanding. An interconnected network of partners collaborates to address common goals and shared challenges.

These priority topics include, among many others:

- Biodiversity, which can include global species migration data, local habitat data, harmonized national observation datasets, and accompanying data on riparian systems, point source pollution measurements, and relevant social and commercial data.
- Disaster risk and management, allowing for the straightforward overlay of parcel data, demographic layers, and critical infrastructure and asset location datasets with environmental data such as flood, landslide, and fire risk maps.
- Public health, which has expanded its use of social vulnerability data and has benefited from enhancements to data quality and analysis methodology following the COVID-19 pandemic.
- Tourism, with countries both promoting the tourism industry as a revenue driver, or monitoring and managing excessive tourist flows, depending on the region.
- Climate mitigation and adaptation, focused on socio-ecological domains, urban planning, and energy transition.
- Synthesized themes covering multiple topics, in the form of a set of national or globally recognized indicators, perhaps the most well-known being the Sustainable Development Goals (SDGs).

Beyond serving individual programs or delivering certain policy outcomes, National Geospatial Portals serve as the storefronts for a country's geographic thinking. To consumers and producers of geographic data, national geospatial portals may represent a national government's commitment to supporting modern GIS, signal depth of engagement on important topics, demonstrate technical capacity, and provide a model for good governance.

4.2 Importance of NGPs in providing access to geospatial data and services

NGPs occupy a unique role as the nationally authoritative destination for geospatial data, which are provided to support local to national agencies, inform the public, or fulfill treaty obligations. In a varied global landscape, technical maturity and funding resources might look very different from country to country, as will the capacity to develop best-in-class foundational data, but every country can apply customer-centric design principles to NGPs that will ensure their Portal is developing to meet modern GIS needs.

"Wicked problems" require the intergenerational transfer of institutional knowledge to ensure that they are sufficiently managed, if not ultimately resolved. Similarly, effective policymaking and long-term planning require consideration of how data will be collected, managed, and provided to facilitate the monitoring and evaluation (M&E) of key priorities (for example, the EU's European Data Strategy topic areas and European Green Deal topic areas are closely aligned). To ensure relevance for analysis and communication on such complex topics over the long-term, NGPs are built to be extensible in terms of capability, and flexible in terms of the data sources they can incorporate and the standards they support.

Modern GIS patterns of work depend on collaboration and interconnectedness, so NGPs are more than data catalogs or galleries for applications—they are systems of engagement for national governments. NGPs model data governance best practices for the communities they serve and ensure the smooth functioning of emerging integrated geospatial infrastructure at the subnational, national, supranational, and global levels.

4.3 Overview of the structure and components of NGPs

A well-architected NGP will enable the most relevant data to be quickly located and consumed by users with different levels of skill and across a variety of systems. Ready-to-use data layers that feed the public and private web applications will use data in place, dynamically indexing content directly from the scientific, agency, and technical authorities that developed them.

National Geospatial Portals provide both the engagement and delivery tiers, serving multiple kinds of users with a wide range of needs.

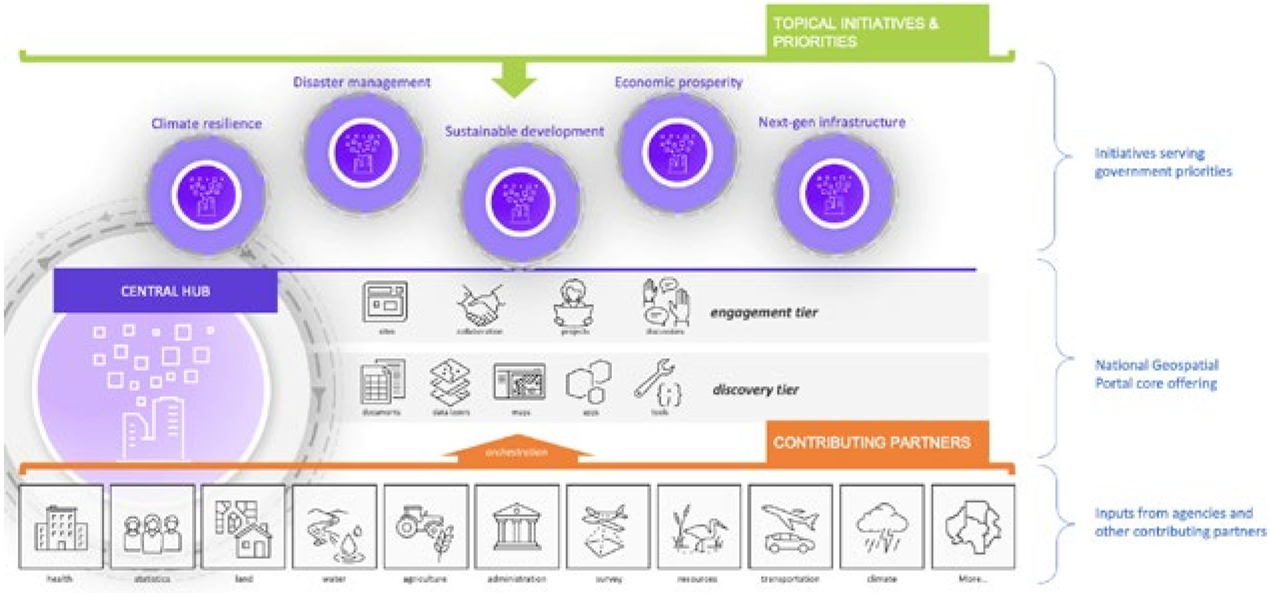


Figure 4

Tightly coupled metadata will be human and machine-readable, meeting requirements for discovery, evaluation, use, management, and future innovations, such as generative artificial intelligence (AI). Partner agencies and other consumers will leverage easy-to-use tools to filter, process, and display NGP data that are relevant for their needs, all while referencing the common, persistent data sources.

4.4 Key features of NGPs

Geospatial infrastructure is the enabler that powers National Geospatial Portals. Geospatial infrastructure manages security, privacy, and accessibility; identity management; spaces for sharing and collaboration; and content negotiation through open standards, APIs, licenses, and marketplaces. Like water and electricity within a physical infrastructure, rich and diverse content flow through geospatial infrastructure – conveyed by applications that bring those data to life. When organizations integrate their geospatial infrastructure, they can quickly and easily work beyond borders, jurisdictions, and sectors in purpose-driven collaborations.

The essential role of geospatial fundamental data themes as the framework or scaffolding of modern information infrastructure is still explicitly recognized. Thus, fundamental data concepts are folded into geospatial infrastructure, not abandoned.[1]

Key features may include:

- Country at-a-glance
- Easy search and discovery tools
- Featured topics
- Map exploration
- Data catalogue
- Collaboration center
- On-the-fly translation/localization
- Audience-centric learning resources

4.5 Real-world use cases

As discussed above, NGPs occupy a unique position within geospatial ecosystems, but their functional patterns of use is not dissimilar to those observed at other levels of government. In practice, innovation often flows from local or regional geospatial portals up to the national level.

We observe several repeatable implementation patterns, including:

- Foundational data initiatives

Figure 5

- Topical initiatives

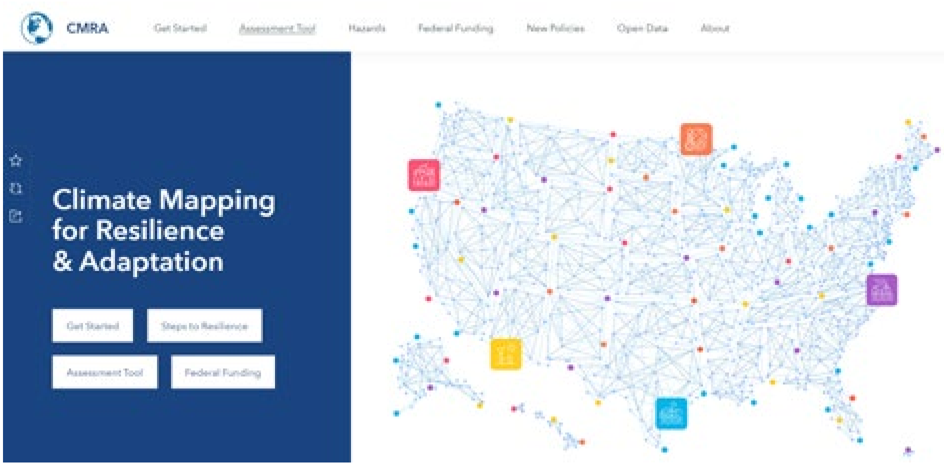


Figure 6

- National Spatial Data Infrastructur (NSDI)

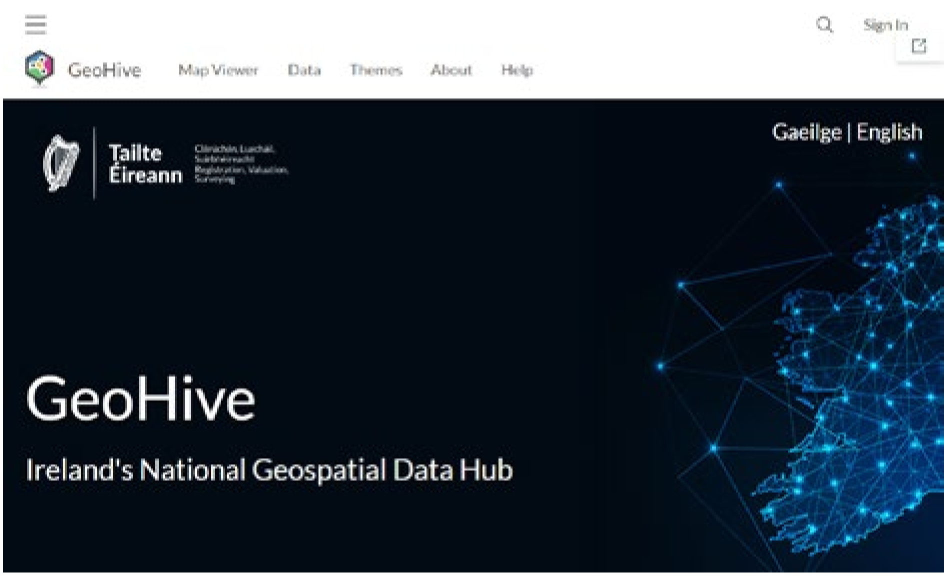


Figure 7

- Whole-of-government collaboration

Figure 8

Other common patterns include:

- Multi-level collaboration
- World Region Initiatives
- Marine and Arctic SDIs
- Regional planning initiatives
- Whole-of-business collaboration
- Sustainable Development Goals
- Emerging Hub networks

To explore these examples and better understand the most common patterns in practice, please use the following QR code or link [3].

4.6 References

1. [1] SDG Data Alliance, 2024. The UN-GGIM Fundamental Data Themes: Explore resources available through the ArcGIS Living Atlas of the World. https://storymaps.arcgis.com/stories/3a4625105caf4f6ca09a197eaea8612e
2. [2] Saligoe-Simmel, J., & Jordan, M. (2023). Working Beyond Borders: GIS for Geospatial Collaboration. Esri Press. [3] Esri. (2023, February 10). Essential guides for "OneMap" administrators. ArcGIS StoryMaps. https://storymaps.arcgis.com/collections/480c3a6b4bfd4f2984c3698558ca3549
3. [4] British Geological Survey. (n.d.). Designing Geospatial Data Portals. https://britishgeologicalsurvey.github.io/design-guide-for-geospatial-data-portals/

LESSON – 5

Data Access and Interoperability

Trainer's Biodata: Dr. Joan Maso Pau

Dr. Joan Maso Pau is a Principal investigator of CREAF leading specialized group on geospatial interoperability, GIS, remote sensing. Developer in C, JavaScript, Python. Active member of the Open Geospatial Consortium (OGC) since 2003. Editor of OGC 07- 057r7 WMTS standard, the new OGC 20-057 OGC API Tiles, and OGC 21-026 Cloud Optimized GeoTIFF. OGC Gardels gold medal in 2018. Active member of GEO (participant in the Data Working group and Chair of the Citizen Science Initiative). Expert in Metadata, Data Quality documentation, Geospatial data visualization, Citizen Science, In-situ data requirements and Green Deal Data Space. Earth and Space Science Informatics (ESSI) former division president in the European Geosciences Union (EGU). Senior member of IEEE and council member of ISDE.

5.1 Metadata, Open data, licenses and FAIR data

Geospatial metadata is defined as "data that provides information about other geospatial data". Geospatial metadata is crucial information about data resources, akin to a library catalogue record, detailing the who, what, when, where, how, and why of a data resource. In the same way that we would never ingest medicine or packaged food without reading the lab el describing the product, we should pay attention to the metadata before using any geospatial asset. Geospatial metadata describes maps, GIS files, imagery, and other location-based data resources.

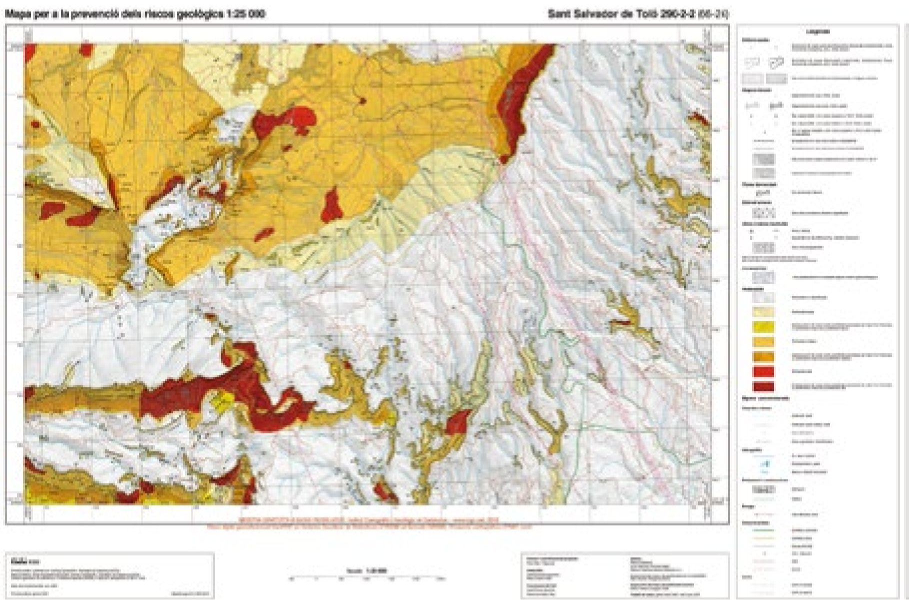


Figure 1: Geological risk prevention paper map. Around the actual data, we can read the title, the location name, the producer information, the scale and projection as well as a detailed meaning of the elements symbolized in the map.

The earliest printed maps had legends that described the source of the map, the producer, the date, and usually the meaning of the symbols used. Paper maps intended for serious users contain a wealth of metadata in their margins in a ver structured way: information including title and other identificiation information, scale, reference system, accuracy, timeliness, sources, producer, and extensive descriptions of the symbology. In the digital environment, the information in the margins of the map needs to be encoded in document that should accompany the data itself.

Today, we download maps and directions from the Internet, relying on them, perhaps more than we should, without knowing much about them – many times, we do not know their origin or how up to date or correct they are. Geospatial metadata standards, like ISO 19115, ensure uniformity in formatting for easy accessibility by both humans and machines. Metadata involves elements like identification, contact, quality, spatial reference, and entity and attribute details to provide a comprehensive understanding of the data resource. In an aerial photograph metadata should contain the geographic location of the center point, the focal length, flying height, date, and time. Without this information the aerial photo is just a picture. With metadata, the aereal picture can be transformed in a nice orthorectified map.

Metadata is used for a variety of applications; these can be classified into four primary functions: find, evaluate, access, and employ, which can be described in detail as follows:

1. Find: metadata enables users to locate geospatial resources and allows producers to advertise their data. Metadata helps organizations find data outside the organization and find partners to share in data collection and maintenance.
2. Evaluate: by having proper metadata elements describing a resource, users are able to determine its fitness for an intended use. Understanding the quality and accuracy, the spatial and temporal schema, the content, and the spatial reference system used allows users to determine whether a resource fulfills their needs. Metadata also provides the size, format, distribution media, price, and restrictions on use, which are also evaluation factors.
3. Access: after locating a resource and determining if it meets users' needs, metadata is used to describe how to access it and, depending on the resource, transfer it to a specific site. Once it has been transferred, users need to know how to process and interpret the data and incorporate it into their holdings.
4. Employ: metadata is needed to support the processing and the application of geospatial resources. Metadata facilitates proper utilization of data, allowing users to merge and combine data with their own, apply it properly, and have a full understanding of its properties and limitations.

We can classify the metadata elements in the following categories:

1. Identification: Identification provides a brief narrative of your data. In other words, it summarizes the purpose of your data in a succinct way. For example, identification assigns the following to your metadata: Title, Description and Keywords
2. Contact: The contact information includes details on who developed and made the data available. For example, it includes the following three entities: Originator, Publisher and Distributor has the main focus on making the data available.
3. Quality: Quality explains the accuracy and standards the data set followed. For example, it includes horizontal and vertical positional accuracy, completeness, integrity, and inspections of the data.
4. Spatial Reference: Spatial reference information assigns a geographic extent and coordinate system. Projection information includes a projection, datum, and units. For example, UTM Zones and State Plane are common coordinate systems. Geographic extent comes in the form of a bounding box, place keyword, or thumbnail.
5. Entity and Attribute: Entities refer to the map data type such as points, lines, polygons, or grids. For example, in a dataset representing cities, each city would be considered an entity. The purpose of this metadata item is to describe how to represent the spatial information in the data. The entity attributes include a description with a list of valid values and domains.
6. Lineage: Lineage describes the sequence of operations and processes that have been applied to the dataset. This could involve data transformation, projection changes, data cleaning, aggregation, or any other processing that alters the original data. Additionally, lineage describes who and when was the data processed. For example, it lists the processing steps and responsible parties. Each processing step has a date when it took place so users can track changes. It's like a changelog listing the evolution of the data from start to finish.

7. Legal: The legal section outlines the constraints for accessing and distributing the data. It describes the liability to assure the protection of privacy and intellectual property. Metadata includes a security classification that handles the restriction over security concerns. For example, confidential, restricted, sensitive, unrestricted, and unclassified are examples of security classification in metadata.
8. Temporal: It focuses on describing the time-related characteristics of the data, such as when the data was collected, the time range it covers, and any temporal attributes associated with the dataset. It not only focuses on when the data were collected but how long it's valid. Plus, it also states the progress such as when there will be future updates. For example, the frequency of updates can be anywhere from daily, weekly, monthly, or annually.
9. Metadata Reference: The metadata reference section is specific to the metadata. This is the individual or organization responsible for creating the metadata. It could be a person's name, a department within an organization, or any relevant entity. It also gives a point of contact when there are uncertainties such as how to cite information when used. The metadata reference has a temporal component for when the user created and will revisit it next.

The best practices for maintaining geospatial metadata standards include:

1. Consistently using standardized vocabularies and unique identifiers when curating metadata elements like keywords and data themes
2. Integrating metadata management into geospatial data management workflows and systems, such as within GIS software and online platforms
3. Regularly reviewing and updating metadata to ensure it remains accurate and up-to-date as the data resources evolve
4. Providing training and guidance to data publishers on best practices for creating high-quality, standards-compliant metadata
5. Adopting common international metadata standards such us the ISO 19115. These provide a comprehensive set of metadata elements to describe geospatial data resources.
6. Ensuring metadata is machine-readable and follows standardized formats like XML schemas to enable automated discovery and interoperability

The key benefits of using common geospatial metadata standards are that they enable the uniform formatting, accessibility, and interoperability of geospatial data resources for both humans and machines. The common geospatial metadata standards are:

1. ISO 19115: Geographic information - Metadata - This is the industry standard for geospatial metadata, developed by the International Organization for Standardization (ISO). It provides a comprehensive set of metadata elements to describe geographic data and services
2. FGDC Content Standard for Digital Geospatial Metadata (CSDGM) - This was the previous U.S. federal standard for geospatial metadata, developed by the Federal Geographic Data Committee (FGDC). It has now been harmonized with the ISO 19115 standard
3. Dublin Core Metadata Initiative (DCMI) standards - These provide a more general set of metadata elements for any type of resource (e.g. a book) that can also be applied to geospatial data.

The ISO 19115 standard defines a schema required for describing geographic information and services. The standard applies to all geographic resources – it is applicable to datasets in series, datasets, individual geographic features, and their attributes. It is also applicable to geospatial services. The standard defines the metadata required to serve the basic metadata functions such as resource discovery, determining a resource's fitness for use, and resource access, as well as optional metadata elements to support a more extensive description enabling a wide range of metadata applications.

ISO 19115-1 defines over 400 metadata elements to fulfill the requirements for finding, evaluating, accessing, and employing geospatial resources. These elements are defined in a data dictionary, based on ISO 11179 Information Technology – Specification and standardization of data elements, which provides a name, definition, obligation, maximum occurrence, data type, and domain for every element. It also provides a schema that defines their relationships using models based on the Object Management Group's (OMG)

Unified Modeling Language (UML), which is also an ISO standard: ISO 19501.

The UML model is an integral part of the entire abstract model for geographic information developed by ISO TC 211 (Chap. 1). The metadata model is presented in 15 packages that combine collections of metadata classes and identify the relationships between these collections. Each package contains one or more classes of metadata, which may be specialized (subclassed) or generalized (superclassed). Each class of metadata contains attributes, which identify the discrete units of metadata.

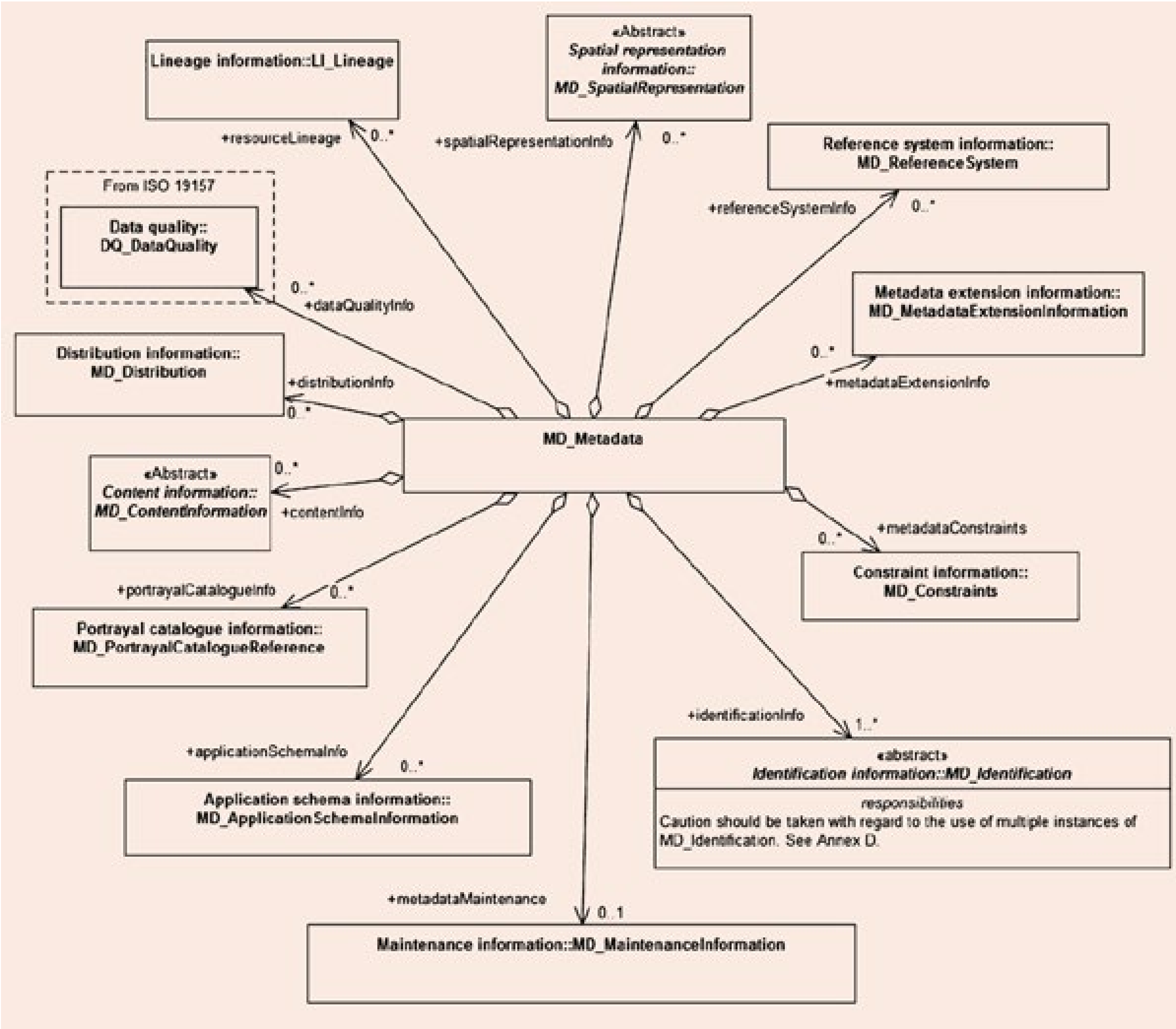


Figure 2. ISO 19115-1 metadata schema classes

Metadata specified in conformance with ISO 19115-1 consists of an MD Metadata class as the root class, which contains attributes (metadata) about the metadata and aggregates other metadata classes that provide metadata about geospatial resources. This UML model specifies metadata about a resource contains the following classes:

1. MD_Metadata – metadata about the metadata
2. MD_Identification – a description that fully identifies the resource, including the spatial extent of the resource and service metadata if the resource is a service.
3. MD_Constraints – identifying any legal or security, access or use restrictions associated with the resource
4. MD_ReferenceSystem – identification of the coordinate or noncoordinate reference system used
5. MD_SpatialRepresentation – a description of the geospatial information types used by the resource to digitally represent the Earth's features
6. MD_Lineage – a description of the sources, processes, and provenance of a resource
7. DQ_DataQuality – using ISO 19157 to provide a description of the quality of the resource
8. MD_Distribution – a description of where and how to obtain the resource
9. MD_ContentInformation – a description of the content of the geospatial information types in the resource

10. MD_PortrayalCatalogueRefererence – a reference to the catalog of symbols/methods use to portray the resource
11. MD_ApplicationSchemaInformation – a description of the machine-readable schema, which characterizes the structure and content of the resource
12. MD_MaintenanceInformation – a description of the scope and frequency of updating and maintenance of the resource
13. MD_MetadataExtensionInformation – a description of any nonstandard user-defined metadata included in the metadata for a resource

The standard further specifies properties, additional classes and subclasses for each of these higher level classes.

The ISO 19115-3 provides a rule-based encoding for ISO 19115 Parts 1 and 2 in XML. It defines a collection of 18 XML schema namespaces roughly equivalent to the UML packages in ISO 19115-1 and 2. It also defines a set of XML schemas representing the aggregation of packages in ISO 19115-1 and 2 for various applications of metadata: metadata for data and services (a complete set of metadata for these two resource types); metadata application (for metadata resource collections with hierarchical metadata), metadata base for the basic mandatory elements, metadata for data transfer; metadata for data and services with geospatial common extensions; and metadata with extensions. ISO 19115-3 includes ISO/IEC 19757-3 Schematron rules that implement the validation constraints included ISO 19115-1 and 2 and also provides Extensible Stylesheet Language Transformations (XSLT) for transforming XML encodings of ISO 19115 using ISO 19139 XML schemas to ISO 19115-3 XML encodings and vice versa. The ISO Metadata XML schemas are located on the ISO standards portal http://standards.iso.org/iso/19115/-3.

The ISO 19115-4 will provide a JSON encoding for the for ISO 19115 Parts 1 and 2.

5.2 Semantic and organizational interoperability

Semantic interoperability refers to the ability of computer systems to exchange data with unambiguous, shared meaning. It involves transmitting the meaning of data along with the data itself by linking each data element to a controlled, shared vocabulary. This shared vocabulary enables machine interpretation, inference, and logic through this shared vocabulary, facilitating accurate and reliable data exchange between information systems. It ensures that the precise format and meaning of exchanged data is preserved and understood throughout exchanges between parties, so that "what is sent is what is understood".

In the geospatial domain, semantic interoperability poses an extra challenge due to the geospatial data heterogeneity, such as differences in data models, coordinate systems, and terminology, that diminishes seamless data sharing and reuse. The solution is applying technologies like the Semantic Web, Linked Data, and GeoSPARQL to publish, discover, and query geospatial data and services based on their semantic descriptions.

Semantic interoperability is a higher level of interoperability that builds upon syntactic interoperability, which focuses on the packaging and transmission mechanisms for data. It is critical for enabling seamless data sharing, integration, and analysis across disparate IT systems and organizations. Achieving semantic interoperability requires the use of standardized vocabularies, ontologies, and information models that provide a common language and framework for describing and interpreting data. This allows different systems to map their local terms and concepts to the shared semantic model, enabling meaningful exchange and understanding of the data.

In contrast with semantic interoperability, organizational interoperability is the highest level of interoperability, focusing on seamless sharing and interpretation of patient data, regulations, and requirements across multiple stakeholders, organizations, and individuals. Achieving organizational interoperability involves considering non-technical aspects like policy, legal, social, and organizational factors to ensure cooperative use of shared data within and across organizations. This level of interoperability enables the comprehensive understanding and meaningful exchange of information, promoting efficient data sharing and collaboration in healthcare settings.

We should distinguish fully semantic annotation of geospatial concept to share data beyond the geospatial domain and exchanging geospatial data between different GIS systems that are specialized for particular disciplines. In the first case you will need ontologies that describe the geospatial data models, coordinate reference systems etc. In the second case we can rely on the assumption that the geospatial domain is part of the universe of discourse and there is a common understanding of the geospatial formats and geometric models. In this circumstance, the semantic should focus on the definition of the properties attached to the geospatial objects or the geospatial domain. For example if the properties are observedProperties defined as numbers, semantic interoperability should focus on defining the meaning of this numbers and the units of measures in which the quantities are expressed. In the semantic web, these concepts are expressed as URIs that can be resolved and understood. There are several vocabularies that can be useful to describe observedProperties or variables such as:

https://community.wmo.int/en/whos-hydrological-ontology
https://www.fao.org/agrovoc/
https://vocab.seadatanet.org/search
http://cfconventions.org/cf-conventions/cf-conventions.html
http://vocab.nerc.ac.uk/standard_name
https://www.eionet.europa.eu/gemet/en/about/
https://www.eea.europa.eu/help/glossary
https://dd.eionet.europa.eu/vocabularyconcept
https://eur-lex.europa.eu/browse/eurovoc.html?locale=en
https://www.earthdata.nasa.gov/learn/find-data/idn/gcmd-keywords
https://ecoportal.lifewatch.eu/ontologies
https://vocabs.lter-europe.net/envthes/en/
https://qudt.org/2.1/vocab/quantitykind
These are resources that define units of measure:
http://vocab.nerc.ac.uk/collection/P06/current/UKMC/
https://qudt.org/2.1/vocab/unit
http://www.ontology-of-units-of-measure.org/resource/om-2

5.3 Data quality assurance and documentation

The value of geographic data is directly related to its quality. The views on quality differ among the various user communities; for example, cadastral applications require positional accuracy within a few centimeters, whereas a nautical chart requires a positional accuracy of only a few meters. Geographic datasets are being increasingly shared, interchanged, and used for purposes other than the producer's intended uses. The purpose of describing the quality of geographic data is to facilitate the selection of the geographic dataset best suited to application needs or requirements.

Geospatial data quality encompasses several key aspects that are essential for ensuring the reliability and usefulness of geospatial information. These aspects include:

1. Accuracy: Accuracy refers to how close a measurement is to its actual value. It is crucial for geospatial data to accurately represent real-world locations and attributes to support informed decision-making and analysis
2. Completeness: Data completeness measures the totality of features present in a dataset. Complete data sets contain all necessary attributes and information required to achieve the intended task or analysis
3. Consistency: Data consistency refers to the absence of conflicts or discrepancies within a dataset. Consistent data ensures that information is coherent and reliable for use in geospatial applications.

These aspects are combined with the three components of the geospatial information: Spatial, thematic and temporal.

Ensuring high geospatial data quality involves assessing and addressing these components to enhance the accuracy, precision, completeness, and consistency of spatial information, ultimately improving the reliability and effectiveness of geospatial analyses and decision-making processes. Indeed, the consequences of poor geospatial data quality can have significant impacts on various aspects of operations and decision-making including : Using data with errors can result in drawing incorrect conclusions, leading to misguided decisions based on flawed analyses, financial losses and lost of opportunities. For example, incorrect spatial information could lead emergency services to the wrong location, potentially delaying life-saving interventions or using inaccurate maps could result in damaging energy installations or water/gas pipelines, endangering individuals and disrupting essential services. Addressing and mitigating these consequences require a concerted effort to improve data accuracy, precision, completeness, and consistency in geospatial datasets.

The topic of quality is spread across three ISO 19100 standards, data quality (ISO 19157), its implementation (ISO/TS 19157-2), and quality assurance (ISO/TS 19158).

The ISO 19157 - Data quality defines the data quality elements that contain quantitative quality information. ISO 19157 defines six types of data quality:

1. Completeness. Presence and absence of features, their attributes and relationships. Negative example: missing road data in a remote part of the province.
2. Logical consistency. Degree of adherence to logical rules of data structure, attribution, and relationships. Example: the application schema distinguishes between public and private buildings. The dataset distinguishes between low buildings and high-rises.
3. Spatial accuracy. Accuracy of the position of features in relation to the Earth. Example: the absolute point accuracy is 10 cm (diagonal).
4. Temporal quality. Quality of the temporal attributes and temporal relationships of features. Example: the date of the data compilation was August 2010.
5. Thematic accuracy. Accuracy of quantitative attributes and the correctness of non-quantitative attributes, as well as the classification of features and their relationships. Example: areas have been classified according to remotely sensed imagery as meadows although they were swamps.
6. Usability. Degree of adherence to a specific set of data quality requirements. Example: a product specification fully applies for the intended purpose; however, the data are 12 years old.

Data quality may be reported using metadata or a stand-alone report. Data quality shall be reported as metadata in compliance with ISO 19157 and ISO 19115-1. To provide more details than reported as metadata, a standalone report may additionally be created. Its structure is not defined.

ISO/TS 19158 - Quality assurance of data supply defines a framework for the producer and customer in their production relationship. It bridges the gap between the quality management systems as defined in the ISO 9000 family of standards and the technically-oriented quality standards of the ISO 19100 family, such as ISO 19157 Data quality.

Through the application of ISO/TS 19158 there are opportunities for a better understanding of requirements by all involved in production and update, especially within multiple-producer environments. Additional benefits may be reduced data throughput time, reduced rework, improved data quality, and increased confidence within a relationship, leading to lower costs for both supplier and organization.

The existence of ISO/TS 19158 does not impose its application, thus releasing smaller companies from potentially unnecessary organizational overheads. The question of whether a demanded quality level has been reached is answered by second-party accreditation. This is performed by the customer regarding the quality of the supplied data without the involvement of a third-party accreditation body.

QualityML is a comprehensive framework and dictionary for encoding geospatial data quality metadata in a standardized and interoperable manner. It is a profile of the ISO geospatial metadata standards, such as ISO 19157, providing a structured set of rules and vocabularies for precisely documenting quality indicator parameters. QualityML defines a hierarchical structure of quality concepts, ranging from high-level quality classes to specific quality measurements and metrics. This allows for detailed and unambiguous representation of data quality. The framework leverages existing standards like UncertML for encoding statistical quality measures, while also extending them to cover a broader range of commonly used quality metrics. QualityML enables the encoding of quality metadata not just at the dataset level, but also at the pixel or feature level, allowing for more granular quality information to be associated with geospatial data. The benefits of QualityML include improved product discovery, better transmission of data characteristics, and the ability for other components (e.g., visualization, discovery, comparison tools) to be quality-aware and interoperable. QualityML has been developed and refined through various European research projects, ensuring its relevance and alignment with industry needs and best practices. QualityML has been implemented in the GeM+ metadata editor. Read more in www.qualityml.org

5.4 Introduction to web services standards and Web APIs

5.4.1 Elements of distributed computing

This section provides the necessary background about network communications and the web protocols. The concepts introduced in this section give us the necessary bases to describe the open geospatial web services in following sections.

5.4.1.1 Client-server architecture

A typical application usually includes three essential elements: presentation, logic and data. The graphical user interface is the responsible element for presentation, the processing code is the responsible element for logic and the database or the file management system is the responsible element for the data. The three elements are interconnected and, when needed, one element makes a request to the other that fulfills the request and responds the necessary information to the first.

The element making the request is called the client and the element fulfilling the request is called the server. The user typically interacts with the graphical user interface (the client) and the client software is the responsible to communicate with the logic and the data (the server). Which physical machine contains the three elements depends on the architecture of the application:

- In a typical stand alone GIS application, the three elements reside in the same computer and you need to be a programmer to actually realize that this separation exists.
- In a typical GIS application accessing the data over a local area network, the presentation and the logic are on one machine but the data is on another machine.
- In a web application, the presentation is on one machine while the logic and the data are on another.

The last alternative is the one selected by the web services and it will be the architecture mainly considered in this chapter.

We have seen that the physical role the three elements depends on the architecture but the logical roles can also depend on the architecture as well, so the logical role of client and server is not fixed. For example, the logic element is the server from the presentation point of view, but it is the client from the data access point of view that acts as a server.

Later we are going to see that the web client-server architecture is a special case of the general client-server architecture. In the web client-server architecture, the client and server communicate using a web protocol.

Figure 3. Client server architecture and the protocol used

5.4.1.2 Messaging

Client and server are two pieces of software that need to communicate talking a common agreed language and using a common set of rules (protocol) in a way that they can understand each other. The role of the protocol is to connect the client and the server in a way transparent to users and that hides the complexity of the system. There are many connection models such as the Remote Procedure Calls (RPC) and Publication-and-Subscribe (PUBSUB). In a RPC, a client sends a request to a remote server and waits for the server to respond. The server executes the procedure, generates the answer and returns the result back to the client that receives it and closes the connection. HTTP uses RPC. RPC does PULL requests because the client initiates the movement (as if the client is pulling a door to enter the communication). In PUBSUB (which does a PUSH request), the client subscribes to some events based on the interests of the user and, later, the server will initiate a communication to the client when there is a need for the client to be notified depending on the client subscription preferences. One typical application in the GIS world is a client that subscribes to changes in the features representing in some area of interest. When changes in the scope of the subscription are made, the server notifies the client about those changes. PUBSUB can be implemented on top of TCP/IP. HTTP cannot do PUBSUB directly but it can emulate it by forcing the client to regularly requesting the server if changes have happened in its subscription, using RPC.

5.4.1.3 Levels of protocol

Web services are developed over a stack of protocols organized in different levels, one on top of the other. To understand the need for a stack of protocols, an analogy can be used: The bank industry had set up a communication protocol with their clients based on sending them some standardized letters where different expenditures and incomes are detailed. Clients regularly receive these messages, review them and react accordingly. This protocol is part of a bank mail notifications protocol where commercial advertisements and other communications are performed by letters all having the bank brand logo very visible in the envelope. This protocol level uses yet another lower level protocol of the postal correspondence. In the postal protocol, letters need to be in an envelope, with a destination postal address and a stamp. Envelopes are deposited in a postal office and they will be classified and will travel a sequence of postal offices until they are close enough to be delivered to the final address. This communication protocol relies on a heterogeneous very lower level protocol: the communications network formed by cars, boats and planes that follow a set of rules through their own media to move around people, goods and, of course, mail (including the bank mail). The bank industry is using the whole stack of protocols without even realize about the complexity of the system. But dependencies exist, and if something goes wrong in a low level protocol (e.g., a lorry strike), the whole stack cannot function properly.

The web protocol is a high level protocol and, below that, other layers of protocol are defined one on top of the other. At the very bottom of the stack, the physical communication layer use a mixture of protocols orchestrating communications among the physical devices and cables of the network, such as mobile phones, Ethernet cables, fiber optic cables, wireless connections, etc, all making Internet communication possible. On top of the physical layer the transport protocol uses Transmission Control Protocol/Internet Protocol (TCP/IP) as a network protocol. It is based on fragments of messages called packets that are transmitted through a sequence of computers that route each packet from the source to the target computer. Each computer has an IP logical address that univocally identifies it and the IP address is used to route the packet from one computer to the next. On top of the TCP/IP, the actual web protocol is defined: Hypertext Transfer Protocol (HTTP). HTTP is an application protocol for hypermedia information systems. Hypertext is structured text that uses logical links (hyperlinks) between documents containing text. HTTP is the protocol designed to exchange or transfer hypertext.

Upper levels of protocol depend on the lower levels; sometimes the separation among levels of protocol is not strict. We are going to see implementation protocols that use (or misuse) some of the HTTP application protocol characteristics.

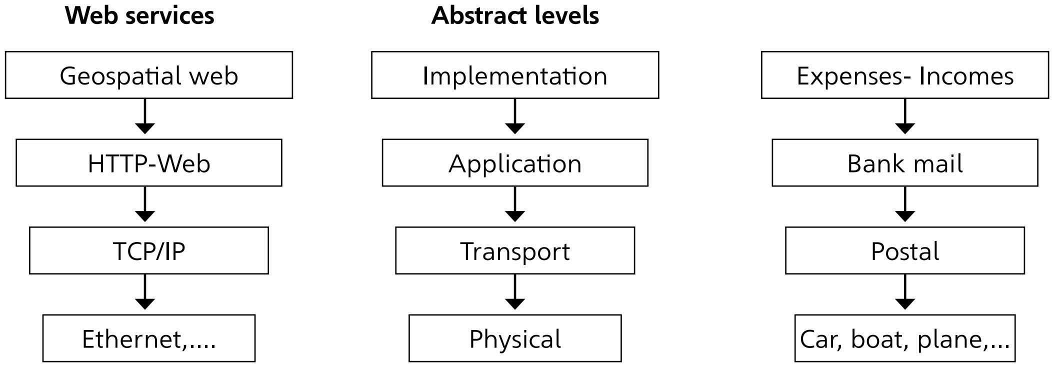


Figure 4. Levels of protocol

5.4.1.4 HTTP

HTTP is a request–response application protocol that is also based on the client–server architecture. The most used client application is the web browser, and a web server application running on a computer hosting a website is the most common server. The client submits an HTTP request message to the server and waits. The server returns a response message to the client, generally containing a single document, being the most common a HyperText Markup Language (HTML) file. The response contains a completion status information about the request in a section of the message that is called the headers and may also contain the requested content in its message body (e.g. the HTML file).

In HTTP, resources are identified and located on the network by Uniform Resource Locators (URLs), using the Uniform Resource Identifiers (URI's). URIs are the bases for requesting documents and are also the bases of the hyperlinks in HTML, acting as a way to link a fragment of an HTML page to another HTML file. A URL can represent pre-stored data or data that is generated dynamically; depending on the implementation in the server side. Commonly, resources correspond to pre-existing files but they can be the dynamic output of an executable application residing on the server side.

HTTP requests use different methods (also called verbs). All methods use a URI, but depending on the method, the client can require a desired action to be performed on the identified resource in the server side. The current specification of HTTP 1.1 defines several methods that have a clear semantics on the action to be done by the server on the URL requested:

- The GET method requests (retrieves) a representation of the specified resource. Requests using GET should only retrieve data and do not contain any payload. The expected response is a document representing the resource as a body of the response.
- The HEAD method asks for the same resource than a GET request, but the response will not contain the resource but only metadata about it in the headers of the response.
- The POST method sends a resource in the body of the message that is requested to be created and stored in the server. The URL of the POST method is NOT the URL of the resource but a new subordinate of the web resource identified by the URI. The resource is expected to be saved and the URL of the new resource returned in the headers of the response.
- The PUT method sends a resource in the body of the message that is requested to be stored under the supplied URI. Generally, PUT is used for updating existing resources, but if the URI does not point to an existing resource, then the server can create the resource with that URI.
- The DELETE method requests that the resource pointed by the URL is deleted.
- The TRACE method echoes the received request so that a client can see what (if any) changes or additions have been made by intermediate servers.
- The OPTIONS method requests information about the HTTP methods that the server supports for a specified resource URL. It can be used to check if a verb is supported before actually requesting it.
- The PATCH method applies partial modifications to a resource.

All general-purpose HTTP servers are required to implement at least GET and HEAD methods, and all other methods are considered optional by the specification.

HTTP is a stateless protocol. In a stateless communication protocol no information is retained by either sender or receiver. In practice this means that requests are independent of each other and the client should not assume that any acknowledge of any previous communication was retained by the server. The stateless design simplifies the server design because there is no need to save and store any details of a request or a response while dealing with conversations between client and server. In practice an application needs to maintain the state of the dialog with the user, so in a stateless communication, details about the state are send to the server in every request; the server need to interpret and many times returned back to the client status information (before closing the a individual request-response circle and loosing the information) to allow for the state of the application to be maintained.

The HTTP protocol was designed to transport information in plain text mode. Due to the nature of the TCP/IP, the computers in the local area network of the client and the server and all intermediate computers that route the communication have access to plain text and can easily read it or even store it. Confidential information transmitted in this mode can fall in the hands of unwanted people, including password, credit cards numbers, etc. To prevent that, a secured version of HTTP was introduced: HTTPS. In HTTPS the URL and the message are encrypted and cannot be read by any computer but the receiver.

5.4.1.5 The web server

A web server is a software (and a hardware that executes it) that implements the server side HTTP application protocol by processing incoming network requests and responding documents. The primary function of a web server is to store, process and deliver web pages to clients. Most common web pages are HTML documents and all related content which may include images, style sheets, scripts (e.g. JavaScript), etc. Common examples of web server software are Internet Information Server (MicroSoft), Apache Server (Apache), NGINX (NGINX Inc.), etc [2].

A web client, commonly a web browser, initiates a communication by making a request for a specific resource URL and the server responds with the requested document or an error message (e.g. if the URL resource does not correspond to any content). The resource might be a real file on the server storage system, but it could also be generated on the fly based on the content of the server database. This is particularly useful for serving geospatial information that it is usually stored in databases that can be dynamically converted into documents (e.g. a GML or a GeoJSON document) on the fly.

The primary function is to serve content by using the GET operation, but complete implementation of HTTP also allows the server to receive content from clients (PUT and POST operations). This feature is commonly used for submitting web forms (e.g. when users answer surveys) or to upload files (e.g. when users store or exchange files on the cloud) and could be utilized to accept and store user edits on a map.

Web servers also support that a URL points to an executable. This executable will be conformal to the CGI, ISAPI or a similar alternative. When the URL to an executable is received, it is executed and it will use the information in the request to perform a task (e.g. generate a geospatial resource). In the geospatial domain, the executable could render a map (from geospatial information in a database) in an image and return it as the response. Modern web servers store documents that contain server side code (such as Active Server Pages (ASP), PHP, etc.) that will be dynamically interpreted by the server when the URL of the server script is requested. Executables and server side scripts are means to augment the behavior of the web server by adding separated elements, while the actual web server main software remains unchanged. Later, we will see that geospatial services will uses this characteristic to implement geospatial services on top of general use web services.

5.4.1.6 Web client

A web client is a piece of software that implements the client side HTTP application protocol by generating network requests and handling and displaying responded documents. The primary function of a web client is to visualize web pages to users. The most common web client is the web browser, which is a generic application for accessing and visualizing many kinds of information available in the World Wide Web (WWW) in the user device. It also has to react to human actions; being the most common one a click on a hyperlink. The web browser is able to visualize HTML pages, images, video, etc. Modern web browsers are able to handle several kinds of information, apply styles to the HTML pages using style-sheets and execute client side scripts (e.g. JavaScript). Common examples of web client software are Internet Explorer, Microsoft Edge, Google Chrome, Mozilla Firefox, Safari, Opera, etc.

Images embedded in HTML files were the first map applications shown in web browsers. Modern web browsers are complicated applications supporting several formats, but they can accept the inclusion of external supplementary code (e.g. with ActiveX). This method is used, for example, to add JPEG2000 imagery viewers.

In recent years, web browsers have evolved fast but their performance cannot be compared to the desktop solutions. In addition, they have security restrictions that severely limit access to local resources (in particular to the local storage drive). To overcome those limitations, application programmers can implement or incorporate web clients in desktop solution specifically tailored for the GIS data requirements that will be capable of combining local files on top of remote data.

5.4.1.7 The URL structure

The URL is a fundamental part of the web request and all HTTP methods use it. In the HTTP protocol, the URL has several functionalities embedded in a single query string: The URL is the way that the client has to identify the server. In addition, a URL points to a resource in the web server space; this resource can be a document or a dynamic server side script or application. In the case that the resource is a dynamic server side script, the URL can include a query string to the server. A URL can also contain a final fragment that the client will process after the server responds. The full URL syntax structure is a sequence of six components disposed like this:

<protocol>://<host>:<port>/<resource>?<query>#<fragment>

For example: http://hostname.com:80/pdfs/mypdf.asp?info=houses#page=3

- *protocol* (in the example "http"): it is the protocol name. In HTTP it can be http or https.
- host (in the example "hostname.com"): it is the name (or the IP address) of the server that contains the resources we are interested in. Commonly it is composed by works separated by dots (e.g. www.maps.com).
- *port* (in the example "80"): It is the TCP/IP port number. If it is missing and http protocol is indicated, 80 is assumed. If https protocol is indicated, 443 is assumed. Some map services uses other ports by default to avoid conflicts with other preexisting services in the server. The Windows distribution of GeoServer setup offers the port 8080 by default. If the default is accepted, all communications to the service need to contain the ":8080" in the URL to work.
- *resource* (in the example "pdfs/mypdf.asp"): It is the path and name of the resource (e.g. a HTML file to retrieve or a server side application or a server side script to execute).
- *query* (in the example "info=houses"): It is the query information that is delivered to a server side application or script. This query can carry parameters to inform the application or the script what to do. Key and Value Pair (KVP) is a particular form of query that represents a sequence of blocks consisting in a key name, an "=" and a value. Each block is separated by an "&". KVP is particularly useful if the resource is a map server that will deliver dynamic maps; this query can transport the extent, the coordinate reference system, the size of the map, the types of features it will contain, the style to portrayal them, etc.
- *fragment* (in the example "page=3"): It is information that will be returned to the client in the response. The client side can use it to do an action. The most common use of "fragment" is to contain a name of an anchor inside a HTML page. Once the browser receives the HTML page, it will scroll down to ensure that the content after the anchor is visible. For example, Adobe PDF uses it to indicate a page number where the book should be presented when it is received and opened in the screen.

5.4.1.8 The HTTP message structure

In the interaction between client and server, a message is send. Once the server responds, another message is send back. These messages are plain text. The general structure of the two messages is the same. Basically, they have two parts: the header and the body. They begin with a header, which contains a sequence of text lines composed by a name, a ":" and a value. A blank line will mark the end of the header and the body of the message will follow. The headers are generally not shown to the user and contain metadata about the communication or the body. They might contain the media type of the body, the length of the body, the calling URL, cookies, security information, language negotiation information, etc. In a request message, headers will also contain the method (e.g. GET, POST) and, in the response, headers will contain the HTTP status number. The HTTP status number is an important part of the response that will indicate if the request has been successfully executed (e.g. by returning 200) or if something went wrong. (e.g. 404 means "file not found" and 500 means "error in a server side application"). The body of the message contains the actual payload. In a GET request, the message has an empty body and the response should contain the document requested in the URL or the result of the execution of an application pointed by the URL. In a POST or a PUT request the body will contain information for the server to save. Even if the HTTP message is plain text, there are ways to transmit a binary payload by re-encoding it in text using one of the common methods available (initially designed for email attachments, e.g. base64). For example, a map server need will use base64 to encode the body of the response containing a map renderization in a JPEG file format that the client has requested with the intention of presenting it in the screen.

5.4.1.9 First attempts to visualize maps in HTML

HTML was initially defined to support hypermedia. To make the presentation more visually appealing, GIF, JPEG or PNG images were included. Soon, it became apparent that a server could be configured to automatically create maps by generating GIF or JPEG. In June 1993, XEROX Corporation generated what is considered as the first interactive map viewer. Based on the selection of the user in a web page, the client software requested maps to a service: an application that was able to retrieve geospatial data from a database and render it at screen resolution and convert it into a GIF file that was returned to the client to be included as an illustration in a HTML page. Many other map browser solutions started to appear, such as the open source MapServer, and many proprietary web pages started to operate, such as MapQuest, all using the same principle. However, none of them could communicate with any other.

In 1997, the Open GIS Consortium (later renamed as Open Geospatial Consortium; OGC) proposed the OGC WWW Mapping Framework with the mission to investigate how to make similar mapping approaches implemented by different vendors interoperable. In 2000 the OGC proposed the version 1.0.0 of the Web Map Service (WMS). WMS was been a huge success and today there are thousands of standard web map services in the world produced by governments, academia and commercial sector and hundreds of web map browsers [4].

The OGC standardize other services that exchange other kind of geospatial information and adopted XML and XML schema as a lingua franca to encode many media types for example service metadata, feature based data (GML), observations based data (O&M), etc., None of the other OGC services has become so popular and has had the same impact as WMS, may be due to fundamental limitations in the early versions of HTML.

5.4.1.10 JavaScript

JavaScript stated as a Netscape's cross-platform, object-based scripting language for web applications. Currently, the most popular web browsers support it and some provide mature debugging tools. Web browsers interpret JavaScript statements embedded in an HTML page and execute the on the fly. The browser reads the HTML page from top to bottom, displaying the results of the HTML and executing JavaScript statements as they are encountered. JavaScript statements embedded in an HTML page can also respond to user events such as mouse clicks, form inputs, and page navigation. For example, a JavaScript function can verify that a user has entered valid information into a form requesting a latitude and longitude. Without any network transmission, the embedded JavaScript on the HTML page can check the entered data and display a dialog box if the user enters invalid data. Modern JavaScript libraries can contain thousands of lines of code and can be used together in a single application. JavaScript can request more data to the service while the HTML page continue responding to the user in a strategy called Asynchronous JavaScript and XML (AJAX). For example, a JavaScript application can request a WFS service more feature data and interpret the response in the form of a GML file, while the user can continue interacting with a map browser.

Web based map clients are coded in JavaScript such as OpenLayers and Leaflet and provide JavaScript APIs with a vast list of functions for developers for presenting and manipulating geospatial data on HTML pages easily.

Recently a small subset of the JavaScript language called JavaScript Object Notation (JSON) has emerged as a simpler and easier to use alternative to XML. JSON is not a programming language but a way to encode objects in JavaScript that has been made independent from it, and can be used elsewhere. JSON does not include the concept of classes and data types but objects can be validated under a certain rules defined in a JSON Schema that acts in a similar way that XML schema does for XML. Another recent advantage of JSON objects is that they can use JSON-LD to be translated to RDF automatically applying JSON context documents [5].

5.4.1.11 Improvements in HTML5

HTML5 is the last effort to improve HTML in a way that can be a good candidate for developing applications on mobile platforms. Listing all the improvements of HTML is out of scope of this chapter but there are two new characteristics that are worth mentioning:

HTML5 introduces a new graphics programming interface called canvas. The <canvas> element is a region of the HTML where JavaScript programmers have full control of the graphical content. It is possible to extract a raster representation of the canvas and change pixel colors one by one. It is also incorporates functions to draw, icons, lines and polygons. We foresee that canvas will enable a completely new generation of geospatial applications, able to present vector data directly and to dynamically change visualization styles for both raster and vector data [6].

5.4.2 Geospatial web services

Geospatial data is constantly being created by cartographic agencies, automatic in-situ sensors, remote sensing satellites, etc. In many countries, it is in the aim of a transparent administration to share, and exchange data with the citizens that can be informed and formulate opinion. In addition, governments have recognized that by making data public, private companies can discover new usages that go beyond the ones initially foreseen by their producers and generate new business models and profits by selling new products and services. In this chapter we are going to review the current geospatial web services that facilitate the dissemination of geospatial data over the web [7].

Most of these geospatial services have been defined by the OGC community, based on a consensus process that favors interoperable geo-processing and provides services to communicate to remote geodatabases. In principle, OGC standards are defined for multiple distributed computing platforms while maintaining common geospatial semantics across the underlying technology. In practice, most of the OGC geospatial services are using the Internet infrastructure and the HTTP protocol in the SOA (and sometimes ROA) architectural style described in a previous section.

5.4.2.1 Services types

You might expect that a single big web service could deal with all geospatial aspects by incorporating all the necessary operations, in the same way that generic a desktop GIS in your computer is able to perform most of the necessary operations. There have been some successful attempts to do this in the commercial sector (such as the ESRI ArcGIS REST API) however, this is not the common situation. Standard geospatial web services are dealing with a single aspect and a single data model. If we want to make sense of this variety of alternative services we need to classify them. The OGC reference model proposes a way of classifying geospatial web services by purpose and the later by data model used [8]. This chapter follows the same approach:

- Visualization services have the purpose of presenting the data in an easy to understandable way. Clients are able to request formats that can be easily placed in a computer screen by client software, such as JPEG, PNG or KML. Servers will be able to respond with visualization formats on demand. To be able to do this, visualization services need to combine the geospatial data with symbolization rules to create a presentation for the geospatial data.
- Data exchange services have the purpose of giving immediate access to data and in some cases permitting the upload of new geospatial data on the server. Clients should be able to request geospatial data exchange formats such as GML, GeoJSON, GeoTIFF or NetCDF. Servers need to respond with the subset of data (as requested) and clients will be able to import the data in their GIS data storage systems. In some cases clients should also be able to send geospatial features to the server that should store them for further use.
- Data scheduling and notification services have the purpose of preparing the future acquisition and notify went it becomes available in data exchange services. The client should be able to specify the use needs in terms of extend, time, data type, etc. The server will schedule the acquisition and will issue a notification when new data becomes available.
- Discovery services have the purpose of storing metadata about the data and services available in the web. Clients will formulate a request describing the type of data they need in terms of extent, projection, resolution, data quality, theme, etc. Servers should be able to analyze the query and respond with a set of metadata records describing the data or services that match the needs of the client.
- Processing services offer geospatial processing routines and hardware. Clients should be able to formulate a request that includes the processing operation, the data to be processed and the parameters to be used to tune the process execution. Servers should be able to perform the requested task and respond with the result of the operation that will commonly be in the form of a geospatial dataset.

5.4.2.2 Common architecture for geospatial web services

As exposed before, the geospatial web services are separated in different types that follow standards developed by the consensus of different groups. Fortunately, there are mainly two common functionalities that all services implement: a common mechanism for formulating requests and a service self description in the form of service metadata (the GetCapabilities operation). These and other minor common characteristics are defined in the OGC Web Service Common standard [9].

In SOA style, OWS Common also defines one or more common encoding to formulate requests to the service based operations. These operations are defined in the standard and are enumerated in the service metadata document. The first encoding is the KVP on HTTP GET URLs. In the query part of the URL a request identifier, a service name and a version number can be specified. Depending of the request name, other keys defined in the corresponding standard can be expected. Sometimes, operation parameter values are too complex to be included in a URL (e.g. when a XML file needs to be send to the server) and XML-POST can be an alternative. In these cases HTTP POST is used to send a XML document containing all parameters needed as instructions to execute the operation in the server. OGC services also support another encode that looks similar to the XML-POST but uses an IT mainstream approach: SOAP. In SOAP requests are embedded in XML files that are called envelops. SOAP services publish another self discovery document called WSDL. Modern OGC web services are defined in an abstract level that can be implemented using one or more of the encoding alternatives described: KVP, XML-POST and SOAP.

A self describing service is a service providing enough information for a client to discover the resources available and the capacities of the service without user intervention. In OWS Common this is done by specifying that all services will provide a service metadata document. The OGC uses a simplified version of the ISO 19115-1 metadata for services. A service metadata document should contain the following information:

- Identification: Contains metadata to identify this server in terms of standards and profiles followed, keywords describing the general content and fees and access constraints.
- Provider: Contains information about the organization operating this service.

In addition, if the service uses the SOA style, it will include:

- Operations: Enumerate the operations that this service supports. One of the operations is for retrieving the service metadata document (GetCapabilities).
- Resources: Enumerates resource identifiers this service offers.

All SOA style services implement an operation to retrieve the service metadata. This operation is always available as a KVP URL that has the query part composed by mainly 3 important keys: Service (containing the name of the service; e.g. WCS), request (that will contain the fixed text GetCapabilities) and AcceptVersions (indicate a list of version that the client understands). Some old service standards use version with a single version number instead of AcceptVersions. An example of GetCapabilities request will look like:

https://acdisc.gesdisc.eosdis.nasa.gov/daac-bin/wcsL3?service=WCS&version=1.0.0&request=GetCapabilities

By default the response of a GetCapabilities is in XML format. In the example, the first sub-element wcs:Service contains the information about services identification and provider. The second subelement *wcs:Capability* enumerates the available operations in the service showing the GetCapabilies and Describe Coverage operations (the list has been shortened for readability). The third sub-element *wcs:ContentMetadata* enumerates the available coverage names with a minimum set of additional metadata for each one.


```xml
<?xml version="1.0" encoding="UTF-8"?>
<wcs:WCS_Capabilities version="1.0.0" updateSequence="2012052200"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:gml="http://www.opengis.net/gml"
xmlns:wcs="http://www.opengis.net/wcs"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.opengis.net/wcs http://schemas.opengis.net/wcs/1.0.0/wcsCapabilities.
xsd">
    <wcs:Service>
        <wcs:description>This Web Coverage Service (WCS) is one of the multiple GES DISC data service
instances used to provide gridded Level 3 daily OMI data products.</wcs:description>
```

```
            <wcs:name>Ozone Monitoring Instrument (OMI) data from NASA Goddard Earth Sciences Data
and Information Services Center (GES DISC)</wcs:name>
            <wcs:label>Ozone Monitoring Instrument (OMI) data from NASA Goddard Earth Sciences Data and
Information Services Center (GES DISC)</wcs:label>
            <wcs:keywords>
                <wcs:keyword>Aura OMI WCS NASA GSFC DISC</wcs:keyword>
            </wcs:keywords>
            <wcs:responsibleParty>
                <wcs:individualName>Long Pham</wcs:individualName>
                <wcs:organisationName>DAAC/GSFC/NASA</wcs:organisationName>
                <wcs:contactInfo>
                    <wcs:phone>
                        <wcs:voice>1 301 614 5132</wcs:voice>
                    </wcs:phone>
                    <wcs:address>
                        <wcs:deliveryPoint>NASA/GSFC Code 610.2</wcs:deliveryPoint>
                        <wcs:city>Greenbelt</wcs:city>
                        <wcs:administrativeArea>Maryland</wcs:administrativeArea>
                        <wcs:postalCode>20706</wcs:postalCode>
                        <wcs:country>USA</wcs:country>
                        <wcs:electronicMailAddress>long.b.pham@nasa.gov
                            </wcs:electronicMailAddress>
                    </wcs:address>
                </wcs:contactInfo>
            </wcs:responsibleParty>
            <wcs:fees>NONE</wcs:fees>
            <wcs:accessConstraints>NONE</wcs:accessConstraints>
        </wcs:Service>
        <wcs:Capability>
            <wcs:Request>
                <wcs:GetCapabilities>
                    <wcs:DCPType>
                        <wcs:HTTP>
                            <wcs:Get>
                                <wcs:OnlineResource
xlink:href=”http://acdisc.sci.gsfc.nasa.gov/daac-bin/wcsL3?”/>
                            </wcs:Get>
                        </wcs:HTTP>
                    </wcs:DCPType>
                </wcs:GetCapabilities>
                <wcs:DescribeCoverage>
                    <wcs:DCPType>
                        <wcs:HTTP>
                            <wcs:Get>
                                <wcs:OnlineResource
xlink:href=”http://acdisc.sci.gsfc.nasa.gov/daac-bin/wcsL3?”/>
                            </wcs:Get>
                        </wcs:HTTP>
                    </wcs:DCPType>
                </wcs:DescribeCoverage>
            </wcs:Request>
            <wcs:Exception>
                <wcs:Format>text/xml</wcs:Format>
            </wcs:Exception>
        </wcs:Capability>
        <wcs:ContentMetadata version=”1.0.0”>
            <wcs:CoverageOfferingBrief>
                <wcs:description>
Spectral Absorbing Aerosol Optical Thickness for best fit aerosol model derived with the Multi-Wavelength
method, scaled by a factor 1000</wcs:description>
                <wcs:name>OMAEROe:AbsorbingAerosolOpticalThicknessMW</wcs:name>
                <wcs:label>Spectral Absorbing Aerosol Optical Thickness for best fit aerosol model derived
with the Multi-Wavelength method, scaled by a factor 1000</wcs:label>
```

```
            <wcs:lonLatEnvelope srsName="urn:ogc:def:crs:OGC:1.3:CRS84">
                <gml:pos dimension="2">-179.875000 -89.875000</gml:pos>
                <gml:pos dimension="2">179.875000 89.875000</gml:pos>
            </wcs:lonLatEnvelope>
        </wcs:CoverageOfferingBrief>
        <wcs:CoverageOfferingBrief>
            <wcs:description>Temporally averaged over the requested time for this product:
Spectral Absorbing Aerosol Optical Thickness for best fit aerosol model derived with the Multi-Wavelength
method, scaled by a factor 1000</wcs:description>
            <wcs:name>OMAEROe:AbsorbingAerosolOpticalThicknessMW_timeAveraged
                </wcs:name>
            <wcs:label>Temporally averaged over the requested time of the Spectral Absorbing Aerosol
Optical Thickness for best fit aerosol model derived with the Multi-Wavelength method, scaled by a factor
1000</wcs:label>
            <wcs:lonLatEnvelope srsName="urn:ogc:def:crs:OGC:1.3:CRS84">
                <gml:pos dimension="2">-179.875000 -89.875000</gml:pos>
                <gml:pos dimension="2">179.875000 89.875000</gml:pos>
            </wcs:lonLatEnvelope>
        </wcs:CoverageOfferingBrief>
    </wcs:ContentMetadata>
</wcs:WCS_Capabilities>
```

Figure 5. Service metadata document for a WCS

ROA style services reuse the HTTP operations and they do not need to define operation on top of HTTP, so that there is no specific section in the service metadata document. They still need to enumerate the list of resources available. This can be done in the service metadata document (that is the case of WMTS 1.0) or might provide a specific resource URL to do it (that is the case of WFS 3.0). Each resource type can be associated to one of more URL templates and the corresponding HTTP operations supported. A URL template is a way to define a pattern for a family of URLs. A URL template looks like a URL but it contains variable names between "{" and "}" that need to be substituted by values to build a valid URL [10].

An example of a URL template used in WMTS to define all tile URLs of layer is:

http://www.maps.bob/etopo2/default/{TileMatrixSet}/{TileMatrix}/{TileRow}/{TileCol}.png

If we substitute {TileMatrixSet} by WholeWorld_CRS_84, {TileMatrix} by 10m, {TileRow} by 1 and {TileCol} by 3, we will get the following valid URL for a particular tile:

http://www.maps.bob/etopo2/default/WholeWorld_CRS_84/10m/1/3.png

Recently, in ROA style a new format to describe a web service API has emerged: OpenAPI (formerly known as Swagger). OpenAPI is a format that describes the resource types in web service by providing URL templates, the definition of all variables in a URL template, and the format and schema of the expected resource representation. WFS 3.0 is the first OGC standard to incorporate this metadata format and it is foreseen that others can follow. OpenAPI documents are written in YAML and will be exposed in a later section.

5.4.2.3 Visualization services

In this section, we describe two OGC services for geospatial data visualization: The Web Map Service (WMS), also published as ISO 19128, and the Web Map Tile Service (WMTS). Both services are mainly intended for map preparation and retrieval. A map is a pictorial representation of geospatial data in an easy to display format (e.g GIF or JPEG) [11]. It is a combination of geospatial data and a set of symbolization rules. Given an area of a device screen where a remote map needs to be shown, a standard client wanting to use a visualization service has two possible alternatives: to formulate a personalized single WMS GetMap request to cover the exact area of the screen with the exact required extent of the map [12], or perform a series of WMTS GetTile requests of small tiles that, once displayed one next to the other, will cover the required area of the screen hiding the exceeded area [13]. After images are positioned in the screen, the user will not see any difference between WMS and WMTS approaches, but when the user tries to navigate quickly through the data, WMTS should outperform WMS due to its internal design. Tiles are optimized for faster delivery, since they can be pre-computed in the server. Furthermore, the fixed set of tiles enables the use of HTTP network mechanisms for scalability such as distributed cache systems [14].

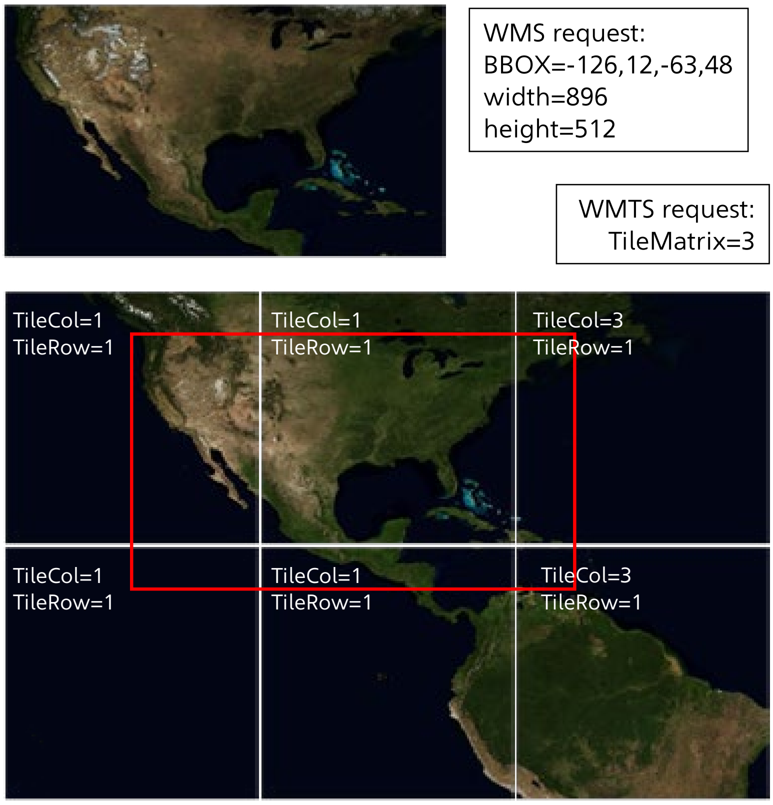


Figure 6. Requesting the same area with a single WMS request (https://neowms.sci.gsfc.nasa.gov/wms/wms?request=GetMap&service=WMS&version=1.3.0&CRS=CRS:84&format=image/jpeg&layers=BlueMarbleNG&style=) or with multiple WMTS requests (https://gibs.earthdata.nasa.gov/wmts/epsg4326/best/wmts.cgi?request=GetTile&service=WMTS&version=1.0.0&TileMatrixSet=500m&format=image/jpeg&layer=BlueMarble_NextGeneration)

In a WMS, a Coordinate Reference System (CRS) for the data is selected for the two dimensions of the map that are going to be represented defining an implicit linear transformation between the desired bounding box in CRS coordinates (called BBOX) and the two dimensions of the image that starts at 0,0 of the top left corner and ends on the WIDTH,HEIGHT of the lower right corner. WMS can create maps at any desired scale (defined here as the size in the screen divided by the size in the reality). In WMS and WMTS the scale is calculated by considering that the map image will be shown at a standard 0.28mm pixel size display. Frequently, the true pixel size is unknown and 0.28 mm was the actual size of a common display back in 2005. This value is still being used as reference even if current display devices are built with much smaller pixel sizes.

In WMTS, the width and height is fixed and the available scales and bounding boxes are discrete and limited to regular matrix that is called the TileMatrix. Given an available scale, the size of a pixel can be calculated as pixelSize=scaleDenominator × 0.28 10^{-3} / metersPerUnit(crs). Tile sizes are tileWidth × pixelSize, tileHeight × pixelSpan. The tiled space starts at a predefined top left corner in CRS coordinates and it is filled with tiles one next to the other until MatrixWidth and MatrixHeigth is reached [15].

In WMTS only a list of TileMatrix (each one with its own scale) is allowed. This generally short list of TileMatrix's is known as a TileMatrixSet. A TileMatrix in TileMatrixSet has a unique alphanumeric identifier in the TileMatrixSet. Some tile based implementations prefer to use a zoom level number, which has the advantage of suggesting some order in the TileMatrixSet. In some other standards, this way of dividing

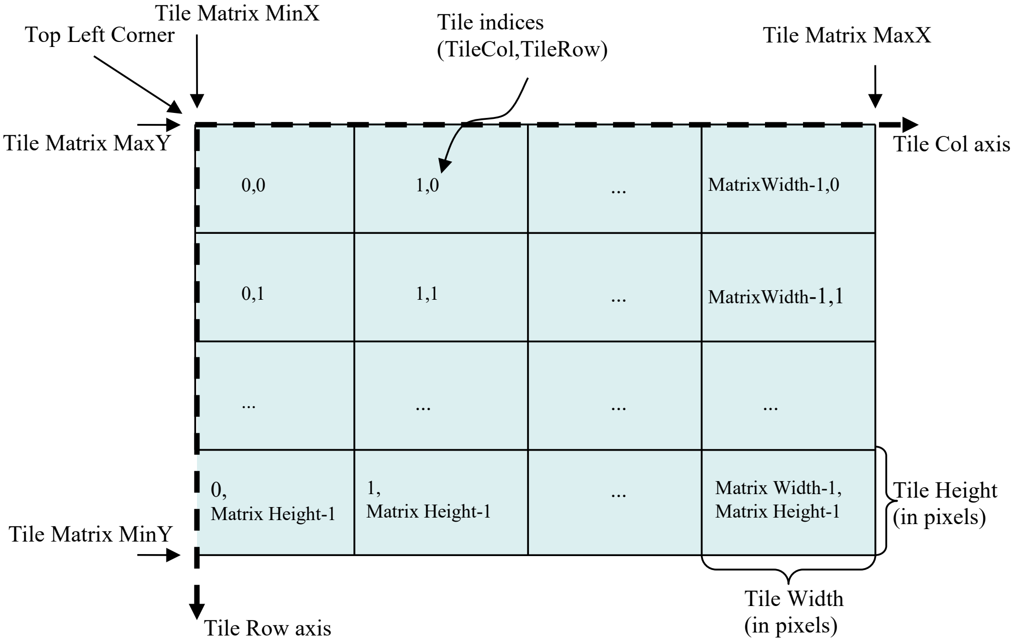


Figure 7. Tiled space

the space is called image pyramid like in clause 11.6 of the OGC KML 2.2 (OGC 07-147r2). JPEG2000 (ISO/IEC 15444-1) and JPIP (ISO/IEC 15444-9) also use a similar division of the space called resolution levels. Nevertheless, in those cases the pyramid is self defined starting from the more detailed TileMatrix (using square tiles), and tiles of the next scales are build by successively aggregating 4 tiles of the previous scale and then lowering the number pixels by a factor of 2. That approach involves a more rigid structure which has scale numvers related by powers of two and tiles that perfectly overlap 4 tiles on the immediately inferior scale denominator. TileMatrixSets are more flexible but encompasses the more restricted KML superoverlays or JPEG2000 resolutions levels.

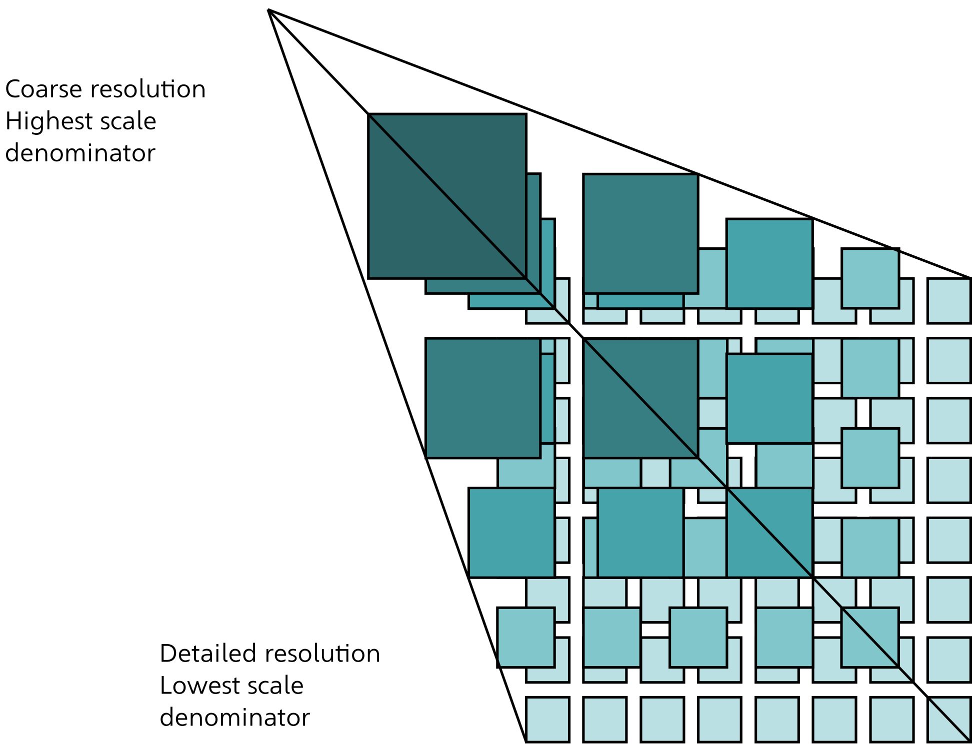


Figure 8. Tile Matrix set representation

Both, WMS and WMTS include three operations, one to get the service metadata (GetCapabilities) one to get a map or a tiled map (GetMap in WMS and GetTile in WMTS) and one to get the actual values represented in a given pixel of a map (GetFeatureInfo). In WMS or WMTS, geospatial information is classified in layers and GetCapabilities enumerates the layer names that can be requested, the possible style names applicable to them and CRS or TileMatrixSets available. With this information, it will be possible to formulate a map or a map tile request using KVP syntax.

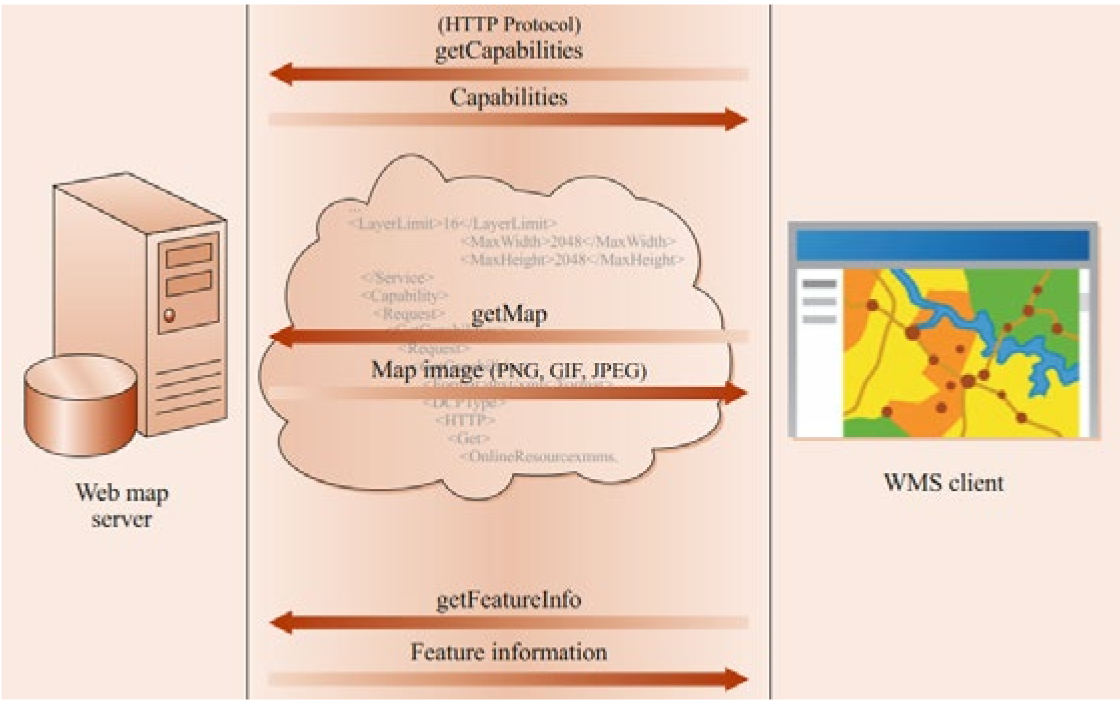


Figure 9. Web map service operations

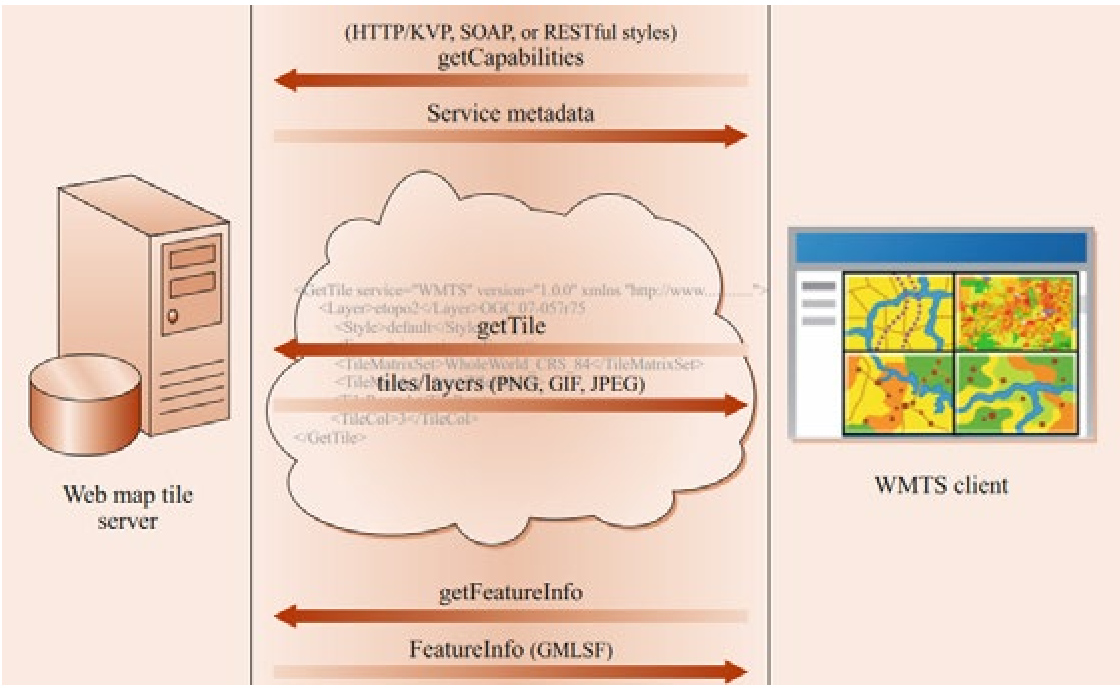


Figure 10. Web map tile service operations

Table 1: GetMap and GetTile requestparametres

WMS		**WMTS**	
Parameter	**Meaning**	**Parameter**	**Meaning**
Service	"WMS"	**Service**	"WMTS"
Request	"GetTile"	**Request**	"GetTile"
Version	"1.3.0"	**Version**	"1.0.0"
Layers	list of layer names	**Layer**	Layer name
Styles	A style for each layer name	**Style**	Style name
Format	Output format	**Format**	Output format
CRS	CRS identifier	**TileMatrixSet**	TileMatrixSet identifier
BBox	Bounding Box 4 coordinates	**TileMatrix**	TileMatrix identifier
		TileRow	Row index of the tile
		TileCol	Collumn index of the tile
Width	Map width		
Height	Map height		
Transparent	"TRUE" or "FALSE"		
Other dimension name	Other dimension value	**Other dimension name**	Other dimension value

A GetFeatureInfo operation has the same parameters as the operation to request maps but adds or overwrites the following ones:

Table 2 GetFeatureInfo requestparametres

WMS		**WMTS**	
Parameter	Meaning	**Parameter**	Meaning
Service	"WMS"	**Service**	"WMTS"
Request	"GetFeatureInfo"	**Request**	"GetFeatureInfo"
Version	"1.3.0"	**Version**	"1.0.0"
InfoFormat	Output format	**InfoFormat**	Output format
I	Row index of a pixel within the map	**I**	Row index of a pixel within the tile
J	Collumn index of a pixel within the map	**J**	Collumn index of a pixel within the tile

It is important to note that clients should be able to superimposed map views of information that come from multiple remote and heterogeneous sources into a single view. To help in this process, servers should be able to deliver transparent pixels when there is no data shown for this layer.

WMTS also includes a ROA style to retrieve tiles (and to formulate the equivalent of a GetFeatureInfo query) based on URL templates instead of SOA KVP style. The ServiceMetadata document provides the URL template for each layer.

A WMS can be extended to allow user-defined symbolization of feature or coverage data using Styled Layer Descriptor (SLD). This standard adds some extra operations to a WMS to define how the Symbology Encoding (SE) standard documents can be used with WMS [16]. Before being able to write a SE document, it is necessary to know more about the layers inside the WMS. With the DescribeLayer operation a user can request information about the features being symbolized (e.g. their feature/coverage type). With this information, the user can define a personalized symbology for the elements present in the layer that can be applied when the GetMap operation is requested. Symbology Encoding (SE) specifies a XML format of a map-styling language of the user-defined styling. A GetLegendGraphic operation produces a legend representation of the user-defined symbology.

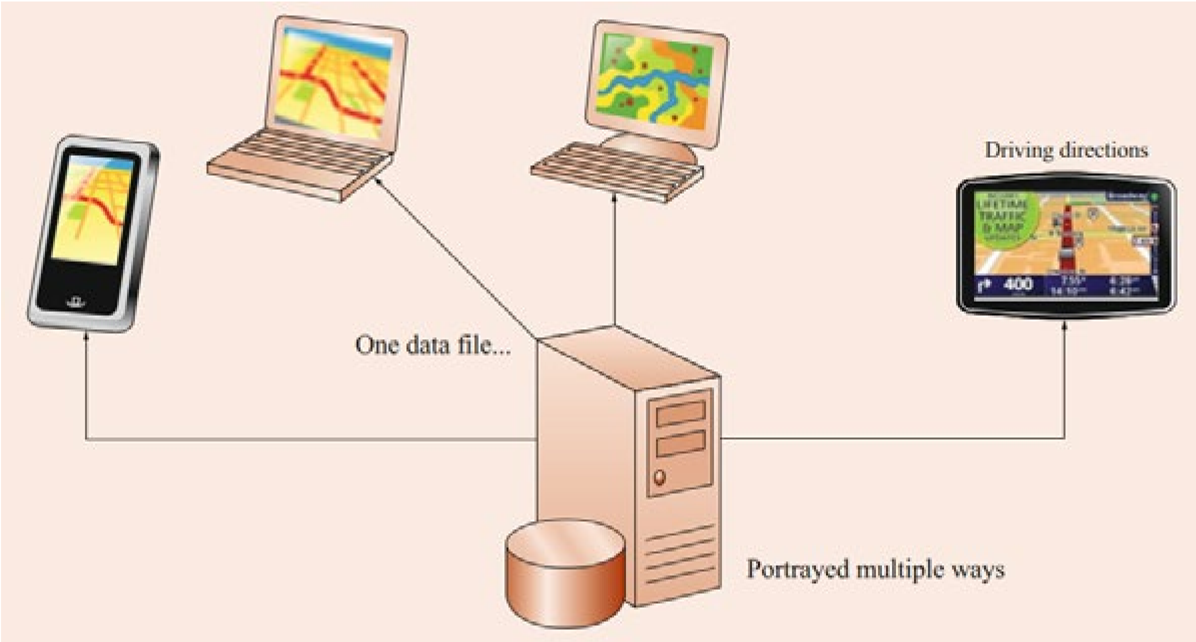


Figure 11. Style Layer descriptor

5.4.2.4 Data Exchange services

In this section we will describe three services for accessing, retrieving and, in some cases, uploading geospatial information in a distributed environment that targets three different data models. The Web Feature Service (WFS) exchanges geospatial information in the vector data models (called "features"), the Web Coverage Service (WCS) retrieves geospatial information in the raster model (called "coverages") and the Sensor Observation Service (SOS) retrieves geospatial information in the sensor models (called "Observations and Measurements", O&M). Since the three services follow the OWS Common model, in a first approach, we can describe them together. The three services provide an operation to retrieve the service metadata. This document enumerates the names of the resources available in the service. With these names, we can get more information about them with a "describe" operation. Finally we can retrieve the actual resources in their representation: WFS delivers the features in GML, WCS delivers the coverages mainly in raster formats (e.g. GeoTIFF, NetCDF...) and SOS delivers the observations in the O&M XML encoding. Sometimes, there is the need to limit a resource collection to a fragment of it by using some filtering capability in the retrieve operation or by using other similar operations. Finally WFS and SOS define transactional operations that allow for creating, updating or deleting resources in the server.

Table 3: Data Exchange services functionalities.

Functionality	WFS 2.0	WCS 2.0	SOS 2.0
Service Metadata	GetCapabilities	GetCapabilities	GetCapabilities
Resources	Feature type names	Coverage names	Observation offerings
Resource description	DescribeFeatureType	DescribeCoverage	DescribeSensor
Retrieve resources	GetFeature	GetCoverage	GetFeatureOfInterest
Retrieve a parts of resources	GetValueProperty	GetCoverage	GetObservation
Create, update or delete resources	Transaction	N/A	InsertSensor, DeleteSensor, InsertObservation

To know more about these data services there is a need to know more about the underlying data model of the resources they use. This is the purpose of the following 3 subsections

5.4.2.5 GML and Web Feature Service

WFS [17] uses the General Feature Model (GFM) and the GML encoding to represent the geospatial features. To understand WFS, first you should understand GML. Unfortunately, understanding GML is not an easy task, in our opinion, due to two main reason: the flexibility of the GFM, combined with the complexities of XML schemas. To represent features in GML, first we need a data model and to encode it into a GML applications schema [18]. In practice this means that features are arranged in feature types. Each feature type will receive a name and will define the list of properties that this feature type has. Let's take the USGS WFS service for transportation data. The GetCapabilities request exposes a list of feature types:

Request:
https://services.nationalmap.gov/arcgis/services/WFS/transportation/MapServer/WFSServer?request=GetCapabilities&service=WFS

List of Feature types: Airport_Runway, Airport, County_Route, Interstate, Local_Road, State_Route, US_Railroad, and US_Route.

To obtain the GML application schema for the feature type "Airport_Runway", we can execute a DescribeFeatureType request [19]: https://services.nationalmap.gov/arcgis/services/WFS/transportation/MapServer/WFSServer?request=DescribeFeatureType&service=WFS&typename=WFS_transportation:Airport_Runway.

In the response (a GML application schema) the feature type tells us that all "runway" features have the same properties and which they are. The most relevant ones are GML MultiSuface geometry (Shape), the FAA name of the airport (FAA_AIRPORT_CODE), the city (CITY) and the runway name (RUNWAY_ID). Now we know what we can expect from the data retrieved and eventually, we could use this information in a GetValueProperty request.

```
<?xml version="1.0" encoding="utf-8" ?>
<xs:schema>
    <xs:element name='Airport_Runway' type='WFS_transportation:Airport_RunwayType'
substitutionGroup='gml:AbstractFeature'/>
        <xs:complexType name='Airport_RunwayType'>
            <xs:complexContent>
                <xs:extension base='gml:AbstractFeatureType'>
                    <xs:sequence>
                        <xs:element name='OBJECTID' type='xs:int'/>
                        <xs:element minOccurs='0' name='Shape' type='gml:MultiSurfacePropertyType'/>
                        <xs:element minOccurs='0' name='PERMANENT_IDENTIFIER'>
                            <xs:simpleType>
                                <xs:restriction base='xs:string'>
                                    <xs:maxLength value='40'/>
                                </xs:restriction>
                            </xs:simpleType>
                        </xs:element>
                        <xs:element minOccurs='0' name='SOURCE_FEATUREID'>
                            <xs:simpleType>...</xs:simpleType>
                        </xs:element>
                        <xs:element minOccurs='0' name='SOURCE_DATASETID'>
                            <xs:simpleType>...</xs:simpleType>
                        </xs:element>
                        <xs:element minOccurs='0' name='SOURCE_DATADESC'>
                            <xs:simpleType>...</xs:simpleType>
                        </xs:element>
                        <xs:element minOccurs='0' name='SOURCE_ORIGINATOR'>
```

```
            <xs:simpleType>...</xs:simpleType>
        </xs:element>
        <xs:element minOccurs='0' name='Data_Security' type='xs:short'/>
        <xs:element minOccurs='0' name='Distribution_Policy'>
            <xs:simpleType>...</xs:simpleType>
        </xs:element>
        <xs:element minOccurs='0' name='LOADDATE' type='xs:dateTime'/>
        <xs:element minOccurs='0' name='FCode' type='xs:int'/>
        <xs:element minOccurs='0' name='RUNWAY_ID'>
            <xs:simpleType>...</xs:simpleType>
        </xs:element><xs:element minOccurs='0' name='FAA_AIRPORT_CODE'>
            <xs:simpleType>...</xs:simpleType>
        </xs:element>
        <xs:element minOccurs='0' name='CITY'>
            <xs:simpleType>...</xs:simpleType>
        </xs:element>
        <xs:element minOccurs='0' name='STATE'>
            <xs:simpleType>...</xs:simpleType>
        </xs:element>
        <xs:element minOccurs='0' name='OwnerType' type='xs:int'/>
        <xs:element minOccurs='0' name='UseStatus' type='xs:int'/>
        <xs:element minOccurs='0' name='SHAPE_Length' type='xs:double'/>
        <xs:element minOccurs='0' name='SHAPE_Area' type='xs:double'/>
      </xs:sequence>
    </xs:extension>
   </xs:complexContent>
  </xs:complexType>
 </xs:element>
</xs:schema>
```

Figure 12. USGS Airport Runway GML applications schema (DescribeFeatureType; simplified)

The three runways of the DCA Washington (Ronald Reagan National Airport) can be retrieved with a GetFeature request limited by airport runway feature type name and bounding box: Request: https://services.nationalmap.gov/arcgis/services/WFS/transportation/MapServer/WFSServer?request=GetFeature&service=WFS&typename=WFS_transportation:Airport_Runway&bbox=-77.092,38.842,-76.992,38.942

WFS services are offering geospatial data that can easily be big. Formulating requests that try to retrieve the complete dataset are normally rejected by services after some minutes of processing. To prevent that, it is very important to know how to limit the GetFeature requests by bounding box (or other restrictive criteria) in order to aim for results that are small enough for the server to handle. In the previous example, we have provided a bounding box that limits the request to the Washington DC area. We see that we got 3 features, one for each runway in the DCA airport. The longest runway is the 01/19 with 5.7km.

```
<wfs:FeatureCollection>
 <gml:member>
  <Airport_Runway gml:id='F10__13355'>...</Airport_Runway>
 </gml:member>
 <gml:member>
  <Airport_Runway gml:id='F10__13507'>...</Airport_Runway>
 </gml:member>
 <gml:member>
  <Airport_Runway gml:id='F10__13526'>
   <OBJECTID>13526</OBJECTID>
   <Shape>
    <gml:MultiSurface srsName="EPSG:3857">
     <gml:surfaceMember>
```

```
                        <gml:Polygon>
                            <gml:exterior>
                                <gml:LinearRing>
                                    <gml:posList> -8575942.8157000002 4701809.001000002
-8575720.8753999993 4699005.684799999 -8575664.9667000007 4699010.873800002 -8575884.6827000007
4701813.5922999978 -8575942.8157000002 4701809.001000002</gml:posList>
                                </gml:LinearRing>
                            </gml:exterior>
                        </gml:Polygon>
                    </gml:surfaceMember>
                </gml:MultiSurface>
            </Shape>
            <PERMANENT_IDENTIFIER>f4156a1d-9259-46d8-abe1-c68b689e2c12</PERMANENT_IDENTIFIER>
            <SOURCE_DATASETID>{49981A0D-FC48-4BBF-872D-8E72CE83EA8B}</SOURCE_DATASETID>
            <SOURCE_DATADESC>Public and Private Airports</SOURCE_DATADESC>
            <SOURCE_ORIGINATOR>Federal Aviation Administration</SOURCE_ORIGINATOR>
            <Data_Security>5</Data_Security>
            <Distribution_Policy>E4</Distribution_Policy>
            <LOADDATE>2016-06-06T14:16:43</LOADDATE>
            <FCode>20100</FCode>
            <RUNWAY_ID>01/19</RUNWAY_ID>
            <FAA_AIRPORT_CODE>DCA</FAA_AIRPORT_CODE>
            <CITY>Washington</CITY>
            <STATE>VA</STATE>
            <OwnerType>1</OwnerType>
            <UseStatus>1</UseStatus>
            <SHAPE_Length>5737.8685680632716</SHAPE_Length>
            <SHAPE_Area>160910.31463661761</SHAPE_Area>
        </Airport_Runway>
    </gml:member>
</wfs:FeatureCollection>
```


Figure 13. USGS Airport Runway seen in the Leaflet WFS test browser (GetFeature response; simplified)

The GML geometrical features can be shown in a feature viewer, with the other properties of the North-South runway.

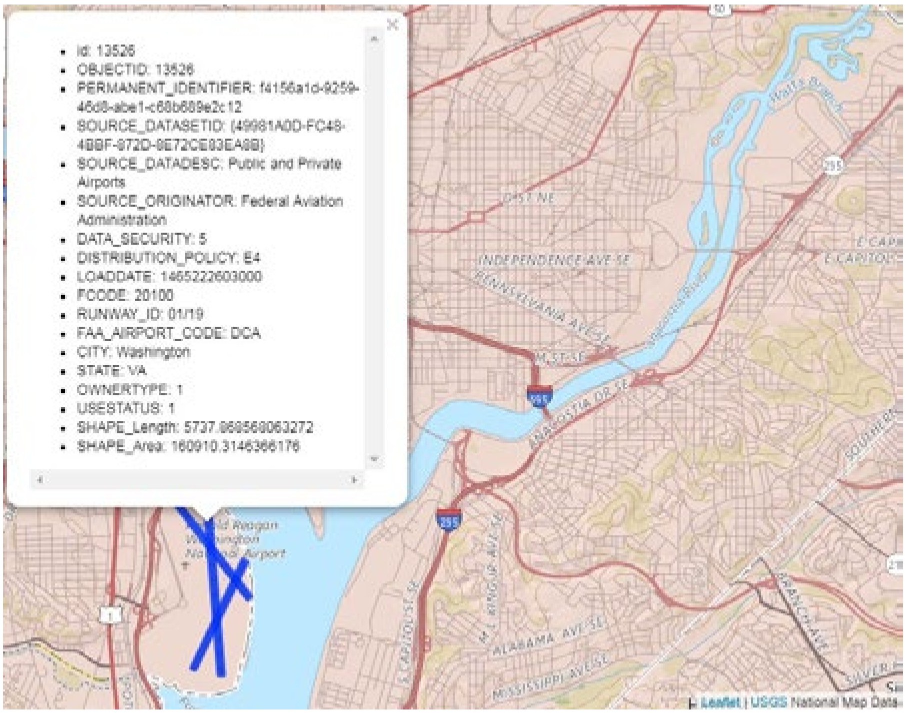


Figure 14. USGS Airport Runway seen in the Leaflet WFS test browser.

In other to simplify the usage of WFS, OGC is experimenting with a new WFS 3.0 that will remove the need for a GML application schema by adopting GeoJSON as a default exchange format. More details about this draft standard will be provided in another subsection.

5.4.2.6 CIS and Web Coverage Service

Commonly, GML is associated to vector data, but actually version 3 introduced raster support. Again GML requires associating a GML application schema to each raster you describe in GML. This is rather tedious a repetitive because most of the raster documents are alike. WCS wanted to adopt GML to describe coverage data but in order to simplify the use of WCS, WCS avoided the need to create a GML application schema by defining a common GML application schema that is valid for all coverages. It is called Coverage Implementation Schema (CIS; formerly known as GMLCov) [20]. Following the CIS model, a coverage can be defined by describing three aspects: DomainSet, RangeType and RangeSet.

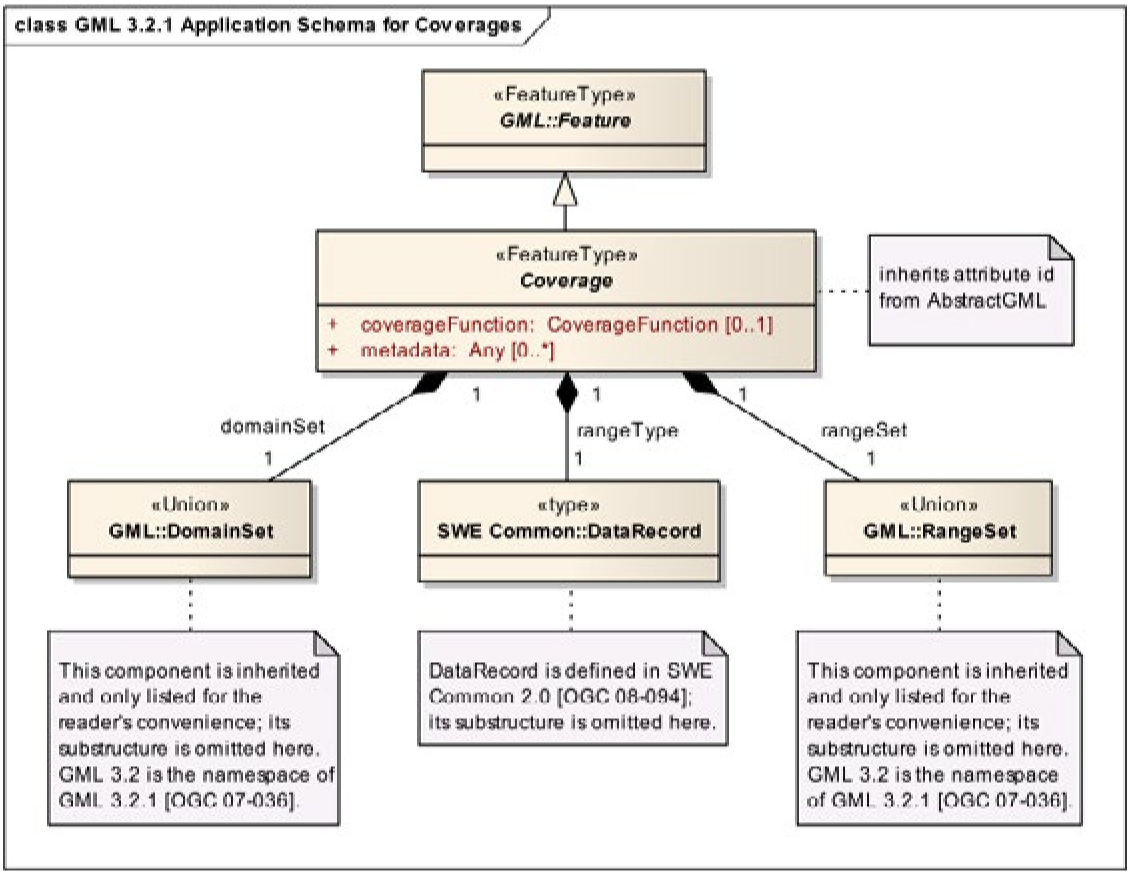


Figure 15. Coverage Implementation Schema (fragment).

In the domainSet, the axes of the multidimensional grid space are defined. In a regular referenced grid, it provides the names of the axes, as well as the origin and the spacing of each axis. In a 2D space it defines two axes that create a regular 2D grid of cells that covers some space, but the cells are still empty and have no values on them. The rangeType defines the types of values (e.g. float numbers), the units and the nodata values that are going to populate the grid. In the rangeSet, the actual list of values that populate the cells is given (e.g. in a DEM it will be a long list of elevations). The domainSet and the rangeType are normally encoded in XML. The rangeSet could also be included in the XML but is commonly referenced to an external raster file such as a GeoTIFF or a JPEG2000 for a more compact notation. This data model has its origins in GML elements that were extracted to define the common application schema for coverages.

In a WCS, information is structured in coverages [21]. A server may have one or more coverage to offer. Let's take the Rasdaman EarthServer demo service. The GetCapabilities request exposes a list of coverages:

Request:
http://ows.rasdaman.org/rasdaman/ows?&SERVICE=WCS&ACCEPTVERSIONS=2.0.1&REQUEST=GetCapabilities

List of Coverage identifiers: AverageChloroColorScaled, AverageChlorophyllScaled, AvgLandTemp, AvgTemperatureColorScaled, BlueMarbleCov, climate_cloud, climate_earth, Germany_DTM, lena, mean_summer_airtemp, meris_lai, multiband, NIR, NN3_1, NN3_2, NN3_3, NN3_4, RadianceColor, RadianceColorScaled, Temperature4D, test_irr_cube_2, test_mean_summer_airtemp, visible_human.

To obtain the domainSet and the rangeType of the CIS schema for the " Germany_DTM", we can execute a DescribeCoverage request:

http://ows.rasdaman.org/rasdaman/
ows?&SERVICE=WCS&VERSION=2.0.1&REQUEST=DescribeCoverage&COVERAGEID=Germany_DTM

```
<?xml version="1.0" encoding="UTF-8"?>
<wcs:CoverageDescriptions>
  <wcs:CoverageDescription gml:id="Germany_DTM">
    <boundedBy>...</boundedBy>
    <wcs:CoverageId>Germany_DTM</wcs:CoverageId>
    <coverageFunction>...</coverageFunction>
    <domainSet>
      <RectifiedGrid dimension="2" gml:id="Germany_DTM-grid">
        <limits>
          <GridEnvelope>
            <low>0 0</low>
            <high>10800 13200</high>
          </GridEnvelope>
        </limits>
        <axisLabels>Lat Long</axisLabels>
        <origin>
          <Point gml:id="Germany_DTM-origin" srsName="http://ows.rasdaman.org/def/crs/
EPSG/0/4326">
            <pos>56.0 5.0</pos>
          </Point>
        </origin>
        <offsetVector srsName="http://ows.rasdaman.org/def/crs/EPSG/0/4326">-0.00083333 0</
offsetVector>
        <offsetVector srsName="http://ows.rasdaman.org/def/crs/EPSG/0/4326">0 0.000833333</
offsetVector>
      </RectifiedGrid>
    </domainSet>
    <gmlcov:rangeType>
      <swe:DataRecord>
        <swe:field name="Elevation">
          <swe:Quantity>
            <swe:label>Elevation</swe:label>
            <swe:uom code="m"/>
          </swe:Quantity>
        </swe:field>
      </swe:DataRecord>
    </gmlcov:rangeType>
    <wcs:ServiceParameters>
      <wcs:CoverageSubtype>RectifiedGridCoverage</wcs:CoverageSubtype>
      <wcs:nativeFormat>application/octet-stream</wcs:nativeFormat>
    </wcs:ServiceParameters>
  </wcs:CoverageDescription>
</wcs:CoverageDescriptions>
```

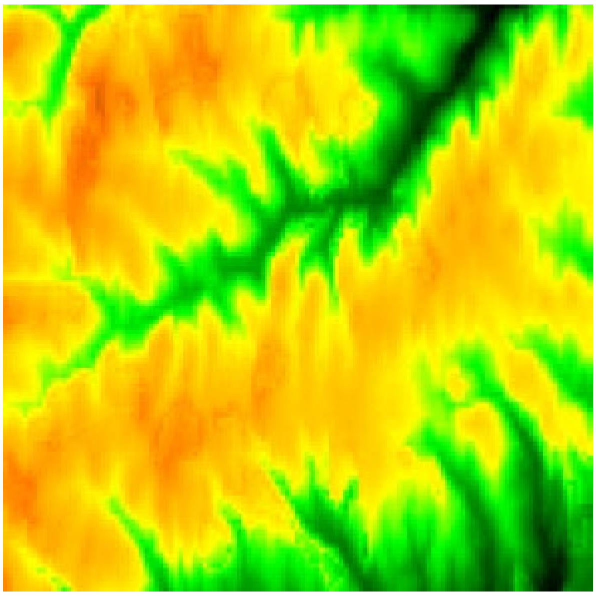
Figure 16. German DTM domainSet and the rangeType (DescribeCoverage response; symplified)

In the description, we can see the definition of a 2D grid of 10800x13200 cells, representing Lat Long that starts at 56.0, 5.0 and has a cell size of 0.00083333 degrees. The cells will be populated with elevations.

The central part of the German DEM can be retrieved with a GetCoverage request indicating the coverageId name and limited by a bounding box: Request: http://ows.rasdaman.org/rasdaman/ows?&SERVICE=WCS&VERSION=2.0.1&REQUEST=GetCoverage&COVERAGEID=Germany_DTM&SUBSET=Lat(50,50.1)&SUBSET=Long(10,10.1)&FORMAT=image/tiff

Again, WCS services are commonly offering geospatial data that is can easily be big. Requests that try to retrieve the complete dataset are normally rejected by services after some minutes of processing. To prevent that, it is very important to know how to limit the GetCoverage requests by bounding box (or other restrictive criteria) in order to anticipate a result that is small enough for the server to handle [22]. In this case we use the subsetting extension of WCS 2.0 to do request a subset for each axis .

The result of this operation will be a small tiff file with the elevations of the requested area. Changing FORMAT=image/tiff to FORMAT=application/gml+xml in the GetCoverage request will result in a full CIS document describing the domainSet, the rangeType, and providing a long list of integer numbers as the rangeSet.



```
<boundedBy>
 <Envelope axisLabels="Lat Long"
srsDimension="2">
   <lowerCorner>49.99958333 9.99958333</
lowerCorner>
    <upperCorner>50.10041667 10.10041667</
upperCorner>
      </Envelope>
   </boundedBy>
   <domainSet>
     <RectifiedGrid dimension="2">
      <limits>
        <GridEnvelope>
```

Figure 17. German DEM fragment shown as a raster and also in a CIS fragment

5.4.2.7 Discovery services

Web search engines are not so good finding geospatial information on the web so specific catalogues for geospatial data discovery were needed. These catalogues cannot be based on the data itself that is, in most cases, not self-describing, but on descriptions about the data (called metadata). A discovery service is a metadata catalogue that has a query interface to request the kind of data that the user needs. The catalogue compares the query parameters with the metadata previously stored and creates an answer in the form of a list of items that potentially can be of the interest of the user. OGC has produced several standards for geospatial catalogue services: One of the most popular is the Catalogue Service for the Web (CSW) [28] based on ISO 19115 metadata [29]. Some people criticize the ISO metadata arguing that it is a too complex model and simpler solutions based on Dublin core has also been produced. Query languages have been also simplified, and OGC has also proposed and extension of the common OpenSearch query language making it spacio-temporal aware: OpenSearchGEO.

This standard works in a very similar way to WFS but the resources it deals with are database "records" instead of geospatial "features". Since there are several variants of CSW catalogue standards, the service metadata document provides an overview of the capabilities of this service including the metadata encodings delivered and the filter languages supported.

Request: http://www.ign.es/csw-inspire/srv/spa/csw?service=CSW&request=GetCapabilities


```xml
<?xml version="1.0" encoding="UTF-8"?>
<csw:Capabilities>
 <ows:ServiceIdentification>...</ows:ServiceIdentification>
 <ows:ServiceProvider>...</ows:ServiceProvider>
 <ows:OperationsMetadata>
  ...
  <ows:Operation name="GetRecords">
   <ows:Parameter name="outputSchema">
    <ows:Value>http://www.opengis.net/cat/csw/2.0.2</ows:Value>
    <ows:Value>http://www.isotc211.org/2005/gmd</ows:Value>
   </ows:Parameter>
   <ows:Parameter name="typeNames">
    <ows:Value>csw:Record</ows:Value>
    <ows:Value>gmd:MD_Metadata</ows:Value>
   </ows:Parameter>
   <ows:Parameter name="CONSTRAINTLANGUAGE">
    <ows:Value>FILTER</ows:Value>
    <ows:Value>CQL_TEXT</ows:Value>
   </ows:Parameter>
   <ows:Constraint name="SupportedISOQueryables">
    <ows:Value>Operation</ows:Value>
    <ows:Value>Format</ows:Value>
    <ows:Value>OrganisationName</ows:Value>
    ...
    <ows:Value>Abstract</ows:Value>
   </ows:Constraint>
   <ows:Constraint name="AdditionalQueryables">
    <ows:Value>Relation</ows:Value>
    <ows:Value>AccessConstraints</ows:Value>
    <ows:Value>ResponsiblePartyRole</ows:Value>
    <ows:Value>OnlineResourceMimeType</ows:Value>
    <ows:Value>OnlineResourceType</ows:Value>
    ...
   </ows:Constraint>
  </ows:Operation>
 </ows:OperationsMetadata>
 <ogc:Filter_Capabilities>
  <ogc:Spatial_Capabilities>
   <ogc:GeometryOperands>
    <ogc:GeometryOperand>gml:Envelope</ogc:GeometryOperand>
    <ogc:GeometryOperand>gml:Point</ogc:GeometryOperand>
    <ogc:GeometryOperand>gml:LineString</ogc:GeometryOperand>
    <ogc:GeometryOperand>gml:Polygon</ogc:GeometryOperand>
   </ogc:GeometryOperands>
   <ogc:SpatialOperators>
    <ogc:SpatialOperator name="BBOX" />
    <ogc:SpatialOperator name="Equals" />
    <ogc:SpatialOperator name="Overlaps" />
    <ogc:SpatialOperator name="Disjoint" />
    <ogc:SpatialOperator name="Intersects" />
    <ogc:SpatialOperator name="Touches" />
    <ogc:SpatialOperator name="Crosses" />
    <ogc:SpatialOperator name="Within" />
    <ogc:SpatialOperator name="Contains" />
   </ogc:SpatialOperators>
  </ogc:Spatial_Capabilities>
  <ogc:Scalar_Capabilities>
```

```xml
    <ogc:LogicalOperators />
    <ogc:ComparisonOperators>
      <ogc:ComparisonOperator>EqualTo</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>Like</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>LessThan</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>GreaterThan</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>LessThanEqualTo</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>GreaterThanEqualTo</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>NotEqualTo</ogc:ComparisonOperator>
      <ogc:ComparisonOperator>Between</ogc:ComparisonOperator>
    </ogc:ComparisonOperators>
  </ogc:Scalar_Capabilities>
  <ogc:Id_Capabilities>
    <ogc:EID />
    <ogc:FID />
  </ogc:Id_Capabilities>
 </ogc:Filter_Capabilities>
</csw:Capabilities>
```

Figure 18. CSW service metadata document response (fragment)

In the same way WFS has a DescribeFeature operation that returns the XML schema (GML application schema) for the requested features, the DescribeRecord operation returns the XML schema for the metadata records. The figure shows a service supporting both Dublin CORE and ISO 19115 metadata records.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<csw:DescribeRecordResponse>
 <csw:SchemaComponent targetNamespace="http://www.opengis.net/cat/csw/2.0.2"
schemaLanguage="http://www.w3.org/XML/Schema">
  <xsd:schema version="2.0.2">
    <xsd:import namespace="http://purl.org/dc/terms/" schemaLocation="rec-dcterms.xsd" />
    <xsd:import namespace="http://purl.org/dc/elements/1.1/" schemaLocation="rec-dcmes.xsd" />
    <xsd:complexType name="SummaryRecordType" final="#all">
     <xsd:complexContent>
      <xsd:extension base="csw:AbstractRecordType">
       <xsd:sequence>
        <xsd:element ref="dc:identifier" minOccurs="1" />
        <xsd:element ref="dc:title" minOccurs="1" />
        <xsd:element ref="dc:type" minOccurs="0" />
        <xsd:element ref="dc:subject" minOccurs="0 />
        <xsd:element ref="dc:format" minOccurs="0" />
        <xsd:element ref="dc:relation" minOccurs="0" />
        <xsd:element ref="dct:modified" minOccurs="0" />
        <xsd:element ref="dct:abstract" minOccurs="0" />
        <xsd:element ref="dct:spatial" minOccurs="0" />
        <xsd:element ref="ows:BoundingBox" minOccurs="0" />
       </xsd:sequence>
      </xsd:extension>
     </xsd:complexContent>
    </xsd:complexType>
  </xsd:schema>
 </csw:SchemaComponent>
 <csw:SchemaComponent targetNamespace="http://www.opengis.net/cat/csw/2.0.2"
schemaLanguage="http://www.w3.org/XML/Schema">
  <xs:schema version="0.1">
    <xs:import namespace="http://www.isotc211.org/2005/gco" schemaLocation="../gco/gco.xsd" />
    <xs:complexType name="AbstractMD_Identification_Type" abstract="true">
     <xs:complexContent>
```

```
        <xs:extension base="gco:AbstractObject_Type">
         <xs:sequence>
          <xs:element name="citation" type="gmd:CI_Citation_PropertyType" />
          <xs:element name="abstract" type="gco:CharacterString_PropertyType" />
          <xs:element name="purpose" type="gco:CharacterString_PropertyType" minOccurs="0" />
          <xs:element name="credit" type="gco:CharacterString_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="status" type="gmd:MD_ProgressCode_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="pointOfContact" type="gmd:CI_ResponsibleParty_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="resourceMaintenance" type="gmd:MD_MaintenanceInformation_PropertyType"
minOccurs="0" maxOccurs="unbounded" />
          <xs:element name="graphicOverview" type="gmd:MD_BrowseGraphic_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="resourceFormat" type="gmd:MD_Format_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="descriptiveKeywords" type="gmd:MD_Keywords_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="resourceSpecificUsage" type="gmd:MD_Usage_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="resourceConstraints" type="gmd:MD_Constraints_PropertyType" minOccurs="0"
maxOccurs="unbounded" />
          <xs:element name="aggregationInfo" type="gmd:MD_AggregateInformation_PropertyType"
minOccurs="0" maxOccurs="unbounded" />
         </xs:sequence>
        </xs:extension>
       </xs:complexContent>
      </xs:complexType>
    </xs:schema>
  </csw:SchemaComponent>
</csw:DescribeRecordResponse>
```

Figure 19. CSW service metadata document response (simplyfied)

Finally, it is possible to retrieve metadata records in ISO XML encoding by specifying the desired schema (resultType parameter) and providing a filter. The simplest filter that emulates the Google search text query is "Like". In the following KVP request, the CQL query language is selected and the "like" query is formulated wiht the word "topo" in between % to indicated that the text is only a fragment of a work; obviously we are aiming to discover "TOPOgraphic" maps.

http://www.ign.es/csw-inspire/srv/spa/
csw?version=2.0.2&service=CSW&request=GetRecords&resultType=results&outputSchema=http://www.isotc211.org/2005/gmd&constraintlanguage=CQL_TEXT&constraint="csw:AnyText Like '%topo%'"

The result is a collection of records.

```
<?xml version="1.0" encoding="UTF-8"?>
<csw:GetRecordsResponse>
 <csw:SearchStatus timestamp="2018-08-26T07:36:21" />
 <csw:SearchResults numberOfRecordsMatched="5110" numberOfRecordsReturned="10"
elementSet="summary" nextRecord="11">
   <gmd:MD_Metadata>
    <gmd:fileIdentifier>
     <gco:CharacterString>spaignCartografiaHistoricalGN201308020055
     </gco:CharacterString>
    </gmd:fileIdentifier>
    ...
   </gmd:MD_Metadata>
```

...
</csw:SearchResults>
</csw:GetRecordsResponse>

Figure 20. CSW service metadata document response (fragment)

Catalogues have some other characteristics, such as the possibility to harvest other catalogues that are not explained here.

5.4.3 Integrated geospatial client

The Geospatial web services described in the previous sections are pieces of the puzzle that need to be built when developing a geospatial client interface with a user. Users of distributed geospatial information should not know about the complexities of the application protocol and they should have the illusion that all happens in a single environment. They should not normally realize that they are accessing distributed data except, perhaps, if there are Internet access problems. To achieve this, an integrated client should automate the service connections and should be able to handle different protocols and formats in a transparent way.

The figure presents a diagram of an integrated client that uses WMS, WFS and SOS requests simultaneously. WMS is used to get orthophotographic imagery, WFS provides nautical charts and SOS is used to get positions of some whale observations. WMS GetMap, WFS GetFeature and SOS GetFeatureOfInterest are formulated to their respective services, each one returning a different format. The client is able to interpret the document formats and present the geospatial information overlapped in the screen. Each time that the user does a zoom or a pan, the whole cycle is repeated. The client should be designed as a multithread application where WMS, WFS and SOS request can happen simultaneously and they are presented in the screen as soon as one of the responses arrives. Due to different delays in the network, sometimes, the SOS observations are represented first, and some other time the WMS imagery is the one that appears before, but in the end, all the necessary information will emerge and the user will have a more fluid experience with the application. If the integrated client is a JavaScript map browsers, implementing multithread applications is relatively easy to program by combining the following strategies: the <img> elements are loaded in an independent thread. GetTimeOut() functions are executed in a new thread and AJAX requests can be asynchronous.

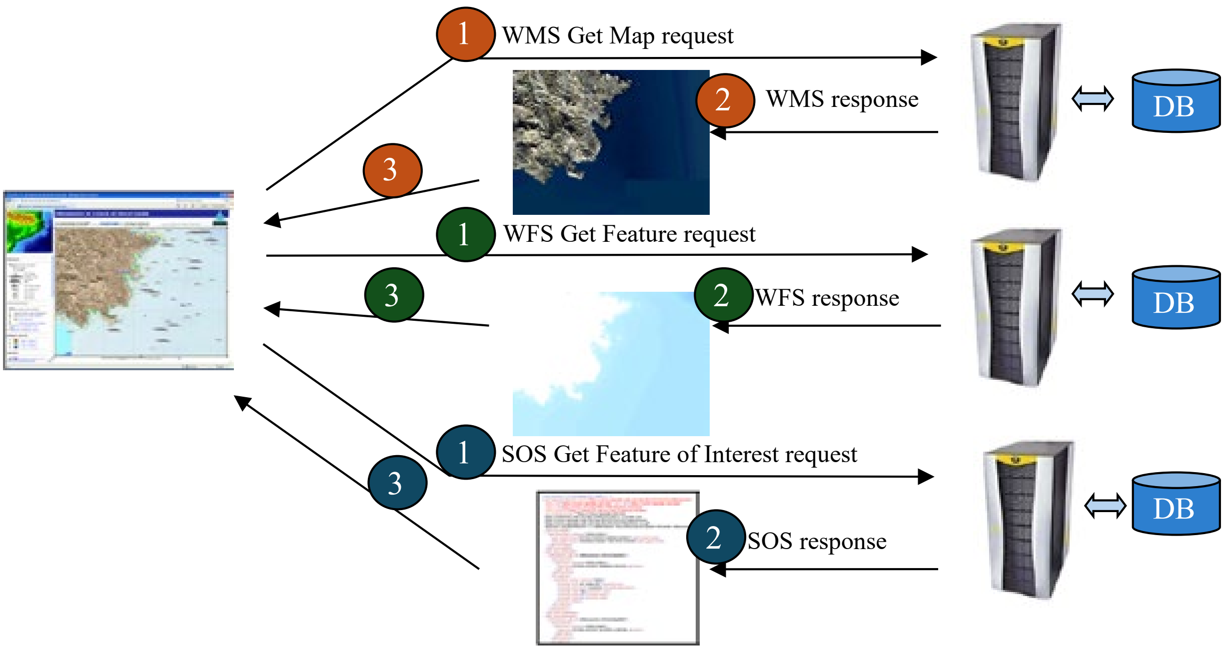


Figure 21. Integrated client with a WMS, a WFS and a SOS

Today, modern GIS applications can also act as geospatial integrated clients and data coming from distributed geospatial services can be combined with local data in an integrated way.

OGC has recently released a standard that was created to allow a set of configured information resources (service set) to be passed between applications primarily as a collection of services as well as in-line content. The goal of OGC OWS Context is to support the distribution of search results, or the exchange of a set of visualized resources such as WFS, WMS, WMTS, WCS and others services included in a common operating picture. Additionally, OWS Context can deliver a set of configured processing services (WPS) and the parameters to allow the reproduction of the processing on different nodes. OWS Context is replacing a previous OGC attempt at providing such a capability (the Web Map Context, WMC) which was limited to WMS.

5.4.4 Deploying Web services

If you have data to offer, you can deploy your own geospatial server. There are many vendors that provide software solutions that implement web services. As a starting point we need a web server where we can deploy the geospatial web server application. The operating system of modern desktop computers can incorporate a web server. In Linux, Apache is the usual alternative and Internet Information Server comes by default in modern versions of Windows. In our experience, adding a geospatial web service to them requires some knowledge on how a web server works and several configuration steps. Instead of using the current web server already installed, it is easy to look for a download package that incorporates their own web server already integrated and configured with the geospatial web server. To illustrate how to setup a service, we will user GeoServer that is an open source implementation for WMS, WMTS, WCS and WFS server, also having extensions for CSW and WPS. GeoServer offers a Windows distribution with a web server integrated that is very easy to setup and configure. At the time to write these lines, the GeoServer official URL for download is http://geoserver.org/release/stable/. The only prerequisite in GeoServer is that it needs the Java Development Kit installed in the machine, before being able to run it (the JDK is not the same as the JRE that is commonly used by many other applications).

Actually, GeoServer default setup automatically starts a geospatial server that includes some sample data. This is useful to demonstrate the correct performance of your new installation but does not help publishing your own data. GeoServer offers a web interface to configure the data that you want to expose as a service. To publish a shape file as a WMS service you should do the following steps:

1. Generate a new workspace for your data. This will separate your "things" from the demo data that is already included.
2. Copy your shape file in the (geoserver install)/data_dir/data/
3. Create a new data source. GeoServer supports many data sources, both files and geodatabases. A shapefile is one of the options under vector data sources.
4. Create a new style. This is done by providing a SLD file that indicates how data should be rendered.
5. Create a new layer. In this process, you select a data source and associate it to a style.
6. Test you new layer as a web service in the layer preview option

Now, your geospatial web server has been set and it is ready to be used by a standards client. The URL of your WMS service is http://my.server.name:8080/geoserver/wms and the data is available as other services such as a WFS service in http://my.server.name:8080/geoserver/wfs.

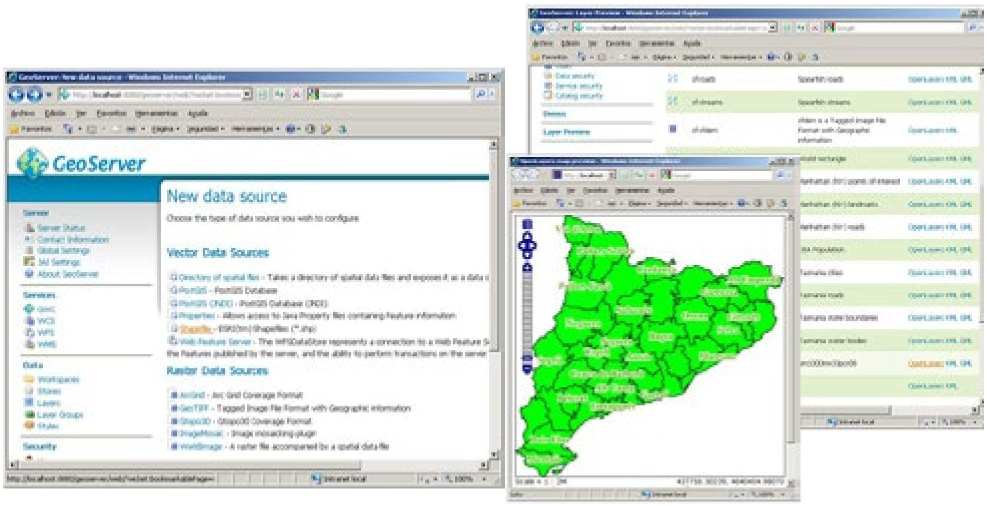


Figure 22. Two steps to configure a layer in a WMS in GeoServer. Setting a data source and previewing the result (the administrative limits of Catalonia shown in the figure are produced by the Institut Cartografic i Geològic de Catalunya)

5.4.5 The emergence of APIs

The services presented in previous subsections are part of the OGC Web Services approach that define implementation level communication protocols between client and server on top of the HTTP application layer. For web developers, it is more natural to import an Application Program Interface (API) than to understand a new communication layer. Originally, the term API was used as a synonymous of "a library to create user interfaces". Today, the term is still used in this sense also for libraries that facilitate the creation of map browser user interfaces. However, the same term has recently been applied to well documented RESTful interfaces that present list of resources and operations. These two different approaches are explained in the following two subsections.

5.4.5.1 Geospatial APIs for creating user interfaces.

Integrated clients implement the complexities of geospatial services and protocols in a way that final users do not need to be aware of the complexities of the distributed computing architecture. In the same way, some software libraries provide APIs warping the complexities of geospatial services and protocols for software developers too. In this case, the software developers use a API to create map applications that run on web browsers. This type of API provides a set of JavaScript functions to simplify the creation of a user interface for presenting and querying geospatial data in a web page. Internally, the library uses geospatial web communications (open standards or proprietary protocols) but developers only need to know the URL of the services and the layers available. By using the API, they will not need to implement the service protocol themselves as the API will execute it for them.

To illustrate how this kind of user interface APIs work, we present a small example with Leaflet (https://leafletjs.com/reference-1.7.1.html). Leaflet is a JavaScript API for building interactive map browsers that is able to communicate with many services including OGC WMS services. In figure 16.32, we can see a fragment of code where the API is imported by adding a style-sheet (leaflet.css) and a script (leaflet.js). The API interacts with a <div> area defined in a HTML page and draws a map inside. The first line of code creates an empty "map" in a region of the Earth defined by some coordinates situated over Barcelona and associates it to the <div> identifier. The second line of code adds a WMS layer to the "map" using the URL of the server (from the Catalan Cartographic and Geological Institute), the layer name and the file format as input. As of July 2021, Leaflet current version was 1.7.1 but the example is backwards compatible, back to version 0.7.3 or below. The library manages the actions of the user on the map and transforms them into client-server communication messages. Web developers do not need to be aware of the communication details and only need to know how JavaScript works and how to use the library.

```
<html>
<head>
	<link rel="stylesheet" href="https://unpkg.com/leaflet@1.7.1/dist/leaflet.css" />
</head>
<body>
	<div id="map" style="width: 600px; height: 400px"></div>
	<script src="https://unpkg.com/leaflet@1.7.1/dist/leaflet.js"></script>
	<script>
		var map = L.map('map').setView([41.90, 2.15], 7);
	L.tileLayer.wms("http://geoserveis.icc.cat/icc_mapesbase/wms/service", {
			layers: 'orto5m',
			format: 'image/png',
			transparent: true,
			attribution: "ICGC © 2021"
		}).addTo(map);
	</script>
</body>
</html>
```

Figure 23. Leaflet API used in a JavaScript code that internally uses the WMS protocol (fragment)

5.4.5.2 Geospatial Web APIs

This kind of APIs provides a completely different approach. In this case, the implementation level protocol is completely removed and the HTTP application level protocol in a ROA style (RESTful) is used instead. The focus is on defining resources, (associating them to URIs) that are transmitted from or to a server using HTTP verbs (with the meaning specified in section 16.2.4). In the end, the API defines a list of URI templates for resources, the valid values for each variable in the URI templates, and which HTTP verbs can be applied to each resource. All this information can be formally presented in an API description document. The OpenAPI 3.0 document format (formerly known as Swagger) is often used as a standard way to provide API descriptions.

The OGC has embraced a process of transforming the OGC Web Service standards (WMS, WFS, WCS, WPS, etc.) into an OGC Web API family of standards that takes advantage of modern web development practices. These standards are built in a coordinated way that will allow for a single integrated API implementation potentially combine characteristics of all the OGC Web API standards. At the same time, the OGC Web API standards are constructed as a set of small "building blocks" that can be combined with other functionalities and resources (including others non specified by the OGC) to assemble novel APIs for accessing a combination of geospatial and non-geospatial content. These building blocks are defined not only by the requirements of the specific standards, but also through interoperability prototyping and testing in OGC's Innovation Program.

As of July 2021, the OGC API – Features (the evolution of the WFS) has been approved by the OGC but many other groups are working in the same direction on several API standard candidates such as the OGC API – Tiles (the evolution of the WMTS), the OGC API – Maps (the evolution of the WMS), the OGC API – Records (the evolution of the CSW), the OGC API – Coverages (the evolution of the WCS), the OGC API – Styles (the evolution of the SLD), the OGC API – Processes (the evolution of the WPS), and the Environmental Data Retrieval (a new convenience API focusing on retrieving environmental data inside points, areas, cubes, trajectories or corridors). The up-to-date panorama of the OGC Web APIs can be found in the initiative page: https://ogcapi.ogc.org.

The OGC Web API family uses a modular approach where each building block is defined as a set of requirements being part of a requirements class. For convenience, some requirements classes are grouped in documents that are called "Parts". Each OGC Web API standard is composed by one or more parts. For example at the time of writing these lines the OGC API – Features has released 2 parts: OGC API – Features – Part 1: Core (OGC 17-069r3) and OGC API – Features – Part 2: Coordinate Reference Systems by Reference (OGC 18-058), and more are expected to be approved soon. To ensure that the OGC Web API all building blocks are compatible and efforts are well align, the OGC is standardizing the OGC API – Common that provides fundamental building blocks for many APIs and developed OGC Web API Guidelines to be followed of the OGC Web API standards.

In the OGC Web APIs, there are three main ways of discovering resources: an API description document, the "follow the links" approach and the predictable common paths. An OpenAPI 3.0 document (https://swagger.io/specification/) is a useful API description document (encoded in JSON or YAML) that lists the available resources in the form of URI templates associated to the HTTP verbs supported for each kind of resource (see Fig. 16.33). OpenAPI 3.0 is the only specified API description document included so far in the current OGC Web API standards. The "follow the links" approach uses links included in many resources starting by the landing page, in the same way that HTML web pages do. Figure 16.34 shows a fragment of the landing page encoded in JSON that contains some links including one pointing to the list of "collections". The "collections" resource will include more links and the linked resources even more links and so on. In particular, the "collections" resource lists the identifiers of each and every available "collection" and provides links pointing to all of them. For each link, there is a "rel" property that describes the type and purpose of the relation. Possible values for the "rel" property can be found in the IANA link relation types web page (https://www.iana.org/assignments/link-relations/link-relations.xhtml. E.g.: "collection") or in the OGC Naming Authority server (http://www.opengis.net/def/rel. E.g.: "http://www.opengis.net/def/rel/ogc/1.0/conformance"). Figure 16.35 shows a graph representing a set of relations included by several OGC Web API standards, informs on how resources (represented as ellipses) are linked together and presents the link relation types used for each link in a label next to the arrows connecting the ellipses. The third way to discovering resources is to try out the common paths and URI templates that are exemplified in the different OGC Web API standards. Most of those paths are optional and server API developers can decide to favor another configuration. However, it is foreseen that most of the developers will use these common URI templates as the bases for their implementations. In some cases, such as the OGC API – Features – Part 1, these fixed paths are considered mandatory.

“Collection” is one of the most important resources in the OGC Web API family of standard because this shared resource gives access to a group of geospatial data available thought the API in several forms. Depending on the number of OGC Web APIs implemented by an API instance (that are listed in the conformance page) and the data capabilities, a “collection” can be retrieved as:

- simple features (e.g. in a vector format)
- a coverage (e.g in raster format; if it is a regular grid coverage),
- maps (by subsetting a styled map representation)
- tiles (by dividing the data in a tileset).

The table 16.4 gives an overview of some of the fixed paths suggested by the OGC Web API family of standards

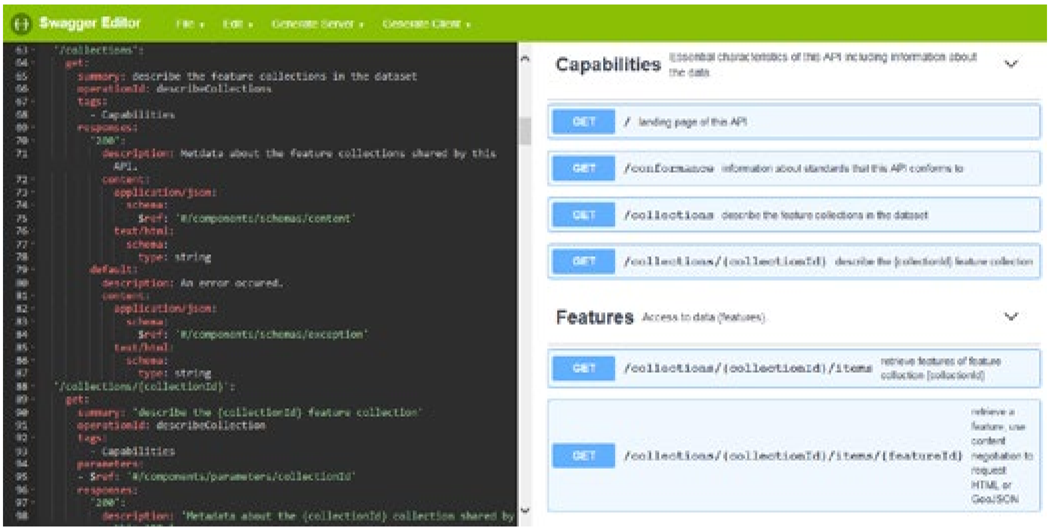


Figure 24. OpenAPI 3.0 document example for an Web API implementing OGC API – Features

---


```
{
    “title”: “Buildings in Bonn”,
    “description”: “Access to data about buildings in the city of Bonn via a Web API that conforms to the OGC API Features specification.”,
    “links”: [
        {
            “href”: “http://data.example.org/api“,
            “rel”: “service-desc”,
            “type”: “application/vnd.oai.openapi+json;version=3.0”,
            “title”: “the API definition”
        },
        {
            “href”: “http://data.example.org/conformance”,
            “rel”: “conformance”,
            “type”: “application/json”,
            “title”: “OGC API conformance classes implemented by this server”
        },
        {
            “href”: “http://data.example.org/collections”,
            “rel”: “data”,
            “type”: “application/json”,
            “title”: “Information about the feature collections”
        }
    ]
}
```


Figure 25. A fragment of a JSON representation of a landing page conformal to the OGC API – Features

---

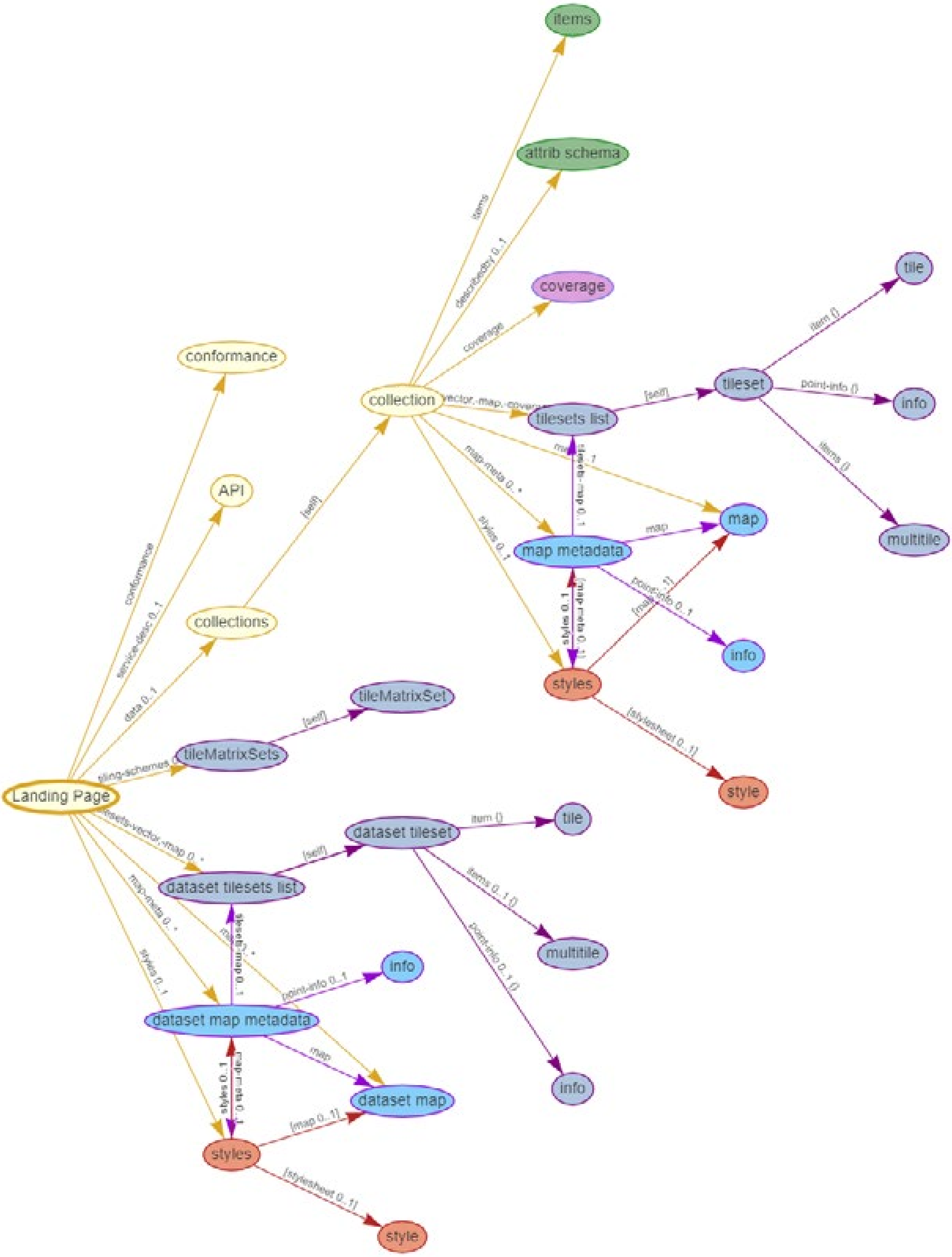


Figure 26. OGC Web API resources (ellipses) and link relations (arrows) graph

Table 4: Correspondence between OGC Web API URIs and OWS Web Service operations

OGC Web API		**OWS Web Service**
URI template	Resource retrieved	Correspondence
/	Landing page	OWS Common GetCapabilities ServiceProvider section
/api	Open API document	OWS Common GetCapabilities OperationsMetadata section
/conformance	List of standard and profiles that this service is honoring	OWS Common GetCapabilities ServiceIndentification section
/collections	Enumeration of "groups" of the data ("collection") present in service	WFS GetCapabilities Content section (FeatureType grouping concept)
/collections/{collectionsId}	Details about a collection	WFS DescribeFeatureType response
/collections/{collectionId}/items	Feature collection description	WFS GetFeature operation selecting a TypeName
/collections/{collectionId}/items/{featureId}	Feature	WFS GetFeatureById response
/collections/{collectionId}/ / tiles/{tileMatrixSetId/{tileMatrix}/{tileRow}/{tileCol}	Tile	WMTS GetTile response
/collections/{collectionId}/ styles/{styleId}/map	Map	WMS GetMap response
/collections/{collectionId}/coverage	Coverage	WCS GetCoverage response
/collections/{collectionId}/coverage/domainSet	CIS domainSet of a coverage	WCS DescriveCoverage domainSet section
/collections/{collectionId}/coverage/rangeType	CIS rangeType of a coverage	WCS DescriveCoverage rangeType section
/styles/{styleId}	SLD/SE style description	WMS SLD GetStyles

Even if many URI templates have a logical correspondence to an equivalent OGC Web Service functionality, there are many practical differences. To cite one difference, OGC Web API favors the use of JSON encodings instead of XML encodings. For example, the OGC API – Features – Part 1 uses GeoJSON as the default format (it still provides optional conformance classes for two XML, based GML simple features profiles 0 and 2). In addition, some of the OGC Web API standards are using the concept of building blocks to specify fragments of resource paths. For example, the path fragment /tiles/{tileMatrixSetId/{tileMatrix}/{tileRow}/{tileCol} can be combined with features to get vector tiles, with a map to get map tiles or with a coverage to get tiles from a coverage (see Table 16.5)

Table 5 The tile path fragement building block combined with other resources

URI template	Meaning	OWS Correspondence
/collections/{collectionId}/ /tiles/{tileMatrixSetId/{tileMatrix}/{tileRow}/{tileCol}	retrieves a vector tile	WMTS GetTile specifying a vector tile format
collections/{collectionId}/map/{Styles)/ /tiles/{tileMatrixSetId/{tileMatrix}/{tileRow}/{tileCol}	retrieves a map tile	WMTS GetTile common use
collections/{collectionId}/coverage/ /tiles/{tileMatrixSetId/{tileMatrix}/{tileRow}/{tileCol}	retrieves a coverage tile	N/A

There will be much more functionalities that are not explained in this section, such as the capacity to create resources that are "jobs" that creates new resources by executing "processes" (a capacity provided by the OGC API – Processes.

5.5 Integration of data using APIs.

Geospatial data integration is the process of collecting data from various sources, with different sources and types (e.g. remote sensing, in-situ, statistics), and unifying them in a single database to create a unified environment for processing, modeling, and visualization. This integration process requires some processing steps transforming coordinate reference systems, data projections, data models and data formats of the different data sources into one common CRS, data model and format. This geospatial data integration can be done permanently in a single repository or can be achieved dynamically on demand. Disparity of licences and contractual agreements, and intellectual property rights frameworks can pose challenges to seamless data integration. Data model coordination, semantic interoperability, different data quality levels, and lack of metadata can pose other difficulties in data integration and increase the effort to ensure successful integration. There are also cultural factors that prevent data integration such as risk-averse attitudes and siloed thinking towards sharing data across different sectors that prevent collaboration and hinder the integration of geospatial data. In particular, companies may treat data as confidential to safeguard their commercial models or competitive advantage, hindering open sharing and integration efforts. Addressing these barriers requires a multi-faceted approach that involves addressing technical, legal, privacy, and cultural challenges to foster effective geospatial data integration for improved decision-making and analysis.

5.5.1 Spatial and statistical data integration

In most countries, mapping and cadastral authorities and statistical offices already cooperate to varying degrees. However, there is a lack of common approaches and general best practices. Integrating geographical and statistical information offers significant opportunities to maximise the utility of data collected for statistical purposes. Key phenomena that need to be precisely understood at local level also need to be more accurately measured, analysed and communicated.

The integration of spatial and statistical data is an important research area that requires collaboration across multiple disciplines, including geography, computer science, statistics, and various domain-specific fields.

The 2020 vision of the European statistical system states that there is a growing need to develop statistics with increasing geographical detail to support national and regional policy making. Statistics with greater geographical detail are needed, for example, in the development, implementation, monitoring and evaluation of EU policies.

Geographic information used in combination with statistics and new data sources such as Big Data and remote sensing can make a substantial contribution to meeting this objective, in particular for social and environmental information. Action taken over the last couple of years includes:

- geo-referencing of official registers and administrative data used for statistics. This was done to achieve better linkage and integration of statistics and geography evaluation of Big Data and Linked Data for statistical production. Examples of official registers used for statistical purposes are: address registers, registers of people, farm registers, tax registers and business registers.
- Advances in technologies like GIS, geocoding, and digital data management have enabled more sophisticated integration of spatial and statistical data in recent years
- Integrating spatial and statistical data allows for more accurate census cartographies, advanced spatial-temporal analysis, and the visualization of non-geographical data in a spatial context
- Developing standardized frameworks, vocabularies, and tools for spatial data integration, such as the OGC Table Joining Service, can help overcome these challenges and enable more seamless data interoperability

Key challenges in spatial data integration include dealing with differences in data formats, scales, accuracy, and classification schemes, as well as managing uncertainty and ambiguity in the data

5.5.1.1 Methods for spatial and statistical data integration

There are several methods most frequently used in the agencies for integrating statistical and geospatial data and taken from the national examples. Those represent different approaches to geocoding, as well as the development of generic spatial join-operations, the usage of 'Linked Data' 20 and 'Ontologies'.

The point-based' system is a simple method to enhanced statistical geospatial interoperability and data integration. The rationale of a point-based system is that it enables geocoding as basis of data integration based on location such as point coordinates. Interaction of statistical and geospatial data with the point-based system can be seen on two levels: the point-based geometries for data integration, and the expansion to the record unit level with Persistent Unique Identifiers.

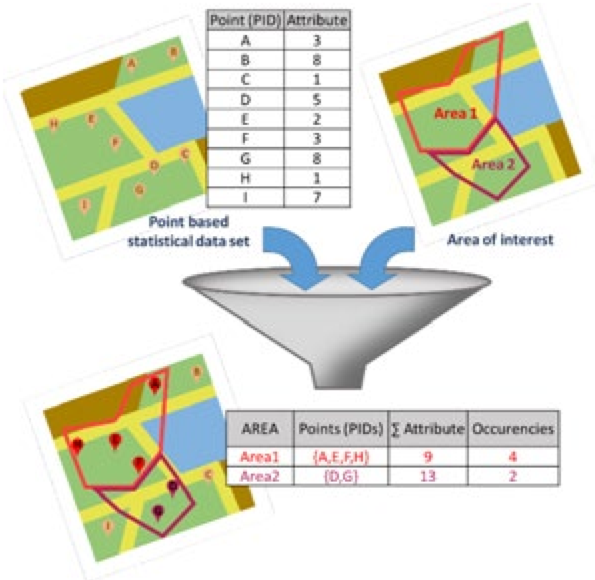


Figure 27. Concept of point based system integration

For example, Spain uses it to overcome a lack of data harmonization. The method allows the NSO to extract the cadastre data that are affecting the NSO without any change of data model in each institution

Statistical tables as well as attribute data are usually labelled with area identifiers, which are also indirect references to location and geospatial data. A standard method to merge geospatial portrayal of different area divisions with table-based attribute data is described by the OGC Standard Table Joining Service (TJS). TJS as such is a single-case integration method using common unique identifiers as links with the data to be integrated and may need a tender process between data providing organizations beyond open data.

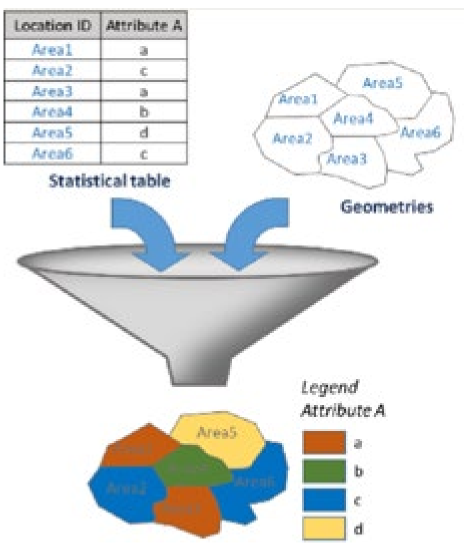


Figure 28. Concept of area base data integration

The OGC Table Joining Service (TJS) is a standard that defines a simple way to describe and exchange tabular data that contains location-based information. It provides a standardized interface for services that allow joining attribute data stored in one database with corresponding geospatial data stored in another database or system. The TJS standard enables the integration of tabular data (e.g., spreadsheets, CSV files) with geospatial data, allowing users to combine and analyze these disparate data sources. To allow that, it defines a structured way to represent the metadata and schema of the tabular data, including information about the columns, data types, and relationships to geographic features. By providing a standardized approach to joining tabular and geospatial data, the OGC Table Joining Service standard helps to improve data interoperability, facilitate data integration, and enable more comprehensive geospatial analysis and decision-making.

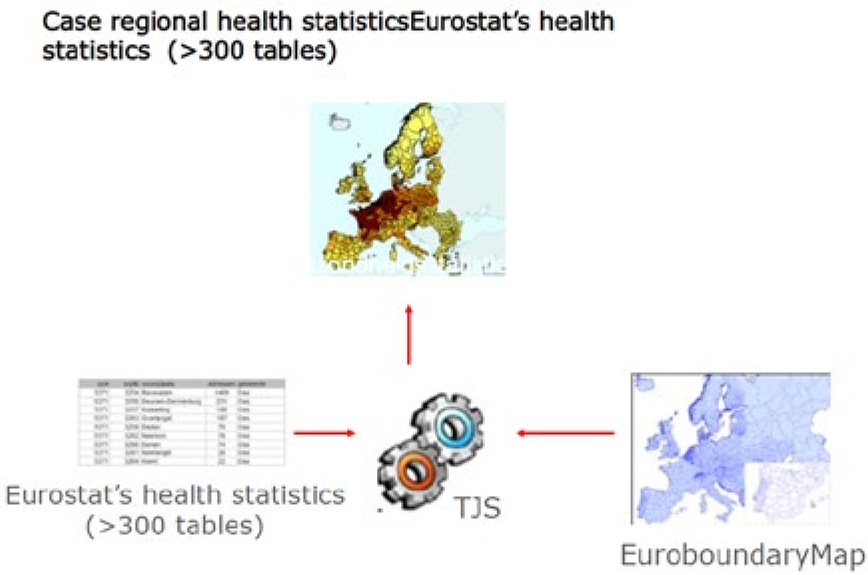


Figure 29. Example of TJS application in the health domain.

Geospatial referencing and data integration need common unique identifiers as geocodes. In addition, spatial aggregation can be performed also by location. This standard spatial join operation in a GIS follows the point-based system. A spatial join is a GIS operation that affixes data from one feature layer's attribute table to another from a spatial perspective. Furthermore, a spatial join operation for the integration of statistical and geospatial data is possible with grid data. It is important that this geospatial/statistical grid is congruent throughout the area that has to be analysed. Different definitions of grids may lead to incompatible analyses. For this reason, statistics in Europe have defined a statistical European grid system within INSPIRE.

A common characteristic and prerequisite for every implementation of data integration, especially the 'point-based' system, is deploying common persistent unique identifiers (PIDs) for statistical data on record unit level as means of geocoding by providing (indirect) spatial reference. Furthermore, PIDs promote management of semantics since PIDs identify the same data objects in different data repositories and vocabulary collections. In many countries, PIDs have been assigned to spatial database objects, especially on European level according to INSPIRE implementing rules on the INSPIRE spatial objects. To promote data linking beyond INSPIRE data and on national level and between different data providers, national guidelines on assigning PIDs and setting up a PID service with a PID (URI) resolver are needed. It is often necessary to link key register data of SDI such as cadastral parcel, address, buildings and dwelling data in order to produce meaningful statistical information. Linking is performed explicitly through identifiers – e.g. NUTS code– or implicitly through geospatial relations – e.g. a building is within a commune. For example, dwellings and population registers will be linked to support policy questions like how many people live in energy efficient buildings. For the use case of routing – e.g. car navigation– the road identifier may be included in the address register. Such linking is activated and further enhanced with shared PIDs that enables geospatial re-aggregations e.g. spatial joins, particularly useful on record unit level, and can be expanded to different time versions by e.g. including time extensions in PID management or even use case defined versions of the PID

RDF is both format and data model to serve as technology for Linked Data. In addition to the method description on the 'Area based data integration', the importance of 'RDF' and 'Linked Data' should be highlighted. Geospatial data, like geographies of areal classifications, and their data models can be transformed in RDF. The same applies to statistical data. Both models are then accessible on the web and can be integrated in a federated query through SPARQL. Statistical data can be made accessible through SPARQL endpoints–e.g. in statistical Extended Knowledge Organization System (XKOS) format–, including the geospatial relations and temporal versions and providing a continuous common semantic structure. Hence, linked data is most applicable for data integration from distributed and heterogeneous data sources provided with appropriate ontologies. Another Data Integration convergence, the Graph-based Linked Data is gradually becoming a mature technology for statistical-geospatial data integration; the data are originally stored in graph structures / databases and not in relational tables. RDF for Linked Data represents an exchange interface and not a storage structure. A storage structure for graphs is established in a graph database e.g. by labelled property graphs. This allows for surpassing differences in data models, user interfaces and distributed data sources. It is providing unrivalled potential for data providers to enrich their information and knowledge supply as well as make their data capture more efficient e.g. through Knowledge Graphs and, step by step, evolving Knowledge Graphs Infrastructure. In a number of NSOs statistical data are already exposed in RDF format. At the European level, the European Data Portal provides SPARQL/RDF endpoints to many European datasets. While rapid development is underway, current standards are not fully taking advantage of linked information technologies.

5.6 Air quality and environmental data integration

The main methods used to measure air quality include:

- Passive measurement:
    - Diffusion tubes: These absorb ambient air samples over time and are then analyzed in a lab
    - Deposit gauges: Large funnels that collect particulates which are then analyzed in a lab
- Active measurement:
    - Automated or semi-automated devices that actively draw in air samples using fans and filters the pollutants
    - These can detect specific pollutants like particulate matter (PM2.5, PM10), gases (ozone, nitrogen dioxide, sulfur dioxide), and others
    - They provide near real-time data that can be used to calculate air quality indexes (AQI)
- Remote monitoring:
    - Satellite-based monitoring using instruments like GOME, MAPS, and sensors on GOES-R and JPSS satellites to measure pollutants like ozone, carbon monoxide, and particulates from space
    - Drones, and other airborne methods can also remotely measure air pollution
- Low cost sensors monitoring:
    - Citizen Science use low cost sensors for air quality, in particular for particulate matter and other variables.
    - Sensors mounted on vehicles like buses, taxis, and cars to collect hyperlocal air quality data through "drive-by" monitoring.

The choice of method depends on factors like cost, accuracy, spatial coverage, and the specific pollutants being measured. Governments and agencies use a combination of these techniques to comprehensively monitor air quality

Some of the key difficulties in integrating air quality data include:

- Spatial coverage gaps: Many regions, especially in low and middle-income countries, have limited ground-based air quality monitoring stations, leading to gaps in spatial coverage. This makes it challenging to get a comprehensive picture of air quality across an entire region or country.
- Heterogeneous data sources: Air quality data comes from a variety of sources, including ground-based monitors, satellite remote sensing, chemical transport models, and mobile sensors
- Integrating these diverse data streams with different measurement techniques, spatial resolutions, and quality assurance protocols is a significant challenge. There is a need for common vocabularies for data quality.
- Data management and interoperability: There are issues around data management, including lack of data descriptions, integrity, transparency, and the development of open-source tools for interpreting and

communicating air quality data from sensor networks. Ensuring interoperability between different air quality data platforms is also crucial.

- Capacity and infrastructure limitations: Many low and middle-income countries lack the institutional capacity, trained personnel, and supply chain viability to develop and maintain sustained air quality monitoring infrastructure. This hinders their ability to collect, analyze, and share air quality data.
- Integrating diverse data streams: Beyond just air quality data, integrating other relevant data sources like traffic, social media, and meteorological information can provide a more holistic understanding of air quality. Developing the frameworks and tools to effectively combine these diverse data streams is an ongoing challenge.

Using the CitiObs project EU project we can illustrate the different steps for air quality integration.

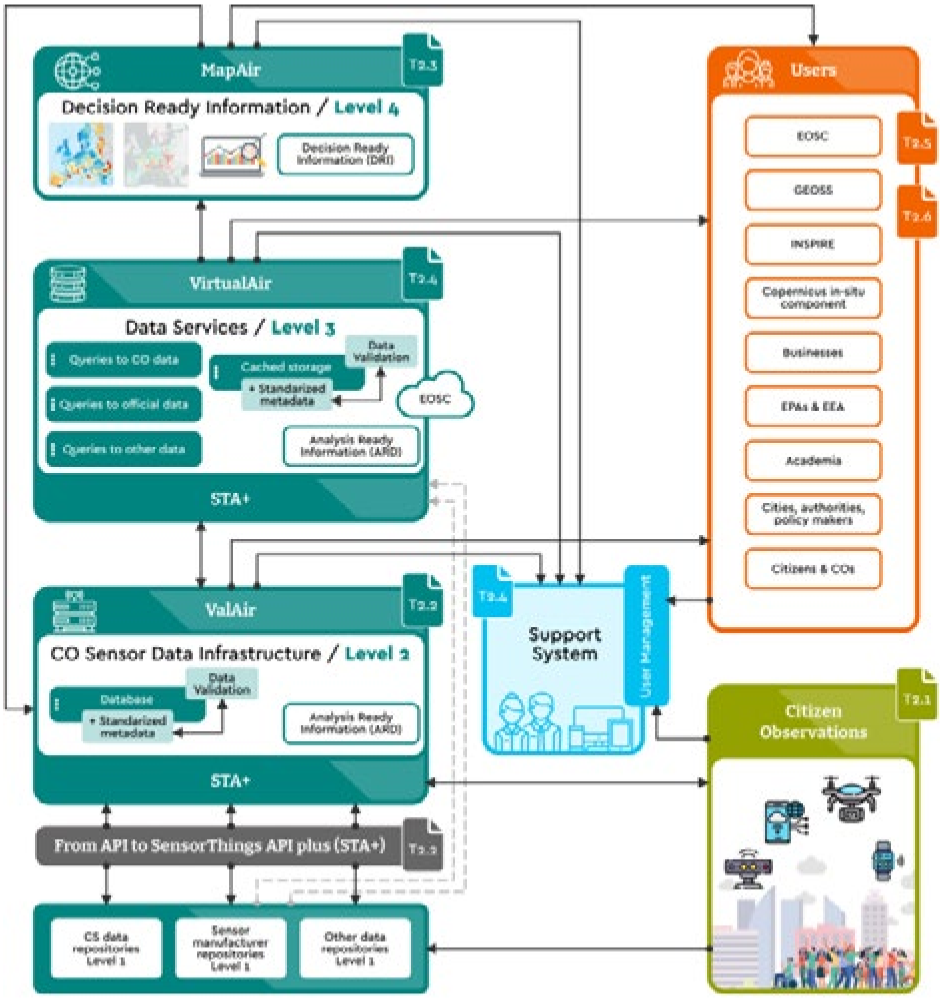


Figure 30.CitiObs architechture.

In the first level, we show several sources of air quality observations generated by sensors. This observations are offered by every individual observatory using the OGC Sensor Things API (STA). In the level 2 a central node does quality checks and generates metadata. Once the information is validated, The level 3 offer a single STA entry point to the validated observations in an integrated way. Finally, Level 4 create map distributions of the different pollutants for decision making. A decision Support System is able to present data from Level 2, 3 and 4.

5.6.1 Sensor Things API

The SensorThings API provides an open and unified way to interconnect Internet of Things (IoT) devices over the Web as well as interfaces to interact with and analyze their observations. Part 1: Sensing was released in 2016 and allowed management and reception of observations or measurements made by IoT sensors. Part 2: Tasking Core, which was released in 2019, provides a mechanism to tell the sensor/actuator what to do.

The foundations of the SensorThings API are the relational connections between entities in the system and the way they are used to model systems in the real world. The entities have a natural relationship which enables any IoT sensing device from any vertical industry to be modelled in the system. An IoT device or system is modelled as a Thing. A Thing has a Location with one or more Datastreams. Each Datastream observes one ObservedProperty with one Sensor of one Thing and has many Observations from the Sensor. Each Observation read by the Sensor observes one particular FeatureOfInterest. Together, these relationships provide a flexible standard way to describe and model any sensing system. It allows SensorThings to be a single data exchange system for heterogeneous devices within any organization.

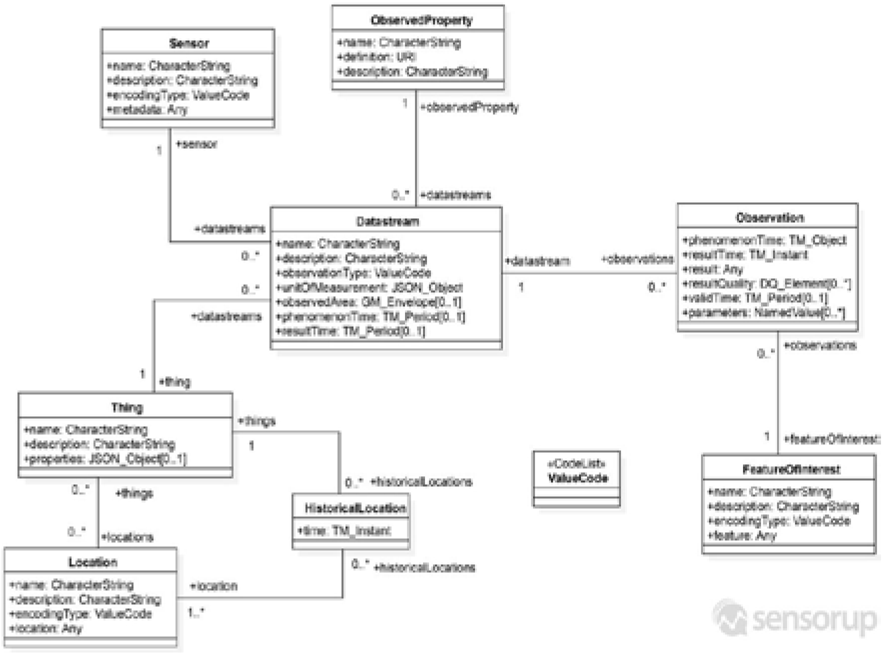


Figure 31. UML Data Model for OGC SensorThings API and defining different entities and their properties, together with the relationship between entities.

One important aspect of the STA is that it provides the definition of variables (property definition in the ObservedProperty entity) and units of measurement (property unitsOfMeasurement in the Datastream entity). This is useful to connect to vocabularies of variables and units of measure.

Another interesting thing of the STA is that a single data model is useful for almost every type of observation. In the following illustration we can see how the STA can be adapter to describe a air temperature of 25 Celsius measured in Barcelona.

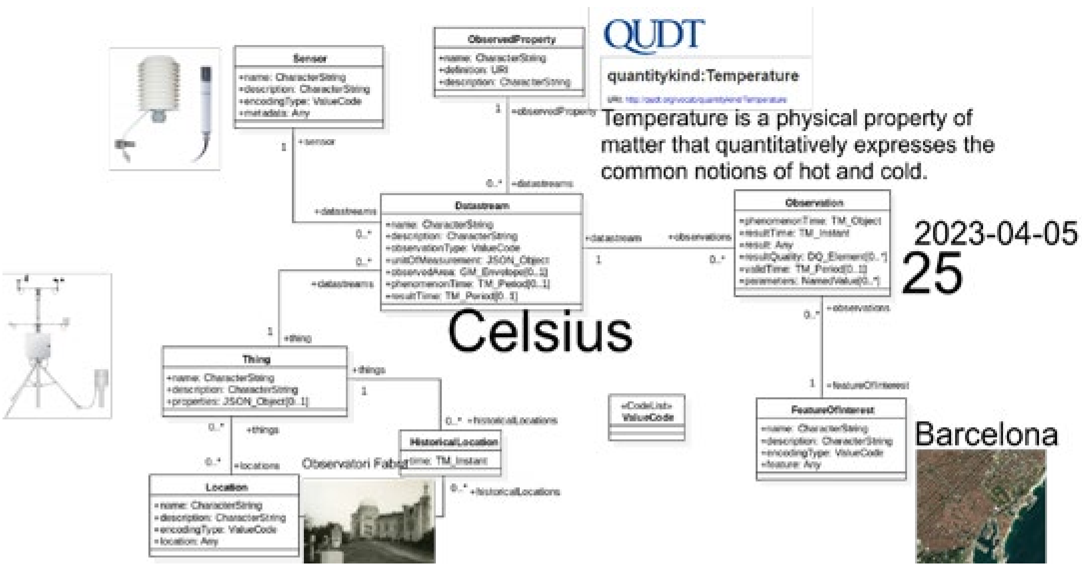


Figure 32. OGC SensorThings API applied to weather observations.

In addition to a flexible data mode, STA provides an API based on OData that can be used to query the service. Assuming that the root of the API is https://citiobs.demo.secure-dimensions.de/staplus/v1.1, the observations associated to the first Datastream can be seen by requesting this: https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Datastreams(1)/Observations

The response will be something like this:


```
{
  "value": [
   {
     "@iot.selfLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(1)",
     "@iot.id": 1,
     "phenomenonTime": "2023-07-24T13:20:56Z",
     "result": "27.25",
     "resultTime": "2023-07-24T13:20:56Z",
     "Datastream@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(1)/Datastream",
     "FeatureOfInterest@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(1)/FeatureOfInterest",
     "Objects@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(1)/
Objects",
     "ObservationGroups@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(1)/ObservationGroups",
     "Subjects@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(1)/
Subjects"
   },
   {
     "@iot.selfLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(9)",
     "@iot.id": 9,
     "phenomenonTime": "2023-07-24T13:20:56Z",
     "result": "27.24",
     "resultTime": "2023-07-24T13:20:56Z",
     "Datastream@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(9)/Datastream",
     "FeatureOfInterest@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(9)/FeatureOfInterest",
     "Objects@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(9)/
Objects",
     "ObservationGroups@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/
Observations(9)/ObservationGroups",
     "Subjects@iot.navigationLink": "https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Observations(9)/
Subjects"
   },
   {…
```


Results contain the values but we do not know what they are about. We can $select the results and phenomenonTime to make the response less verbose and expanding the Datastream we can know more about what they are (we are using $select again to make the response les verbose

https://citiobs.demo.secure-dimensions.de/staplus/v1.1/Datastreams(1)/Observations?$select=result,phenomenonTime&$expand=Datastream($select=name,unitOfMeasurement/name)

The response will be something like this:

```
{
  "value": [
    {
      "result": "27.25",
      "phenomenonTime": "2023-07-24T13:20:56Z",
      "Datastream": {
        "name": "Air Temperature",
        "unitOfMeasurement": {
          "name": "Celsius"
        }
      }
    },
    {
      "result": "27.24",
      "phenomenonTime": "2023-07-24T13:20:56Z",
      "Datastream": {
        "name": "Air Temperature",
        "unitOfMeasurement": {
          "name": "Celsius"
        }
      }
    },
    {...
```

You can learn more about this API by visiting this website: https://developers.sensorup.com/docs

5.6.2 STAplus extension

"STAplus" refers to an extension to the OGC SensorThings API (STA) Part 1: Sensing Version 1.1 standard. The STAplus extension was developed to address requirements from the Citizen Science community, where there are many contributors (citizens) who collectively create observations. The main goal of STAplus is to support a model where observations are owned by different users and can have associated licensing information for reuse. STAplus adds support for expressing ownership (adding Party entity) and licensing (adding License entity) of observations. It allows creating explicit relations between observations (adding Relations entity) and groups of observations (adding ObservationGroup entity) to enable better linked data extraction and semantic querying. The extension is designed to strengthen the "Interoperability" (I) and "Reusability" (R) aspects of the FAIR principles for data management. STAplus extends the STA data model to include new entities and properties related to ownership, licensing, and observation relationships.

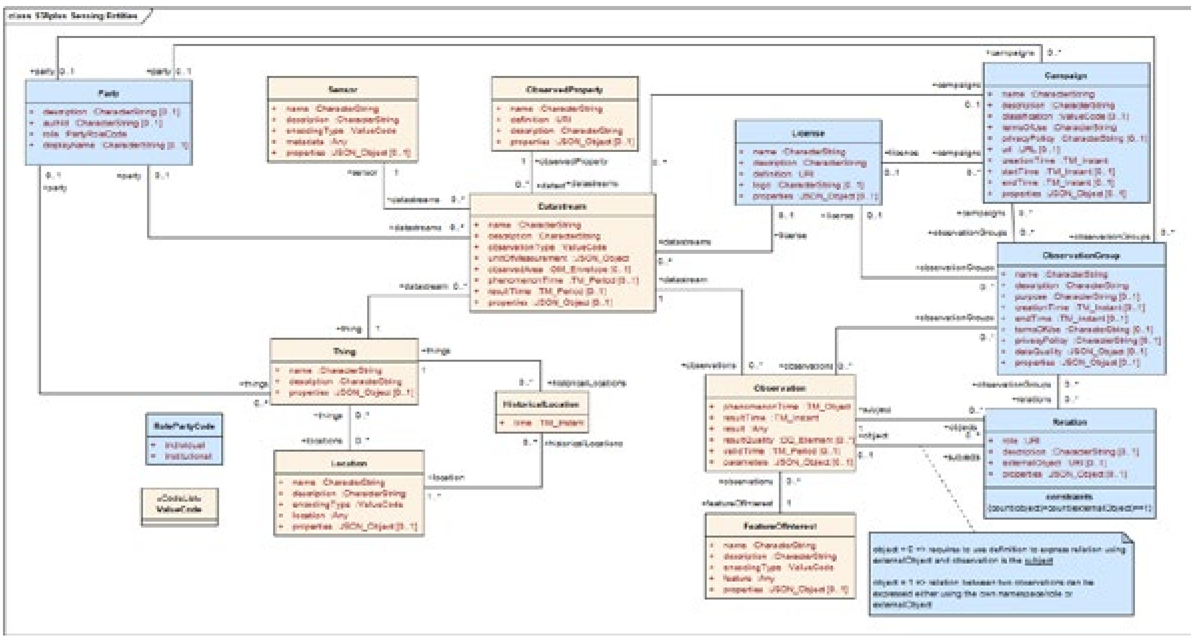


Figure 33. UML Data Model for OGC STAplus

A tool called Tapis has been developed to illustrate how to create queries to STAplus (and SensorThings API). The tutorial of the tool can be found here: https://www.tapis.grumets.cat/help/

5.7 References

Geospatial Metadata Fact Sheet, FGDC 2011. In Internet: https://www.fgdc.gov/resources/factsheets/documents/GeospatialMetadata-July2011.pdf

Data Integration Methods. Analysis of future trends in geospatial data capture, creation, maintenance and management and recommendations for amplified use of good practices UN-GGIM: Europe | Working Group on Data Integration Version 1.0 2021-11-22 In internet: https://www.catastro.meh.es/documentos/internacional/Documento_Integracion_de_datos_UNGGIM_10012022.pdf

Champion M., Ferris C., Newcomer E. and Orchard D. (Eds.): Web Services Architecture, W3C Working Draft November 14, 2002 (W3C, Cambridge 2002), http://www.w3.org/TR/ws-arch/

Peng, Z.R, M.H Tsou (2003). Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Network. Ed. Wiley. 720 pp.

Lucchi, R., Millot, M., & Elfers, C. (2008). Resource oriented architecture and REST. Assessment of impact and advantages on INSPIRE, Ispra: European Communities.

Döllner J. Markkus Jobst, Schmitz P. (2019) Service Oriented Mapping. Lectture Notes on Geoinformations and Cartography, Springer.

W3C: What is Linked Data (W3C, Cambridge 2010) http://www.w3.org/standards/semanticweb/data

Pan, M., Pang, B., Li, C., Wang, D., Min, D., & Zhao, D. (2013, December). Research on Web Mapping of Vector Nautical Charts Based on HTML5. In Information Science and Cloud Computing Companion (ISCC-C), 2013 International Conference on (pp. 611-617). IEEE

Deng M., Di L, Facilitating Data-intensive Research and Education in Earth Science: A Geospatial Web Service Approach. LAP LAMBERT Academic Publishing 2010, 212pp, ISBN-13: 978-3838397146.

OGC Reference Model (ORM) OGC 08-062r7 2011 http://www.opengeospatial.org/standards/orm , http://portal.opengeospatial.org/files/?artifact_id=31112

Whiteside A. and Greenwood J. OGC: OpenGIS Web Service Common Standard 2.0 (Open Geospatial Consortium, Wayland 2010), http://www.opengeospatial.org/standards/is

Gregorio J., Fielding R., Hadley M., Nottingham M. and Orchard D. URI template, Internet Engineering Task Force RFC 6570

Stefanakis E. Web Mapping and Geospatial Web Services: An Introduction. Ed. CreateSpace Independent Publishing Platform.. 2015, 168pp ISBN: 978-1514757550.

Jeff de la Beaujardiere, W. (ed.), OGC Web Map Service Interface 1.3.0, OGC Document # 01-109r1 , 2004-01-20. Available at http://www.opengeospatial.org/standards/is.

Masó J, Pomakis K and Julià N OpenGIS Web Map Tile Service Implementation Standard OGC 07-057r7

Loechel, A. J., & Schmid, S. (2013). Comparison of Different Caching Techniques for High-Performance Web Map Services. IJSDIR, 8, 43-73.

Sample S. T., and Lup E. Tile-Based Geospatial Information Systems. Principles and Practices. Ed.Springer, 2010 ISBN 978-1-4419-7630-7

Lalonde, W. (ed.), Styled Layer Descriptor Implementation Specification 1.0.0, OGC 02-070, August 2002. Available at http://www.opengeospatial.org/standards/is.

Vretanos, P. (ed.), Web Feature Service Implementation Specification 1.0.0, OGC Document #02-058, September 2003. Available from World Wide Web: http://www.opengeospatial.org/standards/is.

Portele C. OpenGIS® Geography Markup Language (GML) Encoding Standard version 3.2.1 OGC 07-036

Davis S. (2007) GIS for Web Developers, Pragmatic Bookshelf; Programers, pp. 254

Baumann P, Hirschorn E and Masó J, OGC Coverage Implementation Schema OGC 09-146r6

Baumann P. (2010) OGC® WCS 2.0 Interface Standard – Core OGC 09-110r3

Zhao P, Di L (2010) Geospatial Web Services: Advances in Information Interoperability, IGI Global

Botts, M., Percivall, G., Reed, C., & Davidson, J. (2008). OGC® sensor web enablement: Overview and high level architecture. In GeoSensor networks (pp. 175-190). Springer, Berlin, Heidelberg

Bröring A. Stasch C. and Echterhoff J. OGC Sensor Observation Service Interface Standard OGC 12-006 http://www.opengis.net/doc/IS/SOS/2.0

Best Practices for Using Sensor Web Technology, 52 North, https://52north.github.io/sensor-web-tutorial/

Alameh, N. (2003). Chaining geographic information web services. IEEE internet Computing, (5), 22-29.

Mueller M. and Pross B. OGC WPS 2.0 Interface Standard, OGC 14-065 http://docs.opengeospatial.org/is/14-065/14-065.html

Nebert D., Whiteside A. and Vretanos P. OpenGIS® Catalogue Services Specification version 2.0.2 OGC 07-006r1

Kresse, W. i K. Fadaie (2004). ISO Standards for Geographic Information. Ed. Springer. 322 pp

Matheus A., Terpstra F. and Masó J. (ed.) OGC Web API Guidelines (2021), available at https://github.com/opengeospatial/OGC-Web-API-Guidelines

Portele C, Vretanos P. and Heazel C. (ed.) OGC API – Features – Part 1: Core, 1.0.0. OGC, Wayland (2019). OGC Document # 17-069r3, 2019-10-14, available at http://docs.opengeospatial.org/is/17-069r3/17-069r3.pdf

Portele C and Vretanos P. (ed.) OGC API – Features – Part 2: Coordinate Reference Systems by Reference, 1.0.0. OGC, Wayland (2019). OGC Document # 18-058, 2020-11-02, available at http://docs.opengeospatial.org/is/18-058/18-058.pdf

LESSON – 6

Partnerships on Spatial Social Challenges

Trainer's Biodata: Dr. Rosemarie Mijlhoff – Portier

Rosemarie guides organizations to initiate multi stakeholder partnerships or help them to break out of a deadlock. For Geonovum she guides the learning processes of multi stakeholder partnerships on developing and embracing a Digital Twin as an instrument for policy and decision-making processes related to the spatial societal challenges the Netherlands faces. Rosemarie holds a PhD on network collaborations, and a MSc. on governance and organizational sciences.

Trainer's Biodata: Lt Gen Girish Kumar, VSM

Lt Gen Girish Kumar, VSM retired as Surveyor General of India (SGI), Survey of India, with almost 39 years of service experience in wide ranging geospatial domain of Surveying, Mapping, Geospatial Policy, GIS applications, Digital data generation, Geospatial Infrastructure, International boundary matters etc. He also served as Additional Director General Military survey (GSGS),Military Survey, Army Headquarters. A Civil Engineering Graduate of Punjab Engineering College (PEC), joined Corps of Engineers of Indian Army in year 1980 and then seconded to SoI in 1982. Undergone various specialised courses in Surveying, Mapping and digital data generation in India and abroad including Production Photogrammetry PG course from ITC, Netherlands. For his exemplary contribution in the geospatial field, got Vishisht Seva Medal (VSM) and Chief of Army Staff Commendation Card during the service. National award for e-Governance in the Gold Category by the Government of India for the year 2013-14 for 3D DSSDI Project of Delhi.

6.1 Introduction

Multi stakeholder partnerships are seen as a holy grail for innovations like the development of digital twins. Both researchers and practitioners say that collaboration processes are often tough and challenging. Professors Huxham & Vangen (2005), famous for their research on collaborative working and inter-organizational partnerships, gave the advice: don't collaborate unless you have to.

In this lesson Rosemarie Mijlhoff shows what types of multi stake holder partnerships exist, and what the challenges are in achieving success as collaborating organizations. She also makes visible and tangible what the hidden dynamics of a collaboration are that participants of partnerships have to be aware of. The benefits of a partnership are also discussed, as well as the self-reflection and competences needed to deal with these hidden dynamics in a way that these act as fuel in support of the partnership rather than an explosive.

During this lesson, the participants will experience how they could bring a multi stakeholder partnership into practice. They will do this by means of the Svamitva project, which Gen Girish Surveyor General of India will introduce during the lesson, and by using the learning methodology Geonovum developed as part of a digital twin innovation project. In this way, the participants gain insight into how they could set up a multi stakeholder partnership and how they could gradually master the competencies needed to deal with the dynamics that occur along the way.

6.2 Theories on multi stakeholder partnerships

The United Nations Sustainable Development Goal 17 which reads, "Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development", recognizes how multi-stakeholder partnerships (MSPs) can help in the achievement of Sustainable Development Goals by mobilizing resources, knowledge, and expertise from all the stakeholders in the society.

MSPs occur when stakeholders feel that there is a need to collaborate, and they can perform more efficiently as a group instead of working independently. Such partnerships include both the State and non-State actors for achieving common objectives.

Why do such Multi Stakeholder Partnerships occur?

Changing Governance: The world has become more globalized. Economic and social issues transcend the national boundaries these days. At the same time people expect the governance to be more decentralized and participatory. MSPs can fill in the gaps left by traditional forms of governance by representing the interests of varied stakeholders.

Human society as Complex Adaptive System: Human societies are complex adaptive systems, which means any change in society happens through combined action of many individuals interconnected in a system and nobody is in full control over anything. The social systems must constantly update to new situations. One way to improve adaptability in such a situation is to improve communication, and MSPs serve as a perfect medium to improve communication between different stakeholders.

Behavioral Aspects: Traditional economic theorists have considered man as a totally rational being solely motivated by wealth and personal gains (Homo economicus). But humans don't behave in such a way. Cooperation and emotions play a large part in our decision making. MSPs foster this spirit of cooperation between different interest groups.

Types of MSPs and their Characteristics:

While there exist no fixed criteria for categorizing Multi-Stakeholder Partnerships, some of the common Multi Stakeholder Partnerships are as following:

- **Sector-Specific Partnerships**: These partnerships focus on a particular industry or sector, such as education or healthcare. They bring together stakeholders from that sector, such as businesses, governments, Civil Society Organizations, to work on common issues in the sector. Example: The Global Alliance for Vaccines and Immunization (GAVI) is a public-private partnership that brings together governments, donors, the private sector, civil society organizations, and research institutions to improve access to vaccines.
- **Territorial Partnerships**: These partnerships focus on a particular geographic area, such as a city, region, or country. They bring together stakeholders from that area to work on issues that affect the entire community.
- **Functional Partnerships:** These partnerships focus on a particular function or activity, such as research and development, standard-setting, or advocacy. They bring together stakeholders who have expertise in that function to work together on a common goal. Example: The World Wide Web Consortium (W3C) is a functional partnership that brings together companies, government agencies, and other organizations to develop standards for the World Wide Web.
- **Multi-Level Partnerships:** These partnerships involve stakeholders from different levels of government, such as local, national, and international. They can be used to address complex challenges that require cooperation across different levels of governance. Example: The Paris Agreement on climate change is a multi-level partnership that brings together countries from around the world to take action on climate change.

The characteristics of MSPs vary depending on the specific partnership, but some common characteristics include:

1. A shared and defined problem situation or opportunity.
2. A commitment from all partners
3. Communication and collaboration among partners.
4. A results-oriented approach.
5. Follows an agreed upon but dynamic process and timeframe.
6. A mechanism for monitoring and evaluation.

6.3 How to form a multi stakeholder partnership

The formation of MSP is a time-consuming process as participating in MSP can tie up limited resources available. There is a need to first address whether MSP is the right choice for a particular issue. Each MSP must go through the following steps:

Step 1- Evaluation: Is MSP the right choice? Measuring the benefits and the pitfalls.

Before forming an MSP there is a need to decide whether an MSP will lead to the intended change. "How is the change going to happen?" and whether MSP is the right choice or not, can be decided by weighing in the benefits and the pitfalls of MSP:

The benefits:

1. Different partners in a multi-stakeholder partnership can bring different resources to the table. While donors & foundations can provide with funding and support, international agencies can bring their technical support and expertise. A market based and commercial approach of Businesses, along with social capital & legitimacy provided by Civil Society Organizations, and regulatory support provided by the Government can help tackle various issues.
2. They can be used whenever there is an alignment of interest between different actors and there is a scope of joint action over any issue.
3. They serve as a medium of knowledge and skill exchange between various stakeholders.
4. They can operate at any level from International to local level.
5. They can promote the spirit of diversity and inclusivity.
6. Results will have broader ownership.
7. Learning and collaboration increase chances of systemic change.

The pitfalls:

1. Partnering across different sectors is difficult. It requires bringing together of people with different approaches and interests.
2. There is always a significant financial and resource cost involved in the formation of such large-scale partnerships.
3. Can only work when there is sufficient representation from stakeholders.
4. They will not often deliver short-term successes.
5. Collaboration processes are often tough, and they may not always lead to the intended results. The hidden dynamics between various stakeholders might come in between achieving success.

Step 2- Designing the process.

The design process of MSP involves:

- 1) Understanding the context and hidden dynamics between different stakeholders.
- 2) Developing a change strategy guided by the MSP principles.
- 3) Deciding on the methodologies and tools that will be used to engage stakeholders in relationship building, analysis, planning, and collective action.
- 4) Designing must also include some level of adaptive planning and the collaborative action proposed for different stakeholders.

Step 3- Implementing.

MSPs are implemented by cooperation of various stakeholders. Working collaboratively to execute plans, address challenges, and monitor progress should be some common goals for all stakeholders. Regularly assessing the partnership's effectiveness and make adjustments as needed in is an important step while implementing any Multi- Stakeholder Partnership.

Competencies needed to deal with hidden dynamics in a Multi-Stakeholder Partnership:

Effective multi-stakeholder partnerships often face hidden dynamics – unspoken agendas, power imbalances, and underlying tensions. These issues can be solved by:

Facilitation of Discussion: Guiding open and inclusive discussions that surface hidden issues while maintaining a focus on shared goals.

Conflict Resolution: De-escalating tensions, fostering empathy, adaptable agility and finding win-win solutions are essential skills.

Transparency and Trust Building: Creating a safe space for open communication and fostering trust among stakeholders is paramount.

Dealing with Asymmetries: In a Multi Stakeholder Partnership, a lot of asymmetries exist between different players. Dealing with these asymmetries is necessary for success of any MSP.

Managing conflicts, building trust, dealing with power asymmetries are paramount to success of any Multi-Stakeholder Partnership.

Step 4- Reflective Monitoring.

The Reflective Monitoring phase lies at the centre of the MSP process model, embedded in the other phases. Reflective monitoring is an integral part of adaptive management and is critical for adjusting to the needs of various stakeholders.

6.4 Case Study: Partnership and Collaboration for SVAMITVA Project: Partnership, Type of Funding and Business Models involved.

6.4.1 Scheme Overview:

The SVAMITVA (Survey of Villages Abadi and Mapping with Improvised Technology in Village Area) scheme is a reformative step towards establishment of clear ownership of property in rural inhabited ("Abadi") areas, by mapping of land parcels using drone technology and providing 'Record of Rights' to village household owners with issuance of Property cards to the property owners. This would enable village household owners, to use their homes as a financial asset for taking loans and other financial benefits and provides the record of rights for inhabited rural areas in village thereby increasing tax collection.

6.4.2 A Collaborative Approach:

The scheme facilitates collaboration at two levels:

1. Government to government Collaborations
2. Collaborations with Private Sector

Government to Government Collaborations: The scheme is implemented with the collaborative efforts of the Ministry of Panchayati Raj (MoPR), State Revenue Department, State Panchayati Raj Department and Survey of India (SoI). The working together of various Government Agencies establishes Government to Government Collaboration between the Union Government, the State Governments and the Panchayati Raj institutions at the grassroot level of governance.

Stakeholder Partnership

The SVAMITVA project leverages a unique partnership model:

- **Central Government:** The Ministry of Panchayati Raj (MoPR) spearheads the project, providing overall guidance and funding.
- **State Governments through State Revenue and Panchayat Departments:** States identify target villages, mobilize local communities, and contribute manpower and resources. The preparation and distribution of property card is also the responsibility of State Governments.
- **Survey of India (SoI):** SoI conducts drone surveys and prepares digital maps of Abadi areas with details of each property duly demarcated.
- **State Panchayati Raj Institutions (SPRI):** SPRIs coordinate with villagers, collect property details, and maintain land records.
- **Banks:** Banks play a crucial role in leveraging land records for loan applications and financial inclusion.
- **Non-Governmental Organizations (NGOs):** NGOs can support community outreach and awareness campaigns.

Partnership Model

The inclusion of every level of governance and the division of responsibilities among different branches of government, from MoPR at the Union Level to the State Revenue Departments and local level Panchayati Raj Institutions, establishes a unique partnership rarely seen in any Government Scheme.

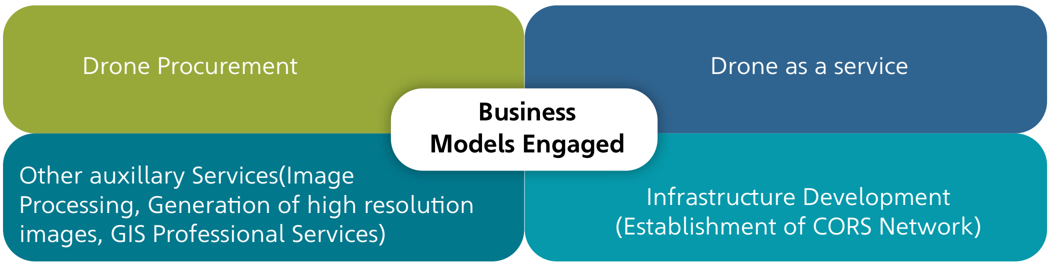
Collaboration with the Private Sector: Scope for private sector participation

- To involve private players, the SOI invited RFE (Request for Empanelment) from reputed Firms for empanelment for "Drone as a service (DAAS)" and also for preparing digital maps under SVAMITVA.
- The service of Private Sector was sought for Image Capturing using Drone including provision of Ground Control, Delivery of quality Ortho Rectified Image (ORI) and DEM and digital maps.
- Along with this, drone procurement, data and image processing services and the involvement of GIS professionals would create collaboration between SOI and private players.
- SVAMITVA covers 6.62 lakh inhabited villages which is just a small portion of the land mass of the country. The remaining part of the country shall also be mapped with the same high-resolution images. The scope of drone mapping is going to exponentially increase in future thereby creating further opportunities for the private sector.

Business Model Involved

Different States, Government Bodies and SoI, have employed different business models & engaged private players in different ways when it comes to implementation of SVAMITVA. Here are some examples:

- **Maharashtra**- For Gaothans (village settlement areas), Maharashtra has sought service of private players for Post Drone Flying work and providing GIS services for Village mapping.
- **SoI**- SoI has involved private players for procurement of Professional Survey Grade UAV/Drone for Abadi areas of the villages. The SoI has also invited bids for using drones as a service. Besides that, SoI has also outsourced the work of establishment of CORS network in the country.
- **Clement Town Cantonment Board**- Has involved private players for the Drone Survey work on the aspects of image capturing, generation of High Resolution, accurate and precise ORI, feature extraction and entry of attributes based upon the workflow for SVAMITAVA scheme.
- **Andhra Pradesh**- Has hired drone/aerial photography as an end-to-end solution to large scale mapping of the entire State.



6.4.3 Funding Type:

The SVAMITVA project primarily relies on the **Union Government** for funding. MoPR allocates funds to states and Survey of India for project implementation. Most of the states have also contributed additional resources based on their priorities.

The component wise fund recipients for the Scheme for implementing different objectives are as following:

1. **Survey of India**: Establishment of CORS network; Large scale mapping using drones, data processing, ground truthing of topographical features
2. **State Revenue Department**: IEC Activities, State Program Management Unit (SPMU) activities, Documentation and Workshops etc.
3. **NIC through NICSI:** National Program Management Unit (NPMU) Activities, Enhancement of Spatial Planning Application, Online Monitoring systems, Central Infrastructure: Hardware & Software.

6.4.4 Benefits of Collaboration:

- As of 1st April 2024, the collaboration of different governments under Svamitva has resulted in drone survey of more than 290,000 villages, digitization of more than 90 million property parcels, and monumentation of more than 1000 CORS.
- The Scheme has given a major push to the procurement of drones from MSMEs and startups. As a result, it has significantly improved MSMEs market access, marketing support, knowledge sharing and participation.
- State authorities now have a record of right of land, which can be relied on by banks while disbursing loans, for settling land disputes, and for increasing tax collection. The expertise of SoI and support from MoPR has significantly helped the State Governments for land record digitization.
- The Scheme has made it possible to formulate better-quality Gram Panchayat Development Plan (GPDP) by making use of GIS maps. In a way the scheme has been beneficial for every level of government.

6.4.5 Challenges during the project:

- There are variations in laws governing the land record-keeping practice across different states. So, implementation may not be the same across different states.
- There are questions regarding the legal validity of property cards generated in the scheme. Though, States have made provision to ensure the legal validity of the property cards in the scheme.

6.4.6 Factors that encouraged partnership in SVAMITVA:

Factor	Reason
Aligning the goals and expectations of stakeholders	Many rural landowners lack documented proof of ownership of their lands. A record of right is in best interest of all the stakeholders.
Establishing a governance structure	A Centrally funded Scheme with well defined roles for SoI, PRIs, State Revenue departments encouraged further partnership.
Identifying and prioritizing business opportunities and challenges	Involving drone companies further encouraged the partnership with private sector.
Increased Transparency and reporting	The monitoring of the progress made in Svamitva web portal, helps in smooth implementation of the scheme.

6.4.7 GKI Enhancement under SVAMITVA:

The partnership and collaboration opportunities created between different government bodies from Panchayati Raj Institutes to State Revenue Departments, and SoI would lead to Geospatial Knowledge enhancement. SVAMITVA exemplifies GKI's goal of transforming raw geospatial data into actionable knowledge. The project utilizes this knowledge to generate Record of Rights for rural inhabited property owners. The collaboration between SoI, State Authorities, geospatial and drone companies would further lead to transition from just data acquisition to real world application and knowledge enhancement.

6.5 Bringing a multi-stakeholder partnership into practice

Bringing a multi-stakeholder partnership into practice involves periodic reflecting and adjusting. The process involves creating a value case for all the stakeholders involved, developing standards for partnership and a need to understand how different parties would collaborate in a Multi Stakeholder Partnership.

The SVAMITVA case clearly highlights how collaboration between different players like Union government, State Governments, SoI and private drone companies, helped in providing "Record of Rights" to village land holders and fostered cooperation between different branches of government.

The long-term success of any governance initiative can be achieved if solutions are jointly initiated. For this reason, MSPs are increasingly being encouraged by development policy promoters. From initiatives like Svamitva and Digital twin project of Geonovum, it can be inferred that some general steps in bringing Multi Stakeholder Partnerships into practice should include:

- Identifying the issue
- Stakeholder Mapping
- Setting Common Goals
- Developing Governance Structure
- Building Trust and Communication
- Resource Mobilization
- Implementing Action Plan
- Monitoring and Evaluating Progress
- Sustaining and scaling impact

The success of any such partnership also depends largely on the players involved and the social capital present in the society. If people, processes, and practices all work in cooperation, such Multi Stakeholder Partnerships become a success story.

6.6 References

4. Huxham, C., & Vangen, S. (2005). The theory and practice of collaborative advantage. Routledge.
5. Multi-stakeholder partnerships | Department of Economic and Social Affairs. (n.d.). https://sdgs.un.org/topics/multi-stakeholder-partnerships
6. Mikheyev, A. (2005). Multi-stakeholder partnerships: definition, principles, typology, and partnering

process. In Conference Proceedings: UNESCO between Two Phases of the World Summit on the Information Society (pp. 299-310).

7. An Introduction to multi-stakeholder partnership (2016). The Partnering Initiative. https://thepartneringinitiative.org/wp-content/uploads/2023/01/Introduction-to-MSPs-Briefing-paper.pdf
8. Brouwer, H., & Woodhill, A. J. (2016). The MSP guide: How to design and facilitate multi-stakeholder partnerships.
9. Multi-Stakeholder Partnerships (MAP) – Collective Leadership Institute › Collective Leadership Institute. (n.d.). https://www.collectiveleadership.de/blog/article/map/
10. SVAMITVA web portal. https://svamitva.nic.in/svamitva/
11. Svamitva. (n.d.). https://www.surveyofindia.gov.in/pages/svamitva
12. Framework for implementation of Svamitva Scheme (2021). https://svamitva.nic.in/DownloadPDF/Svamitva_Scheme_Framework(2021-2025).pdf

LESSON – 7

Leveraging Latest Data and GeoAI Technologies for National Mapping

Trainer's Biodata: Dr. Kumar Navulur

Dr. Navulur is founder and CEO of GeoCognition. He received his PhD from Purdue University and was recognized as 2019 Distinguished Purdue Alumni for his contribution to the remote sensing industry. He has over 30 years of experience in the Geospatial industry. Dr. Navulur is the author of the book "Multi-Spectral Image Analysis Using Object Oriented Paradigm" that is used in curriculums in 20+ universities globally. Dr. Navulur is author of several publications and articles in industry magazines. He has several patents in the field of remote sensing and his latest focus includes use of gaming technology and GeoAI/ML technologies to improve geospatial workflows.

7.1 Introduction

One of the key aspects of GKI is the foundational data. Recent years have seen an evolution of New Space technologies that are capturing the planet in multiple modalities (EO, SAR, LiDAR, RF, etc.) and in multiple dimensions (spectral, spatial, and temporal) that can help nations with successful implementation of GKI. New Space developments include technology advances in the field of rocket launches, miniaturization of payloads and sensors resulting in reduced costs of satellites, ISL's (Inter Satellite Links), satellite on-board processing technologies, increasing network of ground stations, and others. Also, new government space and data policies are allowing commercial companies to advance remote sensing technologies. Aerial technologies have also advanced in the last few years with commercial companies collecting imagery at ultra-high spatial resolutions (<-15cm) as well as LiDAR data at high densities (10-20 points per sq m) for creation of DEM's (Digital Elevation Models) as well as 3D mesh models. Drone technology has also rapidly evolved in recent years allowing for ultra-high resolution capture of imagery and 3D data for local areas and their ability to rapidly mobilize allows for frequent map updates. Rapid advances in autonomous driving vehicles resulted in HD mapping technology from terrestrial vehicles that includes street views and detailed 3D surface maps to capture road furniture such as traffic signs, stop signs, lane markings, etc. Public agencies such as NASA and ESA have continued their scientific satellites missions, Landsat and Copernicus missions respectively, providing global imagery datasets at medium spatial resolution. Several NGO's (Non-Governmental Organizations) such as Open Street Map (OSM), Google.Org, and others have developed global map/vector datasets such as building footprints, road networks, etc. that can be leveraged for national mapping. National Mapping agencies can leverage multiple datasets available in the commercial industry as well as from public sector agencies to execute their mapping missions.

The last few years have also seen the increasing adoption of AI/ML technologies for variety applications including GenAI (Generative AI), LLMs (Large Language Models), Deep Learning (DL) and Machine Learning

(ML) models. The invention of Transformers by Google coupled with increased computational infrastructure has drastically increased the adoption of AI/ML technologies in our daily lives. Further, improvements in the labeling technologies/platforms as well as segmentation technologies further enhanced the use of AI/ML technology. Geospatial industry saw an increased adoption of AI/ML (GeoAI) for object and feature extraction, change detection, and other mapping applications. GeoAI is poised to become integral to geospatial community and national mapping agencies should embrace GeoAI for national mapping and maintenance in 2D as well as 3D.

Geospatial datasets such as imagery, elevation models, road vectors and others can be leveraged for a variety of applications beyond traditional mapping. EO imagery can be used for a variety of missions including agriculture, environment, transportation, natural resources management, atmospheric monitoring, census, etc. One of the key applications of EO imagery is satellite derived shallow water bathymetry (SDSB) which allows for mapping and updating shallow water bathymetry in open oceans as well as for inland waterways. DEM's play a critical role for flood plain/watershed modeling, transportation, insurance, topographic maps, and others. 3D texture models are becoming important for various missions including digital twins for smart cities, modeling and simulation, autonomous driving, and other applications. A coordinated geospatial data acquisition mission will benefit national mapping agencies in efficiently leveraging various geospatial datasets for multi-purpose missions.

7.2 Various datasets available for national map updates

There are several datasets available in public as well as private sector for national mapping. The main datasets used for mapping include Electro-Optical (EO) Imagery, Synthetic Aperture RADAR (SAR), Digital Elevation Models including Digital Surface Models (DSM) and Digital Terrain Models (DTM), LiDAR and other 3D point clouds, 3D textured Models, Bathymetry, Ground Control Points (GCP's), Image Control Points, and various Vector Datasets including roads, buildings, coastlines, field/parcel Boundaries, and others.

7.2.1 EO Imagery

EO imagery is one of the main data sources used for mapping missions by the National mapping agencies. EO imagery sources include satellites, planes, drones and terrestrial sensors. Key dimensions for selection of EO imagery include: Spatial resolution (pixel resolution/ground sampling distance (GSD)) that defines the finest object that can be extracted from an image, Spectral resolution (number of spectral bands, wave lengths of spectrum, width of spectral bands to capture spectral signature of interest), and temporal resolution (revisit frequency over area of interest). Spatial resolution can vary from 5cm for applications such as planimetric mapping to 3meter-15meter GSD for nationwide change detection. Standardized schema such as NIIRS(1) (National Imagery Interpretation Scale) describe the objects that can be detected at multiple pixel resolutions and help agencies to select the right spatial resolution for a specific mission. Also, it is important to keep in mind that image quality for the same spatial resolution can vary from one data source to another based on image quality parameters such as Signal to Noise Ratio (SNR), radiometric resolution (preferred >10 bits of data per pixel), and others that influence the overall usability of EO images. National mapping agencies should also be aware of techniques such as pan sharpening, super resolution, and resampling that simulate a spatial resolution that is much finer than the native GSD in selecting an imagery source. Spectral resolution is an indicator of the number of spectral bands capture by the EO sensor. Typical sensors used for mapping missions include a minimum of 4 bands spanning across Visible and Near Infra-Red (VNIR) part of Electro Magnetic (EM) spectrum of sunlight. Depending on the number of bands, EO imagery can classified as Multi-Spectral (MS) bands (2-10 bands), Super Spectral (SS) (10-20 bands), and Hyper-Spectral (HS) (>20 bands). Most of the AI/ML applications use multi-spectral visible bands (Red, Green, Blue (RGB)) for object detection while applications such as landuse/landcover, agriculture, etc. need additional band such as a NIR band, that contain unique spectral information for feature extraction. There are few sensors in the market that collect imagery in the Short Wave Infra-Red (SWIR) part of the EM spectrum. SWIR imagery can be used for specialized applications such mineral mapping, methane detection in the atmosphere, soil and canopy moisture detection, and others. Mid Wave Infra-Red (MWIR) and Long Wave Infra-Red data (LWIR) carry information about thermal responses of land features and is an emerging field that can be leveraged for unique missions including measuring energy efficiency of buildings, monitoring emissions in Oil & Gas industry, and other applications. Table 7.1 shows a partial list of companies that provide global satellite imagery that can be used for national mapping:

Source/Data type/EM Spectrum/GSD	Public/Private
Landsat 8 (MS) VNIR, SWIR, MWIR	Public (NASA)
Sentinel 5 (MS),	Public (ESA)
Maxar (SS), VNIR-SWIR	Private (USA)
Airbus (MS), VNIR	Private (Europe)
Planet (MS), VNIR	Private (USA)
Satellogic (MS), VNIR	Private (Argentina)
BlackSky (MS), VNIR	Private (USA)
SIIS (MS), VNIR	Private (S. Korea)
ISI (MS), VNIR	Private (Israel)
Superview (MS), VNIR	Private (China)
Jilin (MS), VNIR	Private (China)
Satlantis (MS), VNIR	Private (Europe)
SatelliteVu (MS), MWIR	Private (UK)
Albedo (MS), VNIR	Private (USA)
Pixxel (HS), VNIR	Private (India)
Orbital Sidekick (HS), VNIR, SWIR	Private (USA)
Earth Daily Analytics (VNIR, SWIR, MWIR)	Private (Canada)
Axelspace (MS), VNIR	Private (Japan)

Table 7.1 Source of Satellite EO imagery

There are several commercial companies that provide aerial and drone imagery and few companies such as NearMap and Aerometrix are collecting imagery globally. There are select few companies including NCTech that are offering street-view terrestrial imagery as an alternative to HD datasets from companies such Google, Apple, HERE, and others.

National Mapping agencies can also take advantage of COTS (Commercial Off The Shelf) global/national/regional image mosaic products offered by companies such as Maxar, Airbus, NearMap and others. These datasets are typically built on global specification (e.g. UTM WGS84) and agencies can leverage these COTS products to rapidly create imagery basemaps of their nation.

Commercial companies offer various business models to access their imagery and products as well as offer sovereign access to satellites to nations. Agencies can acquire new images for their areas of interest by tasking the satellites as well as access older data from the image archives, with latter being relatively cheaper option compared to tasking. Several of the commercial companies have built platforms and API's (Application Programming Interface) to provide access to new tasked imagery, archive imagery as well as COTS products. For scientific applications, some of the companies are offering Analytics Ready Data (ARD or Data Cube) that is spectrally, spatially, and temporally corrected imagery, from both public and private sources of imagery. Further, there are several data aggregators such as EUSI, UP42, SkyWatch, SkyFi, Appollo Mapping and others that provide access to multiple sources of EO imagery using one platform.

Positional accuracy of the EO imagery is an important factor to consider that influences the map scale of national maps. Satellite image accuracy is dependent on pointing accuracy of the satellites as well as spatial errors propagated through various image processing steps. Ancillary data such as Digital Terrain Model (DTM), and Ground Control Points (GCP's) play a key role in the final accuracy of processed images. Further, positional accuracy of image mosaic products includes error propagation from bundle adjustment/

aerial triangulation of tie points from multiple images. National agencies should require demonstrated positional accuracies of EO Imagery and associated products before using them for national map creation and maintenance.

7.2.2 Synthetic Aperture RADAR (SAR)

SAR imagery with it's ability to see through clouds and day/night and all weather imaging can be useful for national mapping agencies for several missions including geology, maritime, surface deformation, infrastructure maintenance, disaster response, and others. Various dimensions of SAR imagery to consider for mapping include: Spatial Resolution, Wavelength (X, C, L), Polarization, and Imaging Modes (Spot, Strip, Scan). SAR satellites acquire imagery by scanning the ground in two dimensions and the associated spatial resolution of SAR data is defined by impulse response (IPR). The IPR is a two-dimensional entity that is characterized by the range-dimension width (the width of the IPR in the ranging dimension) and the cross-range (or azimuth) dimension width. An image is built up from the reflected signals in both dimensions. By default, SAR sensor resolution is defined by the slant-range plane. Varying microwave lengths of SAR are suitable for different mapping missions. L, C and X- bands are the most widely employed in SAR instruments and table 7.2 summarizes the missions that can be supported by these bands:

Band	Mission
L (15-30cm)	Geophysical monitoring, Biomass and vegetation mapping, InSAR
C (3.8-7.5cm)	Global mapping, Change detection, Monitoring of areas with low to moderate vegetation; Ice, ocean, maritime navigation
X (2.4-3.8cm)	Urban monitoring

Table 7.2 SAR Applications

Most SAR systems provide dual and quad-polarized images, essentially giving multiple images of the same scene. Quad-polarized SAR, also referred to as Polarimetric SAR (PolSAR), captures diverse structural and texture information and allows the recognition of different scattering mechanisms. The specific frequency, look angle, polarization, and illuminated area of a SAR dataset determine which applications the dataset is appropriate for.

Several commercial companies offer a range of image capture modes that define the spatial resolution and the area captured in a SAR scan. SPOT modes typically offer the highest resolution with relatively small area (5kmx5km to 10km x 10km at sub 1 meter resolution), Strip mode offers 1-2 meter GSD and larger footprint than spot model, and Scan model at >2m GSD with large areas ranging in thousands of sq km.

There are several public and private sources of SAR imagery and Table 7.3 provides a partial list of SAR sources:

Company/Region	Public/Private
Sentinel 1/Europe	Public
ALOS/Japan	Public
Airbus (Tandem-X)/Europe	Private
MDA (Radarsat-2)/Canada	Private
eGeos (COSMOS-SkyMed)/Europe	Private
ICEYE/Europe	Private
Capella Space/USA	Private
Umbra/USA	Private
Synspective/Japan	Private

Table 7.3 Sources for SAR imagery

Commercial SAR satellite companies offer various business models for tasking SAR satellites as well as accessing their image archives. Similar to EO companies, customers can access tasking and imagery via platform and associated API's. There are few aerial companies providing SAR imaging services that can be also leveraged by national mapping agencies.

7.2.3 Elevation Data

Elevation is one of the key foundational data for GKI. Elevation data types include Digital Terrain Models (DTM), Digital Surface Model (DSM), 3D Texture models, and Point Clouds. Sources of elevation data can range from sensors in space, aerial as well as terrestrial platforms. Key dimensions for the selection of DEM (DTM and DSM) data source include spatial resolution of gridded data or density for point clouds for LiDAR and associated positional accuracy in X,Y,Z dimensions.

SRTM (Shuttle RADAR Topographic Mission) global elevation datasets, at 90 meter or 30 meter resolution, has been the commonly used elevation data across the globe. National elevation programs have been in place in several countries that have collected DTM data at 10 meter spatial resolution and in the last few years, programs such as 3DEP (3D Elevation Program) in the United States of America (USA) are capturing the entire nation at 1 meter postings. These programs are primarily using aerial LiDAR technology with satellite and IfSAR (Interferometric SAR) based elevation models complementing the LiDAR datasets. Metropolitan areas in select countries across the globe have been mapped at much higher spatial resolution/ postings using aerial stereo photogrammetric as well as LiDAR technologies. Aerial oblique photography has been one of the commonly used technologies for creating photo-realistic 3D textured models of cities. Terrestrial Street-View photogrammetry for HD mapping for autonomous vehicles has also been used to create 3D data. There are several efforts, research and commercial, to combine the terrestrial and aerial 3D datasets to create a cohesive 3D dataset with photo realistic representation of an area. Table 7.4 details some of the global sources of DEM data.

Elevation Source/Product/Posting/Technology	Public Private
SRTM (NASA)/DTM/30m or 90m (SAR)	Public
Airbus/DTM/10 m (SAR)	Private
Maxar/DSM-DTM-3D/<=1 m (EO)	Private
NTT Data/DTM/3 m (EO)	Private

Table 7.4 Sources of DEM data

There are few companies that are providing global satellite derived shallow water bathymetry data including EOMAP and TCarta. Other technologies used for high accuracy bathymetry maps include SONAR and LiDAR techniques and there are several companies offering services for maritime mapping of shallow waters and littoral zones.

7.2.4 Ground Control Points (GCP's)

GCP's are an important aspect of ancillary data for enhancing and validating the positional accuracy of 2D images as well as 3D data. Surveying technologies such as RTK (Real Time Kinematic) GPS provide centimeter-level accuracy GCP's in real-time. Traditionally GCP's were collected on the ground for precisely geolocating 2D datasets, they can also be leveraged to validate the Z dimension of 3D datasets. Photo Identifiable Features (PIF) of GCP's can be existing known points of interest and or some nations have created a nation-wide network of PIF's, similar to surveying monuments, that can serve multiple missions including georeferencing, cadaster creation and maintenance, transportation planning, and others. There are few companies such as CompassData from the USA that offer global GCP's from an archive.

National agencies requiring GCP's only for georeferencing of imagery and datasets can use other alternatives to RTK GPS surveying, which can be resource intensive and costly. There are GPS technologies available in the market that can achieve sub 50 cm accuracy at much lower costs than RTK GPS. Similarly, image chips from high accuracy imagery sources, both 2D and 3D, can also serve as sources for GCP's.

7.2.5 Open Source Datasets

There are several open source vector datasets that national agencies can leverage for national maps. Open Street Map[3] (OSM) is a free open source vector dataset that has global coverage of roads, buildings, and other features and is updated by millions of volunteers daily. Humanitarian Open Street Map (HOT)[4] is a non-profit organization that supports disaster response across the globe and produces vector datasets that agencies can use to respond to disasters. Overture Foundation[5], founded in 2022 under the Joint Development Foundation, is dedicated to the development of reliable, easy-to-use, and interoperable open map data that powers current and next-generation map products. Microsoft has released millions of building footprint data[6] it has derived from it's global aerial imagery missions and these datasets are available using Bing Maps API and have been also integrated into OSM. There are few entities that have created global population density estimates including ORNL[7] (Oak Ridge National Laboratory), WorldPop[8], and Columbia University[9], NY USA that could be of help for national mapping agencies.

There are several public as well as private sources now addressing climate change and are proving global datasets for trace gases and atmospheric pollutants. ESA, NASA, JAXA, and few other public agencies have scientific missions that provide information about green house gases such as methane, nitrous oxide, formaldehyde, and others using satellites in LEO (Low Earth Orbit) and GEO (Geostationary Earth Orbit). Various NGO's such as EDF (Environmental Defense Fund), Carbon Mapper, and others are now providing/planning to provide free access to methane and carbon dioxide data, at global scale. Further, several private entities are now launching satellites for atmospheric mapping as well.

7.3 Use of technology for map maintenance and update

7.3.1 AI/ML Overview

AI/ML technologies are poised to transform map making and National agencies should prepare current and future workforce to adopt these technologies for map making and maintenance. AI/ML can be considered as a discipline such as Physics with various technologies such as Machine Learning, Deep Learning/Neural Networks, Generative AI, and Large Language Models (LLMs) as part of larger AI/ML framework (Figure 7.1).

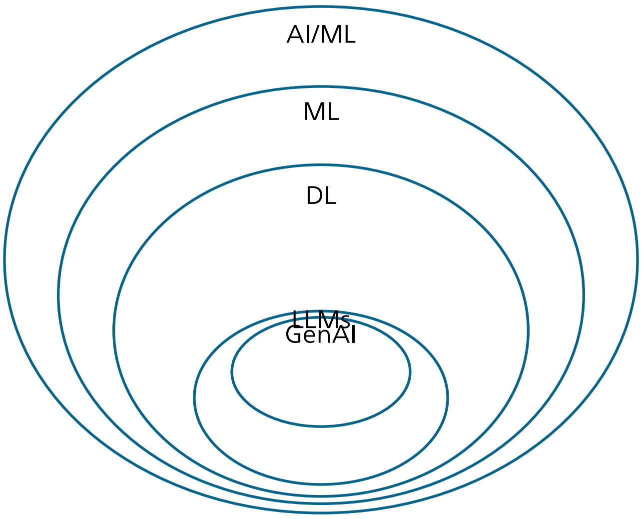


Figure 7.1 Overview of AI/ML technologies

AI/ML is a rapidly evolving technology that is transforming our way of life. While techniques such as ML/DL have been around for decades, it was the invention of transformers by Google engineers that resulted in the rapid evolution of the technology in the last few years. Computing technology from companies such as NVIDIA, AMD, Intel, Google, Alibaba and others played a key role in the development of large AI/ML models. Multi-billion-dollar AI/ML investments from Microsoft, Google, Meta OpenAI, and other technology enterprises is contributing to increased technological advances in this field.

Similar to the data value chain, the ML/DL value chain is comprised of different levels (Figure 7.2)

Figure 7.2 Value chain of ML/DL models

Most of the applications such as object detection, feature extraction, and image classification all fall into the descriptive category. Change detection between two or more temporal sequence of imagery or geospatial datasets also falls into the descriptive category. Figure 7.3 shows an example of detecting number of planes in an image.

Figure 7.3 Plane detection from an aerial image using DL, a descriptive phase of ML/DL value chain

There are few services companies that provide diagnostic service but they typically leverage human analysts to diagnose the information derived from imagery/geospatial sources. Diagnostic phase typically requires the use of multiple data sources (multi-INT) to understand what is happening in an image. An example of diagnostic step can be understanding the agriculture output of a nation and correlating excess, or shortage of food supplies based on weather conditions. Patterns of life, both nature and human, can be predictable most of the time and cell phone companies are able to predict daily behaviors of people and recommend traffic situation along a route routinely taken by an individual. A GeoAI example of predictive step could be predictive socio-economic recovery of a neighborhood, post disaster, by observing traffic patterns on streets. Prescriptive phase of DL/ML is presently used in aircraft industry where the technology is being used to prescribe when a maintenance is required for an aircraft engine. An example of geospatial mission that national agencies can use Predictive recommendations from ML/DL include recommendations for pesticide applications based on potential pest infestations in agriculture. For GeoAI, the diagnostic, predictive, and prescriptive phases on ML/DL value chain offer tremendous opportunity for new research that can aid with various national mapping missions.

The AI/ML techniques can be grouped into two distinct categories: Descriptive and Generative. ML and DL techniques typically are descriptive in nature while Generative AI techniques and algorithms are used to generate new data or content that resembles, and often extends beyond, the original training data. Unlike traditional AI methods that focus on recognizing patterns in data or making predictions based on existing data, generative AI is capable of creating entirely new data instances that have not been explicitly seen before (Figure 7.4)

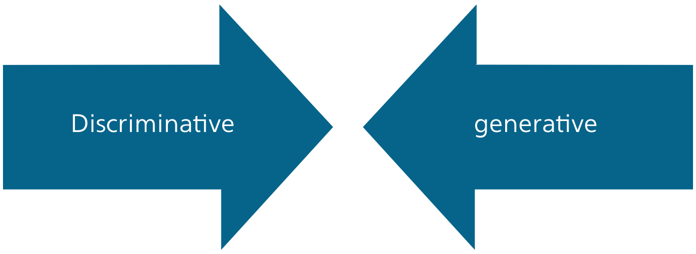


Figure 7.4 Traditional AI models are discriminative while GenAI models create new information

There are several approaches to generative AI, including: Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Autoregressive Models, Transformer Models such as OpenAI's GPT (Generative Pre-trained Transformer) series. Generative AI has a wide range of applications for GeoAI.

Large language models (LLMs) overlap with generative AI models and are designed to understand and generate human-like text. These models are built upon deep learning architectures, particularly transformer architectures, and are trained on vast amounts of text data. Some of the key characteristics and features of large language models include: Scale, Transformer Architecture, Pre-training and Fine-tuning, Generative and Predictive Capabilities, Contextual Understanding, and Versatility. Examples of large language models include GPT (Generative Pre-trained Transformer) models developed by OpenAI, BERT (Bidirectional Encoder Representations from Transformers) developed by Google, and T5 (Text-To-Text Transfer Transformer) developed by Google Brain.

There are few examples in the geospatial industry today where GeoAI professionals are starting to take advantage of GenAI and LLM's to address geospatial applications. These models have the potential to completely transform national mapping workflows in the future.

7.3.2 AI/ML workflows

First step in a typical AI/ML workflow (Figure 7.5) includes data preparation which consumes large amounts of resources and time. Data preparation includes creation of labels for training, testing and validation of the model performance.

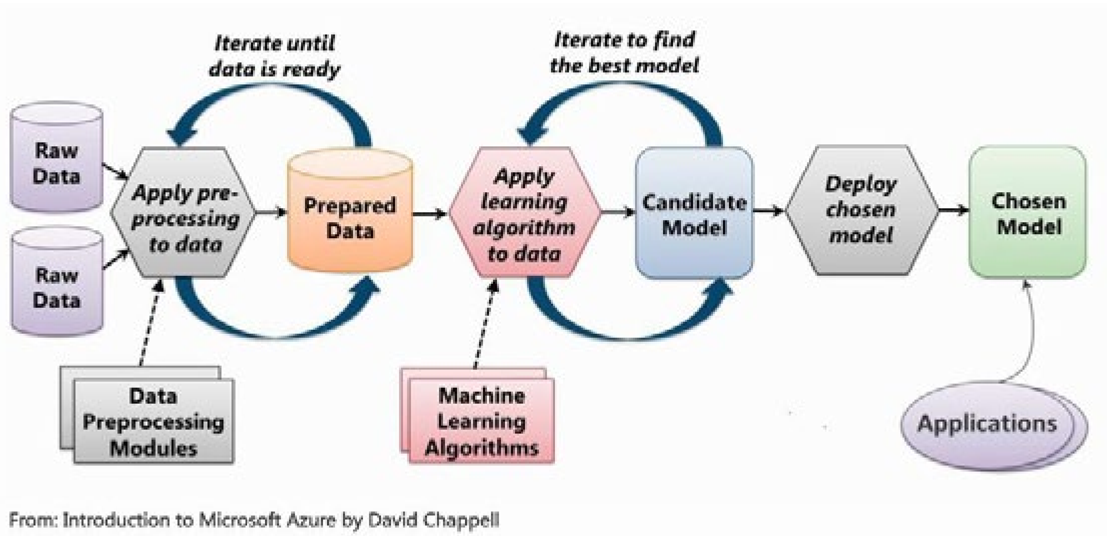


Figure 7.5 Typical AI/ML Workflow

Labels for features of interest are include image tiles with fixed dimensions (e.g. 64x64 pixels, 128x128 pixels, …) and involves drawing bounding boxes and/or polygons around features of interest (Figure 7.6).

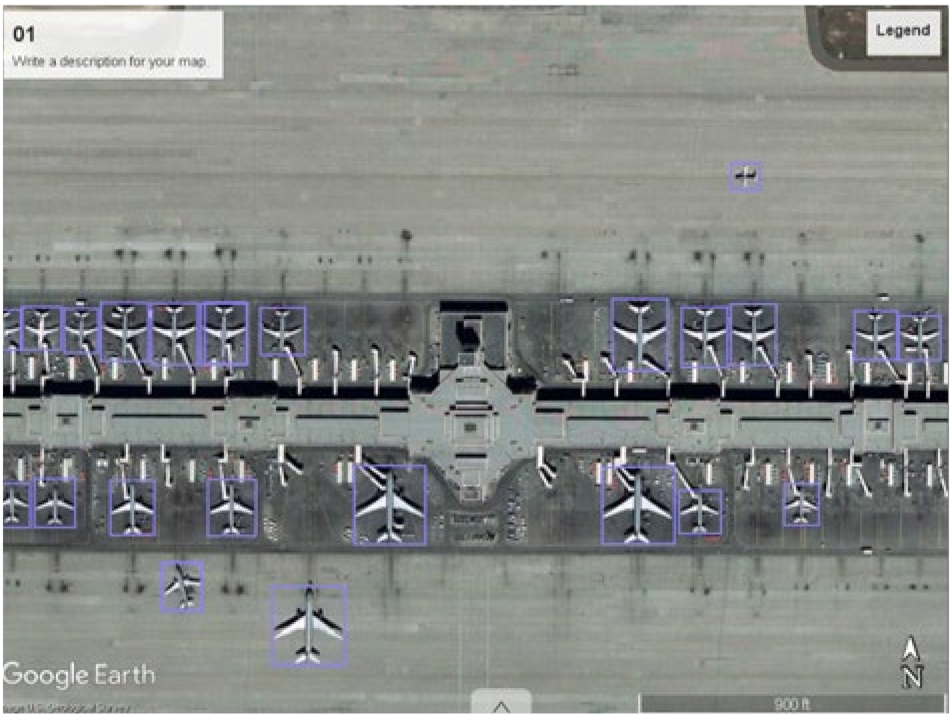


Figure 7.6 Labeling process includes drawing bounding boxes around features of interest

In many machine learning tasks, fixed image sizes for label images is not strictly required, but they can be beneficial depending on the specific requirements of the model and the nature of the task. Convolutional Neural Networks (CNNs) commonly used for tasks like image classification, object detection, and segmentation, typically require fixed-size input images. This is because the convolutional layers in CNNs have fixed-size filters that slide over the input image, and the size of the output feature maps depends on the size of the input image. Therefore, using fixed-size images ensures consistency in the input size across different samples and allows for efficient processing in the network. In data Preprocessing, even if the model does not strictly require fixed-size images, it's often beneficial to preprocess your data to have fixed-size inputs. Resizing images to a common size can simplify data preprocessing and model training pipelines. Additionally, it can improve computational efficiency and memory usage during training. Some models, such as certain types of recurrent neural networks (RNNs) or attention-based models, can handle variable-length inputs more naturally. The choice of model architecture may also influence whether fixed image sizes are required. Some architectures, such as fully connected networks, may require fixed-size inputs, while others, like convolutional and recurrent networks, can handle variable-size inputs more flexibly. There are several commercial companies and opensource tools that offer labelling technologies that can be leveraged by national agencies. There are ongoing efforts at global standards organizations such as Open Geospatial Consortium (OGC) to establish standards for labeling. The bounding boxes or polygons are saved in a predefined formats and directories, depending on the AI/ML model used for analysis.

Synthetic label generation has been employed in GeoAI to simulate various geographic backgrounds, varying atmospheric conditions, varying spatial resolution and other image artifacts. Synthetic label generation practices leveraging CAD models of objects of interest and insert them into imagery with varying backgrounds. Synthetic data labels could be of interest to national agencies for specific objects/features that might not have enough real world geospatial data labels. There are several open sources for labels for EO and SAR imagery datasets (Table 7.5) for national mapping agencies to get started with GeoAI applications.

Label Sources
SSDD (EO and SAR ship detection)
SSDD+
SAR SHIP Dataset
AIR-SARShip-1.0/2.0 Dataset
HRSID
LS-SSDD-V1.0
Official SSDD
SRSD-V 1.0
RSDD-SAR
iVision MRSSD
xView3

Table 7.5 Free Label Sources for EO and SAR imagery

There are evolving new developments in the AI/ML industry that are driving towards reducing manual labeling efforts. Zero shot training and Segment Me Anything (SAM) are some of the examples of these developments that will help national agencies in their mapping efforts. There is also research being conducted to extend the LLM's to recognize various types of semantic objects in an image with techniques such as Vision Language Models (VLMs) that eventually minimize the need for large label dataset creation.

Once the labels are created, next step in AI/ML workflow is to create subset the labels into 3 categories for: 1. Training 2. Testing and 3. Validating the models. Next step includes selecting and running an AI/ML of choice (discussed in the next section) and validating the results. The training, testing and validation of the model is reiterated by modifying various parameters (e.g. Epochs, label editing, etc.) and post processing steps (e.g. lower probability objects) until desired accuracy of object/feature detection is achieved. The final version of the models can be deployed to support real time operations by the national mapping agencies. Reinforcement techniques is another technique that involve continuous training and validation of the models by using the new datasets to account for varying imaging artifacts.

7.3.3 AI/ML approaches for image characterization

AI/ML models for image characterization fall into 5 categories: 1. Image Classification 2. Object Detection 3. Oriented Bounding Box Detection 4. Sematic Segmentation and 5. Pose Detection. Image classification models are designed to identify image tiles with one or more features of interest such as image tiles with agricultural fields (Figure 7.7). Image classification techniques are also ideal for area reduction for broad area search missions as well as for change detection.

Figure 7.7 Image tiles with agricultural fields across the globe

Object detection is the most common application in GeoAI and is used to identify objects such as buildings, cars, ships, planes, etc in images. Objects in the image tile are identified by bounding boxes (Figure 7.8) and each bounding box is associated with a confidence of prediction for one or more labels/ classes.

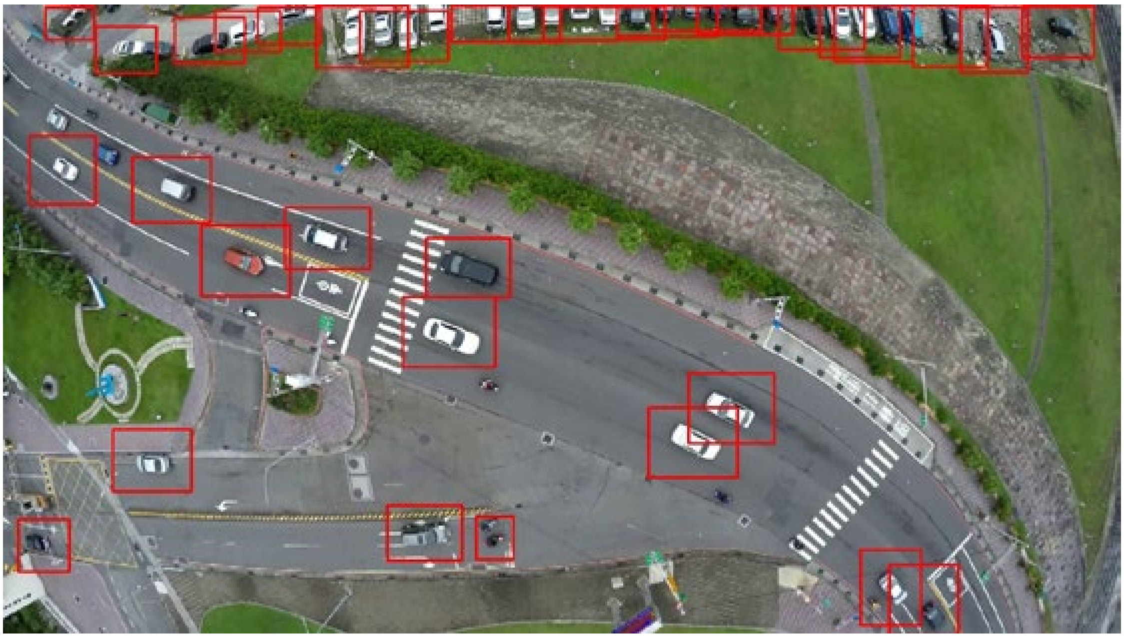


Figure 7.8 Object detection using AI/ML models

Oriented Bounding Box object detection is an improvement over object detection where the bounding boxes are oriented in the direction of the object of interest (Figure 7.9). This technique leverages Segment Me Anything (SAM) model to determine the size and direction of the objects of interest.

Figure 7.9 Ship detection with oriented bounding boxes

Image Segmentation identifies the outlines of various objects in an image and is useful for various GeoAI applications including extraction of building footprints, agricultural field boundaries (Figure 7.10), flood extent and others.

Figure 7.10. Segmentation techniques to extract agricultural field boundaries

Pose detection in the field of AI/ML involves estimating the pose (object position and orientation) of an object from an image or video. It aims to identify the spatial locations of key object joints (also known as keypoints). The relative positions of these keypoints can be used to distinguish one pose from another. GeoAI can leverage pose detection for object tracking missions such as MTI (Moving Target Indicator) of objects in 2D and 3D space and could be an area of interest to national mapping agencies in the future (Figure 7.11)

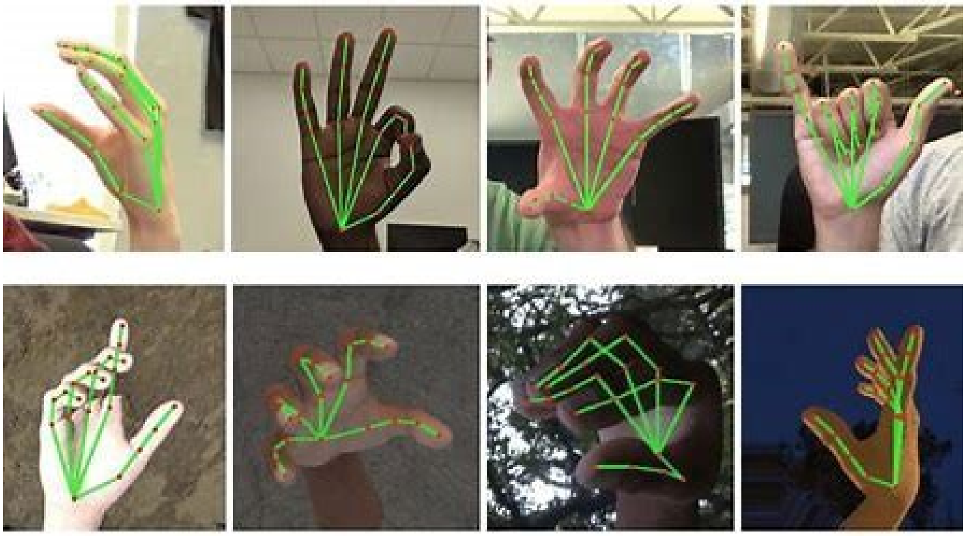


Figure 7.11 Pose detection can be used in GeoAI for MTI

7.3.4 AI/ML Techniques

Similar to the remote sensing image classification techniques, AI/ML techniques follow 2 approaches for image classification: 1. Supervised and 2. Unsupervised. Supervised Classification approaches require apriori training datasets that the GeoAI models leverages to characterize objects/features in an image. Unsupervised classification is gaining ground for object detection, outside the GeoAI field, and does not need any training datasets for image classification. A combination of opensource labeling datasets such MicroSoft COCO (Microsoft Common Objects in Context), a large-scale object detection, segmentation, key-point detection, and captioning dataset and emergence of LLM's for semantic correlation are driving the use on Unsupervised classification approaches. We will focus on Supervised AI/ML model training which is commonly used in GeoAI.

In general, there's two different approaches for AI/ML supervised classification: One stage, where a fixed number of predictions on grid are made or Two stage, that leverages a proposal network to find objects and then use a second network to fine-tune these proposals and output a final prediction. Each approach has its own strengths and weaknesses. An example of two stage detection is YOLO (You Only Look Once) where a pretrained model with pre- assigned weights can be adjusted to predict objects on interest in new images. Two stage detectors are the most commonly used techniques in GeoAI. Single shot detector (SSD) is a single stage detector where a pre-trained model for image classification is used as the backbone network. The model can be tweaked for a specific detection task. One major distinction between YOLO and SSD is that SSD model attempts to directly predict the probability that a class is present in a given bounding box whereas the YOLO model predicts the probability of multiple potential label classes. Another technique for AI/ML classification is Anchor-free object detection that has gained attention due to its speed and generalizability. Anchor-free methods directly predict object locations without predefined anchors/ boxes. Instead of bounding boxes, they predict points or key-points associated with objects. They are more generalizable and can extend to tasks like key-point detection and 3D object detection. Anchor-free object detection offers advantages in terms of simplicity, generalizability, and speed, making it a promising approach for improving small-size object detection models. Object training from scratch is another approach suitable for GeoAI for features that are not commonly used in pretrained models. There are several tools available for this approach. National agencies can start with two-stage models such as YOLO and keep track of new trends that can improve GeoAI workflows in the future.

7.3.4.1 AI/ML Frameworks

Frameworks are the backbone of AI/ML models. As the field of AI/ML advances, its intricacy grows, emphasizing the significance of frameworks in simplifying its processes. Historically, successful technologies have leveraged frameworks for efficient development. Acquiring proficiency in AI/ML frameworks not only saves time but also optimizes the development process. Some of the common frameworks used in the industry today include:

1. TensorFlow: TensorFlow is a free end-to-end open-source platform that has a wide variety of tools, libraries, and resources for AI/ML. It was developed by the Google Brain team and initially released on November 9, 2015. You can easily build and train Machine Learning models with high-level APIs such as Keras using TensorFlow. It also provides multiple levels of abstraction so you can choose the option you need for your model.

2. Caffe CAFFE (Convolutional Architecture for Fast Feature Embedding) was originally developed at the Berkeley Vision and Learning Center at the University of California and released on 18 April 2017. It is a deep learning framework written in C++ that has an expression architecture easily allowing you to switch between the CPU and GPU. Caffe also has a MATLAB and Python interface. Caffe is the perfect framework for image classification and segmentation as it supports various GPU- and CPU-based libraries such as NVIDIA, cuDNN, Intel MKL, etc. Caffe can currently process over 60M images in a day with a single NVIDIA K40 GPU which makes it one of the fastest options today. Because of all these reasons, Caffe is extremely popular in startups, academic research projects, and even multinational industrial applications in the domains of computer vision, speech, and multimedia.

3. Apache Spark: Apache Spark is an open-source cluster-computing framework that can provide programming interfaces for entire clusters. It was developed at Berkeley's AMPLab at the University of California and initially released on May 26, 2014. Spark Core is the foundation of Apache Spark which is centered on RDD abstraction.

4. PyTorch: Pytorch is a Machine Learning library that is based on the earlier open-source Torch library. It was initially released in October 2016 and is in primary use now that Torch is not actively in development anymore. PyTorch provides TorchScript, which facilitates a seamless transition between the eager mode and graph mode. Moreover, the torch.distributed backend provides scalable distributed training for Machine Learning and optimized performance.

5. Amazon SageMaker: Amazon SageMaker is a fully integrated development environment (IDE) for Machine Learning that was initially released on 29 November 2017. Amazon Web Services provides this Machine Learning service for applications such as Computer Vision, Recommendations, Image, and Video Analysis, Forecasting, Text Analytics, etc. Amazon SageMaker allows you to build, train, and deploy machine learning models on the cloud. The Amazon SageMaker Autopilot also has an automated machine learning capability that allows you to do all this automatically. Amazon SageMaker also allows you to create Machine Learning algorithms from scratch because of its connections to TensorFlow and Apache MXNet.

6. Accord.NET: Accord.NET is a Machine Learning framework that is completely written in C#. It was developed by César Roberto de Souza and was initially released on May 20, 2010. Accord.NET provides coverage on various topics like statistics, machine learning, artificial neural networks with various Machine learning algorithms like Classification, Regression, Clustering etc. along with audio and image processing libraries. Accord.NET libraries are available as source code, executable installers as well as NuGet packages. (Wherein NuGet is a free and open-source package manager that was created for the Microsoft development platform)

7. Microsoft Cognitive Toolkit: Microsoft Cognitive Toolkit is a Machine Learning or specifically, Deep Learning framework that was developed by Microsoft Research and initially released on 25 January 2016. You can easily develop popular deep learning models such as feed-forward DNNs, convolutional neural networks and recurrent neural networks using the Microsoft Cognitive Toolkit. This toolkit uses multiple GPUs and servers providing parallelization across the backend. You can use the Microsoft Cognitive Toolkit in a customizable manner as per your requirements with your metrics, networks, and algorithms. You can use it as a library in your Python, C++, or C# programs or you can use BrainScript, it's own model description language.

Machine Learning is a rapidly evolving field that has seen a significant surge in adoption by companies seeking to revolutionize industries. As this technology progresses, the need for frameworks becomes increasingly important to simplify processes and ensure efficient development. These frameworks provide the necessary resources to create advanced Machine Learning models tailored to specific requirements.

7.3.4.2 AI/ML Models

AI/ML models use a mathematical formula to make predictions about future events. They are trained on a set of data and then used to make predictions about new data. Some common examples of ML models include regression models and classification models. A deep learning model, or a DL model, is a neural network that has been trained to learn how to perform a task, such as recognizing objects in digital images and videos, or understanding human speech. Deep learning models are trained by using large sets of data and algorithms that enable the model to learn how to perform the task. The more data the model is trained on, the better it can learn to perform the task. DL models are composed of multiple layers of neurons, or processing nodes. The deeper the model, the more layers of neurons it has. This allows the model to learn more complex tasks by breaking them down into smaller and smaller pieces. For example, ResNet is a deep learning model for computer vision tasks such as image recognition. It is one of the deepest models currently available, with a version that contains 152 layers (ResNet-152). VGG (Visual Geometry Group) Deep Convolutional Neural Network Architecture   YOLO, or "You Only Look Once," is a deep learning model for real-time object detection. Surpassing YOLOv4 and YOLOR, the latest versions, YOLOv7 and YOLOv8, are super fast and very accurate, the current state of the art for several AI vision tasks.   Some of the most popular open-source AI models include: You Only Look Once (YOLO), SAM (Segment Me Anything), Regional-Convolution Neural Networks (R-CNN), and others.

AI/ML models are typically optimized for speed Vs accuracy. An example of various model sizes available for YOLOV810 range from nano (YOLOv8n) to extra large (YOLOv8x) with nano being is the fastest and smallest, while extra large is the most accurate yet the slowest among them. In addition, YOLO model iterations are managed by epochs. By trial and error and associated model performance statistics for best and last epochs, analysts can identify the right number of epochs to be used for to extract objects of interest. While most of the national mapping efforts need accuracy and can leverage extra large size models, situations such as disaster management can leverage small model for faster response. Python is commonly used programming language for running AI/ML model with several open source tutorials on how to run the models with custom datasets of interest to national mapping agencies and several open source tools are available for analysts to test different model sizes for GeoAI.

Post processing is one of the last steps in analyzing an AI/ML model performance that can improve the object detection accuracy. Some of the commonly used post processing steps include filtering the model results by size (absolute Vs relative size to image tile), Intersection and Unions of various bounding boxes identifying objects of interest, confidence threshold of various labels, and others.

Last few years have seen increasing adoption of AI/ML in the geospatial industry and there are examples of using AI/ML object/feature extraction at national/global scales. The following examples show the use of GeoAI for geospatial applications:

Figure 7.12 Plane detection from Aerial Imagery in Roboflow with post processing of Overlap threshold

Figure 7.13 Building footprint extraction using ArcGIS AI/ML platform

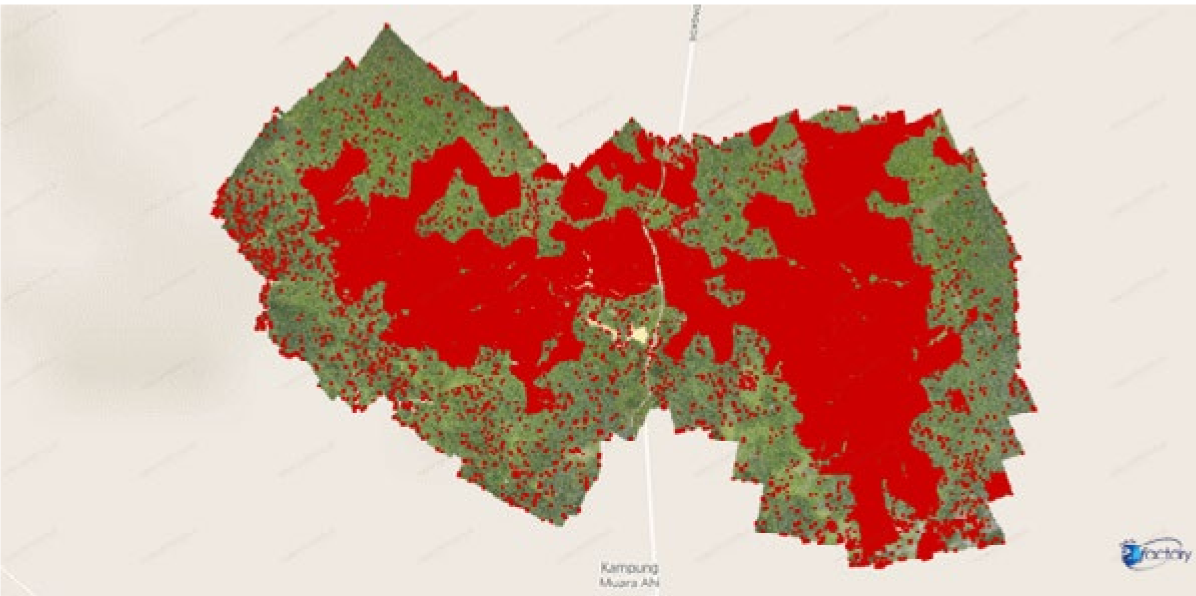


Figure 7.14 Plam tree extraction EoFactory.AI platform

7.3.4.4 GeoAI Orchestration

Combining generative AI with spatial reasoning and analysis techniques is the new frontier for automating authoritative and trustworthy spatial queries. Knowledge-based AI techniques have been around in geospatial sciences for decades while Machine-learning techniques and LLMs identify patterns from data, knowledge-based techniques rely on automated reasoning with symbolic representations of data. GeoAI Orchestration combines multiple AI and analytical tools into workflows. Knowledge-based reasoners and authoritative data repositories are used to generate reliable and trustworthy responses to questions. AI "orchestration" is emerging as a foundational approach to combining the exceptional text-extruding capabilities of LLMs with reliable knowledge-based spatial reasoning and analysis

7.4 Applications of national maps beyond traditional mapping

Several of the datasets such as imagery, elevation models, GCP's, transportation networks, and others can be used for multiple missions within national agencies ranging from agriculture, census, transportation, environment, natural resources management, coastal management, disaster response and others. Countries such as USA have created federal geospatial programs such as 3DEP (3D Elevation Program), NAIP (National Agriculture Imagery Program), US Census TIGER Road Network, and others that are leveraged by a combination of agencies across state/local/federal as well as my military agencies. Some of the agencies in USA were able to save hundreds of millions of dollars by leveraging the national geospatial programs. Table 7.6 shows various national missions and data requirements that can leveraged for a unified nationwide geospatial program.

Mission	Data	GeoAI Application	Requirements
Census	EO Imagery	Building footprints, change detection	Blds: EO - RGB - 50cm Change: 2m-5m MS
Agriculture	EO Imagery,	Field Boundaries, Crop Type, Change Detection	Field Boundaries: EO RGB 50cm Crop Type: EO Imagery, MS, 50cm Change Detection: EO Imagery MS 2m-5m
Forestry	EO Imagery, DEM	Tree Identification, Tree Height/width, De/Reforestation	Deforestation: EO MS 2-5m Tree Height: DEM, 50cm postings Change Detection: EO Imagery MS 2m-5m

Mission	Data	GeoAI Application	Requirements
Cadaster	EO Imagery, GCP's	Agriculture Parcels, Urban Parcels, Change Detection	Ag Parcels: EO RGB 50cm; GCP's <1 meter accuracy; Urban Parcels: EO RGB 5-15cm; GCP's: 5cm accuracy Change Detection: EO 1-2 meters
Disaster Planning/ Response	EO/SAR magery, DEM	Flood plain mapping, Damage assessment	Flood plain: DTM, 1 m postings Disaster: EO Imagery, SAR Imagery, 50cm
Water Resources Management	EO Imagery, DEM's	Watershed Modeling	EO Imagery: MS 1-5 meters DTM: 1-2 Meters
Mineral Resources	EO Imagery, DEM	Mineral Deposits, Mining	Mineral: EO Imagery; HS: 1-2 Meters Mining: DSM's <1m
Atmoshperic Monitoring	EO Imagery	Methane, Particulate Matter	EO Imagery: MS or HS: 5-50 meters
Coastal Monitoring	EO Imagery	Coastlines, Bathymetry	Coastlines: EO RGB <1 meter Bathymetry: EO MS, <1 meter
Transportation	EO Imagery, DEM	Road/Rail Mapping, Road furniture for Autonomous Vehicles	Road Rail: EO Imagery: <=50cm Autonomous: EO Imagery <=15cm; StreetViews: 3D <=30cm
Telecommunications	EO Imagery, 3D	Cell tower detection	EO Imagery: MS <50cm 3D: <=30cm
Oil & Gas	EO Imagery, DEM	Detect roads, infrastructure	EO Imagery: MS: <50cm DTM: <=50cm

7.5 References

1. NIIRS – Wikipedia, https://en.wikipedia.org/wiki/NIIRS
2. SAR data applications in earth observation: An overview - ScienceDirect
3. www.OpenStreetMap.org
4. www.hotosm.org
5. https://overturemaps.org
6. Microsoft Releases Building Footprints | Bing Maps API
7. https://landscan.ornl.gov
8. WorldPop

LESSON – 8

Fundamental Geospatial Data Themes

Trainer's Biodata: Prof Dr Zaffar Sadiq Mohamed-Ghouse

Prof Dr Zaffar Sadiq Mohamed-Ghouse is a Director of Advisory & Innovation at Woolpert, a leading global architecture, engineering, and geospatial firm. He has over 25 years of experience in providing strategic advice and solutions to several governments and international organizations on geospatial and space-related matters, such as positioning networks, data policies, spatial digital twins, geospatial ecosystems, natural resources, and Geospatial master plans. He has led several multi-million-dollar national mapping projects for various countries and regions, including Australia, New Zealand, Malaysia, Indonesia, India, United Arab Emirates, and Saudi Arabia.

As a Fellow of the Institution of Engineers Australia, a Fellow of the Royal Geographical Society, and an Honorary Fellow of the Geospatial Council of Australia, Zaffar is recognized as one of the spatial leaders in Australia and globally. He is also the Chair of the United Nations Global Geospatial Information Management - Private Sector Network, where he facilitates collaboration and innovation between the private sector and the UN-GGIM community. He is passionate about advancing the spatial industry and its contribution to the sustainable development goals. He has published extensively in GIS, surveying, remote sensing, data management, and spatial data / knowledge infrastructures and Geospatial Ecosystems. He is the recipient of several prestigious awards, including the Professional of the Year 2022 by the Asia-Pacific Spatial Excellence Awards and University of Melbourne's Prestigious Thornton Smith Medal for outstanding service to Geospatial and Surveying Community. Zaffar has academic appointments at University of Melbourne, Anna University and National Institute of Advance Studies. He served on the Programme Board of Geospatial Knowledge Infrastructure since its inception. He is currently the Vice president of the International Society of Digital Earth and serves as Board of Directors of the Open Geospatial Consortium and Programme Board of Group on Earth Observation.

8.1 Introduction to global fundamental geospatial data themes, as adopted by UN-GGIM

Data and the insights derived from it are today central to the development of industry, the functions of the government, as well as research. Aided by an ongoing process of Digital Transformation that promotes or highlights the role of data as a digital currency within real-world organizations, their activities, and transactions or interactions between organizations, individuals, or the physical environment. The significance of ensuring the availability of authoritative data for the sustenance of emerging data economies across the globe, at local, national, and regional scales, can simply not be emphasized enough.

UNGGIM, at its seventh session under decision 7/104, adopted the proposed minimum list of Global Fundamental Geospatial Data Themes. At its eighth session, detailed theme descriptions were presented.

"Fundamental data sets are the minimum primary sets of data that cannot be derived from other data sets, and that are required to spatially represent phenomena, objects, or themes important for the realization of economic, social, and environmental benefits consistently across a polity at the local, national, sub-regional and regional levels." [1]

Fundamental geospatial data themes are a form of public infrastructure and underpin geospatial data-driven business and policy decisions. Providing a consistent approach to the definition and documentation of national-level spatial information. Thus, enabling the ubiquitous use of national spatial datasets in the everyday business of government and private sector entities alike, leading to national productivity gains and greater levels of innovation across all areas of government and industry. Implementing the themes will necessitate the integration of information from National Geospatial Information and Mapping Agencies, National Statistical Offices, and other institutions to produce standardized, authoritative data, for use within a given domain. The data themes are available to inform the data value chain from the local to the state level, focusing on fostering a national-level information management environment.

8.2 Description and significance of data themes

This section seeks to provide a general overview of the 14 global fundamental geospatial data themes, by providing a description, establishing the need for a specific theme, describing datasets that would fall under a given theme, and where the datasets are typically sourced from. The existing global standards of relevance to each theme are also listed, in view of enabling standardization of foundational data themes globally.

1 https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/Fundamental_Data_Publication.pdf

8.2.1 Global Geodetic Reference Frame

Definition

The Global Geodetic Reference Frame is the framework which allows users to precisely determine and express locations on the Earth, as well as to quantify changes of the Earth in space and time.

Datasets Description and Source

Examples of data that would form a part of this theme would include existing Geodetic Reference Frames such as the International Celestial and Terrestrial Reference Frames (ICRF and ITRF) as well as Geodetic Reference Frames that are defined and established nationally. Data belonging to this theme is often a product of component technique observing systems such as Global Navigation Satellite Systems (GNSS), Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR) and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS), and their associated data centers, analysis centers or physical infrastructure. Such as gravimetric products and physical height systems.

Rationale

Referencing spatial data to a commonly agreed upon and established Geodetic Reference Frame, allows for its interoperable use, as it can now be confidently reinstated, integrated, or projected through time and for use at a different epoch.

Common Standards

- International Earth Rotation and Reference System Service (IERS) conventions that contain models, constants, and standards.

Use Case

Norwegian Mapping Authority's geodetic Earth observatory in Ny-Ålesund, supports the precise monitoring of issues such as ice loss and sea-level change, by enabling precise observations on the surface of the Earth.

8.2.2 Addresses

Definition

An Address is used to identify a plot of land, a building or part of a building, or some other construction, together with coordinates indicating their geographic position.

Rationale

An address is often the unit to which a public service, such as water, is provided. Addresses also enable effective communication with and amongst citizens, informing them of policies applying to them, and notifying them of relevant incidents, playing a vital role in connecting otherwise unrelated information. Geocoding addresses relates such information to geographic location. This allows for location-based data analytics and data mining.

Data Sources and Custodians

Address data may be created and maintained at local level, while ideally being compiled into a single national register.

Datasets Description

Addresses comprise at least one locator (building, floor or apartment number and/or name), a two-dimensional geographic position and a number of address components which place the address within other features such as a road, a locality, an administrative unit or postal code, taking the form of a structured label.

Common Standards

- INSPIRE Data Specification on Addresses – Technical Guidelines 3.1.
- ISO 19160-1:2015 Addressing -- Part 1: Conceptual model.
- ISA Program Location Core Vocabulary; and,
- ISO 19160-4(UPU, Universal Postal Union) Addressing--Part4:Internationalpostaladdress components and template language.

Use Case

Addresses are critical for the delivery of goods, e-commerce, emergency services and other public, private, and government processes.

8.2.3 Buildings and Settlements

Definition

A Building refers to any roofed structure permanently constructed or erected on its site. Settlements are collections of buildings and associated features.

Rationale

This theme allows for the physical description of real-world entities seen as construction. This theme's main use is to locate population and its distribution. Enabling use cases, including collection of statistics, provision of public services, resource management, emergency management, planning for urban development, and application to mapping and navigation.

Common Standards

- INSPIRE Data Specification on Buildings
- CityGML
- BIM Data Standards.

Datasets Description

Minimum attributes for both are the location geometry and an identifier of some type. Additional attributes of buildings include its address and its functional classification. For settlements additional attributes might include an indication of population size.

Data Sources and Custodians

Building and Settlement datasets are usually maintained by public authorities, at national and sub-national/regional or local level. Ideally compiled into a single national register.

Use Case

Building databases are directly a reference dataset for the location of settlements. An estimate of where people live is a good basis to model economic activity, planning disaster response and incumbent resource allocation, and the impact response for disease outbreaks.

8.2.4 Elevation and Depth

Definition

An Address is used to identify a plot of land, a building or part of a building, or some other construction, together with coordinates indicating their geographic position.

Rationale

An address is often the unit to which a public service, such as water, is provided. Addresses also enable effective communication with and amongst citizens, informing them of policies applying to them, and notifying them of relevant incidents, playing a vital role in connecting otherwise unrelated information. Geocoding addresses relates such information to geographic location. This allows for location-based data analytics and data mining.

Data Sources and Custodians

Address data may be created and maintained at local level, while ideally being compiled into a single national register.

Datasets Description

Addresses comprise at least one locator (building, floor or apartment number and/or name), a two-dimensional geographic position and a number of address components which place the address within other features such as a road, a locality, an administrative unit or postal code, taking the form of a structured label.

Common Standards

- INSPIRE Data Specification on Addresses – Technical Guidelines 3.1.
- ISO 19160-1:2015 Addressing -- Part 1: Conceptual model.
- ISA Program Location Core Vocabulary; and,
- ISO 19160-4(UPU, Universal Postal Union) Addressing--Part4: International postal address components and template language.

Use Case

Addresses are critical for the delivery of goods, e-commerce, emergency services and other public, private, and government processes.

8.2.5 Functional Areas

Definition

Functional Areas are the geographical extent of administrative, legislative, regulatory, electoral, statistical, governance, service delivery and activity management areas.

Rationale

Linking data to administrative and functional geographies provides the spatial component that further enables data integration and broader comparability. It's also the key to informing stakeholders on the status of international, national, and sub-national policy objectives and program goals/ deliverables. Functional areas form the link between data collection and implementation of actions. They can be used to visualize data and trends.

Data Sources and Custodians

Data relating to this theme is usually available from public sources. These sources may be at different levels of government.

Datasets Description

Functional Areas are essentially human-defined virtual areas, often organized in a hierarchical way. Their key attributes are geometry, level (in the hierarchy), code, name, and function.

Common Standards

- INSPIRE Data Specifications for Administrative Units.
- INSPIRE Data Specifications for Statistical Units.
- INSPIRE Data Specifications for Protected Sites.
- ISO 14825 Intelligent transport systems-Geographic Data Files (GDF)-GDF5.0.
- ISO 19152: Land Administrative Domain Model (Spatial Unit Group).

Use Case

Functional areas include protected sites, planning zones, statistical units, flood zones, school catchments, agricultural zones, administrative areas.

8.2.6 Geographical Names

Definition

They are location identifiers for cultural and physical features of the real world, such as regions, settlements, or any feature of public or historical interest.

Rationale

Geographical Names provide orientation and identity to places and have the potential to inter-relate and cross-reference disparate data sources, both spatial and non-spatial. Geographical names are the most common, understandable, and widely used entry-point for broader searches for geospatial data and information and are therefore, necessary as search criteria in gazetteers, geoportals, spatial data catalogues etc. Geographical names are also required for a wide range of topographical and thematic map output at any scale.

Data Sources and Custodians

National geographical names datasets are usually maintained by public authorities for features on land, coastal or marine areas.

Datasets Description

A named place (e.g. settlement, mountain, bay) may have several names in different languages. Many named features have indeterminate boundaries but, where feasible, their delineation should be included.

Common Standards

- Technical reference manual for the standardization of geographical names, (UNGEGN), 2007, ISBN: 92-1-161500-5.
- INSPIRE Data Specification on Geographical Names – Technical Guidelines 3.1.
- ISO 639 Language Code List for the language of origin of geographical names.
- UTF-8-character set (UNICODE) for the exchange of syllabics, diacritics and other special characters.

Use Case

The Geographical Names theme may comprise both geographical names as attributes of feature types that are already in another fundamental data theme, such as Transport Networks, Water or Functional Areas, Buildings/Settlements and as named places, as feature types, that are not yet in another theme. Also, a named place (e.g. settlement, mountain, bay) may have several names in different languages.

8.2.7 Geology and Soils

Definition

Geology is the composition and properties of geologic materials (rocks and sediments) underground and outcropping at the Earth's surface. It includes bedrock, aquifers, geomorphology for land and marine environments, mineral resources, and overlying soils. Soils are the upper part of the Earth's crust, formed by mineral particles, organic matter, water, air, and living organisms.

Rationale

As the interface between rock, air and water hosting most of the biosphere, Geology and soils information has the potential to inform us about best practices in land management, hazard avoidance, soil erosion or salinity, soil pollution, nuclear waste storage, crop suitability, and conditions that affect the structural engineering of buildings.

Data Sources and Custodians

Global Lithological Map (GLiM), OneGeology, National Geologic Surveys, Harmonized World Soils Database; and National Soil Surveys.

Datasets Description

Geology data is generally characterized according to composition, structure and age. It also provides knowledge about aquifers, i.e. subsurface units of rocks or sediments of sufficient porosity and permeability.

Common Standards

- Geology
  - USGS NCGMP'09,
  - GEOSciML.
- Soils:
  - FAO: World Reference Base for Soils Resources 2006; and,
  - USDA NRCS SSURGO Data Model.

Use Case

Geology and Soils information have the potential to better inform us about best practices such as land management hazard avoidance, soil erosion or salinity, soil pollution and degradation, nuclear waste storage, the regolith, crop suitability and conditions that affect the structural engineering of buildings. This theme is also a key component in the analyses to estimate fossil fuel reserves, and site suitability for renewable energy sources.

8.2.8 Land Cover and Land Use

Definition

Land Cover represent the physical and biological cover of the Earth's surface. Land Use is the current and future planned management, and modification of the natural environment for different human purposes or economic activities.

Rationale

The use of the land informs land management impacts, especially on changes in natural resources, agriculture, conservation, and urban developments. It is required at a disaggregated level to allow local planning to manage and monitor land use at land parcel level.

Data Sources and Custodians

Classified Earth observation (EO) data, potentially as a Data Cube; National datasets relating to environmental information and land parcels; and, international organizations, Regional United Nations Centre, different levels of public authorities (municipalities) and the private sector.

Datasets Description

Both Land Cover and Land Use are separated into different classes based on an agreed classification schema which is usually hierarchical. The data can be represented either as polygons or as a raster. It may also be found as attributes of a land parcel.

Common Standards

- ISO 19144-1:2009 – Geographic Information Classification system – Part 1 Classification system structure (last reviewed in and confirmed in 2015).
- ISO 19144-2:2012 - Part 2 - Land Cover Meta Language (LCML) (there are limitations on this standard).
- ISO 19115:2003 Geographic information – Metadata; and,
- INSPIRE data specification on Land Cover and on Land Use.

Use Case

Land Cover is required, for example, for developing land management policy, understanding spatial patterns of biodiversity and predicting effects of climate change and may also help to forecast other phenomena, such as coastal erosion and flooding. It is critical data in national assessments of biodiversity, conservation efforts, and water quality monitoring.

8.2.9 Land Parcels

Definition

Land Parcels are areas of land or more generally of the Earth's surface (land and/or water) under common rights (such as ownership or easements), claims (such as minerals or indigenous land) or use.

Rationale

They are a necessary part of a good secure land tenure system which in turn contributes to economic development by enabling investments. Land parcel data is required for land management, infrastructure management and spatial planning.

Common Standards

- INSPIRE Data Specification on Cadastral Parcels – Technical Guidelines 3.1.
- ISO 19152: Land Administration Domain Model; and,
- International Land Measurement Standard (ILMS).

Datasets Description

The Land Parcels theme mainly comprises the feature land parcel with three basic attributes: geographic location; unique identification of the parcel; and type of parcel. Land parcels may be associated with land registries (or equivalent) that establish the rights (and possibly the restrictions and responsibilities) that a party (a natural or legal person) has on a land parcel - on ground, below ground or above ground.

Data Sources and Custodians

Land parcel data is frequently found in registers supporting the land tenure and transfer system of a country. It may also be part of a taxation or planning system.

Use Case

Land parcels may be associated with land registries (or equivalent) that establish the rights (and possibly the restrictions and responsibilities) that a party (a natural or legal person) has on a land parcel - on ground, below ground or above ground.

8.2.10 Physical Infrastructure

Definition

The Physical Infrastructure theme includes industrial & utility facilities, and the service delivery facilities associated with administrative & social governmental services.

Rationale

The infrastructure of a country is how services are delivered to the population, be that hospitals, schools, energy or telecommunications. As such spatial data on the physical infrastructure present within a given polity enables numerous real-world use-cases.

Common Standards

- INSPIRE environmental facilities specification and guidance.
- INSPIRE Utility and governmental services specification and guidance.
- INSPIRE production and industrial facilities specification and guidance.

Datasets Description

These are human-made features, and this theme relates mostly to the built environment, and focuses on the service infrastructure. Key attributes are the geometry, an identifier, a name (if any) and the type of service (school, hospital, …) and other relevant attributes.

Data Sources and Custodians

National and local government Utility and services companies

Use Case

A nation's physical infrastructure consists of a broad array of systems and facilities that house and transport people and goods and provide services. Among other things, this infrastructure includes industrial and utility facilities and administrative and social governmental services such as public administrations, civil protection sites, water reticulation, sewerage, stormwater drainage, schools and hospitals and postal and telecommunications services.

8.2.11 Population Distribution

Definition

The Population Distribution theme covers the geographical distribution of people, including population characteristics.

Rationale

It's vital to understand the spatial distribution of the population and its characteristics, as well as how population impacts urbanization, regional development, or sustainability. It's important to understand how policy and programs impact migration and population growth, as well as economic activity, social well-being, and quality of life.

Datasets Description

Usually expressed as a count of people within an area unit such as a census tabulation area, or an estimate within a city, postal code area, country, state, or province. Include specific data cohorts such as at-risk populations, such as elderly, family composition, children, indigenous population, immigrant/ethnic origin, and education.

Common Standards

- U.N. Demographic manuals.
- ISO: 18391:2016.
- INSPIRE data specification on Population Distribution.

Data Sources and Custodians

United Nations Population Division, individual country census agencies, National Statistical Institutes, and commercial providers such as Michael Bauer Research GmbH, Environics, Esri, and Facebook.

Use Case

This theme supports policies to improve and create sustainable rural and urban living conditions. Enables understanding of how policy and programs impact migration and population growth, as well as economic activity, social well-being, and quality of life.

8.2.12 Orthoimagery

Definition

Orthoimagery is geo-referenced rectified image data of the Earth's surface, from satellite or airborne sensors.

Rationale

It's useful for either human visualization or machine interpretation. This image data is a record of the Earth's surface at the time of imaging, which then has immense historical value in the future. Orthoimagery may be exploited using algorithms for automatic processing to extract features and information such as buildings, roads, vegetation, soil moisture & water content, cloud cover, and to detect changes such as land use.

Common Standards

- INSPIRE Data Specification on Orthoimagery – Technical Guidelines.
- FGDC content standards for digital orthoimagery.
- USGS National Geospatial Program. Digital Orthoimagery Base Specification V1.0.

Datasets Description

The Orthoimagery theme includes image products generated from sensors aboard drones, aircraft or satellites. Once captured by the sensor, these images may be subject to a variety of treatments which are designed to increase data interpretation capabilities.

Data Sources and Custodians

National Mapping Agencies usually carry out periodical space/aerial photo capture and orthophoto production through national programs. Public and private remote sensing satellite operators capture massive amounts of satellite images which are distributed through multiple channels, both commercial and non-commercial.

Use Case

Once captured by the sensor, these images are subject to some of these processes which are designed to increase interpretation capabilities. It can be used to support urban planning, risk identification and resilience planning, while supporting many other fundamental uses.

8.2.13 Transport Networks

Definition

Transport Networks are the suite of road, rail, air, cable and water transport routes and their connectivity.

Rationale

Essential for the effective planning, development, delivery and maintenance of transport infrastructure or transport asset management. This data theme also enables the generation of information on mobility and logistics over the transport infrastructure.

Common Standards

- INSPIRE Data Specification on Transport Networks – Technical Guidelines 3.2.
- ISO/AWI 8148:2012 Geographic Information –location-based services – Linear referencing System.
- ICAO Standards and Recommended Practices.
- ISO 14825 Intelligent transport systems-Geographic Data Files (GDF)-GDF5.0.

Datasets Description

Their main attributes are location, name, identification codes, category, classification (road, rail, air, water) and connectivity. Transportation data includes topographic features related to transport by road, rail, water, and air. It is important that the features form networks where appropriate, and that links between different networks are established.

Data Sources and Custodians

National Mapping Agencies, Transport Ministries, and International Civil Aviation Authority

Use Case

An example use case would be a map of global shipping routes and harbors using data from the FAO, NOAA, and Esri. With say the thickness of the lines connecting global ports representing the mass of goods transported.

8.2.14 Water

Definition

The Water theme covers the extent and conditions of all water features including rivers, lakes and marine features.

Rationale

Study of the Earth's water cycle helps understand how it interacts with the environment and how much is available for human use. Conversely, it's also a potential danger to people and property. Coastal and transitional waters as well as the shoreline and the shore are relevant since these areas of land-water proximity are significant in terms of environment and intense economic activity.

Datasets Description

Features include rivers, lakes, reservoirs, marine & glacial features, and groundwater. All features will have a geometric location and identifier of some type such as a name. Possible attributes will vary across quality: e.g. physical parameters such as temperature, pH, e-coli, turbidity, salinity, etc.; quantity: e.g. volume, direction, velocity; and form: e.g. ice, snow, fresh, salt, season patterns e.g. monsoons.

Use Case

Provide data to characterize all aspects of freshwater, including surface and ground water resources and the interface between fresh and sea water.

Common Standards

- INSPIRE data specification on Hydrography.
- S-44 - IHO Standards for Hydrographic Surveys.
- S-57 - Transfer Standard for Digital Hydrographic Data.
- S-100 - IHO Universal Hydrographic Data Model.
- ISO 1900 - Series of geographic standards for hydrographic, maritime and related issues.
- ISO 14046:2014 - Environmental management -- Water footprint.
- ISO/TC 147 - Water quality; • UNSD - International Recommendations for Water Statistics (IRWS).
- OGC® WaterML 2 Part 3 - Surface Hydrology Features (HY_Features) - Conceptual Model

Data Sources and Custodians

AQUASTAT - FAO database on water, UN-Water statistics and UN Statistics Division environment indicators, International Hydrological Program (IHP) Water Information Network System (IHP-WINS), Global Water Forum and Global Water System Project, and the International Water Management Institute (IWMI) Eco-Hydrological Databases.

8.3 Implementing the Fundamental Geospatial Data Themes

Implementing the Fundamental Geospatial Data Themes in line with facilitating a national level information management environment would call for addressing the practical challenges associated with the heterogeneity in the production, management, supply, and governance of geospatial data across multiple levels of government. Foundation Spatial Data Framework (FSDF), an initiative from ANZLIC (Spatial Information Council) addresses the challenge above through means of a data specification framework underpinned by model driven approaches. A data specification framework supports the development and maintenance of a suite of foundation data products, by providing the ability to:

- capture and model user requirements for data products.
- document current data products.
- unambiguously define and document concepts, relationship and classifications used within and across domains (Fundamental themes).
- design and implement new products, meeting articulated requirements, using agreed concepts.
- support gap analyses between current and future states for spatial products; and
- support evolution of a suite of interrelated products.

The key components of a data specification framework would include:

- a suite of modular interdependent models and controlled vocabularies that define foundational spatial data and the way in which it is used.
- modelling tools and processes – the tools, systems and processes used to create access and exploit models and vocabularies; and
- model and vocabulary governance -the roles, processes, rules, and mechanism for the governance of models throughout their entire lifecycle from creation, publication, use and retirement.

Modelling is a critical step in the data specification process, which incorporates product design and implementation. FSDF adopts an information modelling approach known as Model Driven Architecture, which places the structural definition of the information at the centre of the design process. Formal modelling enables the development of interrelated spatial data products, followed by efficient production, maintenance, and use of suites of interoperable, standards-based foundation spatial data. The information modelling process can be split across the three stages of conceptual modelling, logical modelling, and physical modelling.

- Conceptual Models: Describes the domain i.e. data theme, in terms of its composite concepts and relationships, as required to meet the themes use-cases. It generally includes the following elements:
    - > the definition of the spatial and temporal representations of and between spatial objects.
    - > reference to common spatial and temporal reference systems as well as multilingual thesauri.
    - > the patterns for unique object identifiers; and
    - > any cross-domain constraints.
- Logical Models: Describes the concepts, their relationships, properties, and the data types required to meet the theme specific use-cases and ensure that the information products are related and coherent. Provides a detailed representation of the theme's data, independent of any particular data management technology.
- Physical Models: Describes the models previously defined within the specific context of an implementation platform. Physical models are specialisations, or profiles, of the logical models, in that they constrain the features, attributes and cardinalities, of the model to meet one or more specific use cases.

The entire process of information modelling is depicted in the figure:

Information models play a major role in achieving more efficient Geospatial data supply chains that deliver timely and high-quality interoperable foundation products. Three primary actors are involved in supply chains: provider, intermediary, and user. These supply chains span multiple agencies and as data progresses through the supply chain, it may undergo data format conversion, interpretation, data transformation, integration, and harmonisation with other data to produce an end product. A data specification framework underpinned by information modelling approaches can result in the following patterns of Geospatial data supply chains:

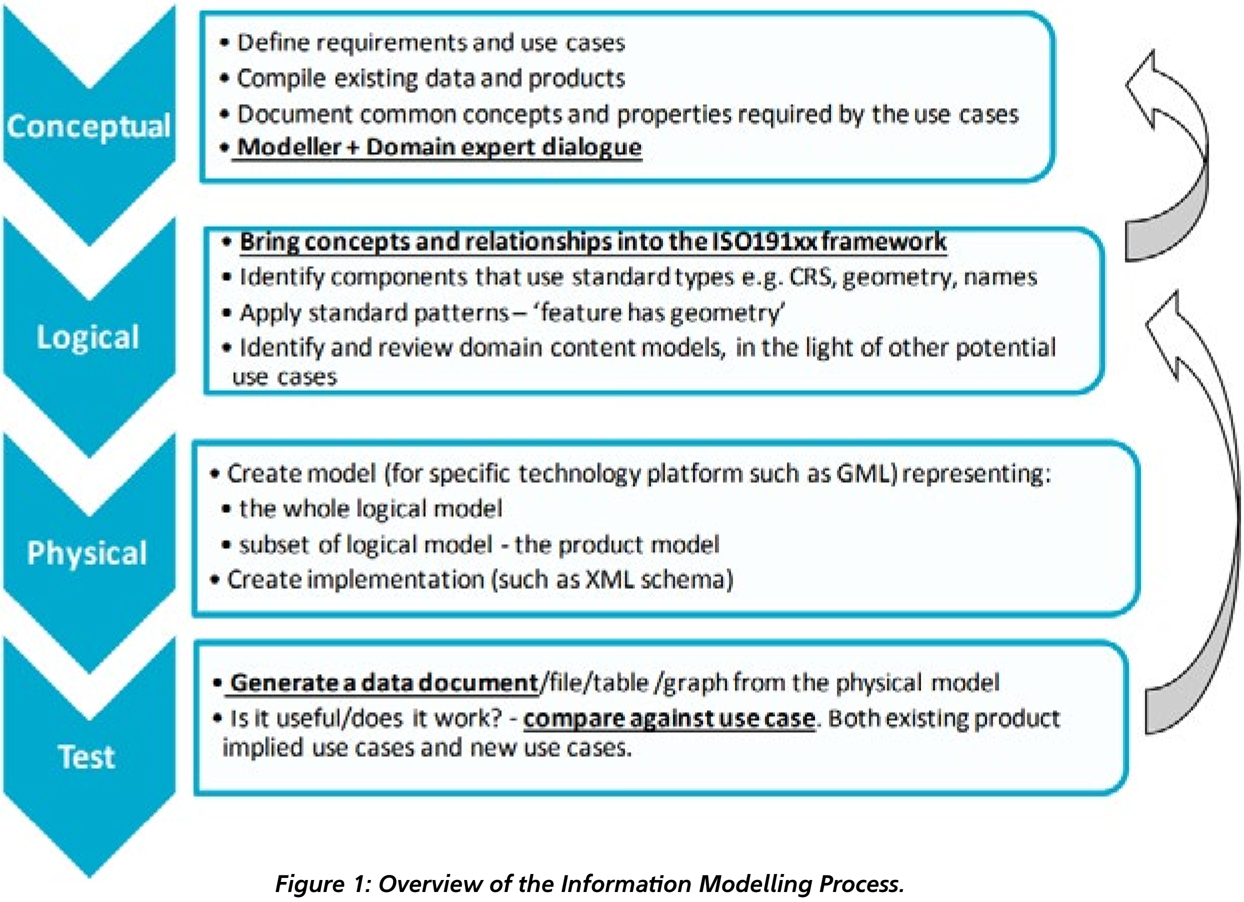


Figure 1: Overview of the Information Modelling Process.
Source: FSDF Data Specification Framework.

Information models play a major role in achieving more efficient Geospatial data supply chains that deliver timely and high-quality interoperable foundation products. Three primary actors are involved in supply chains: provider, intermediary, and user. These supply chains span multiple agencies and as data progresses through the supply chain, it may undergo data format conversion, interpretation, data transformation, integration, and harmonisation with other data to produce an end product. A data specification framework underpinned by information modelling approaches can result in the following patterns of Geospatial data supply chains:

- **Anarchic** (or point to point): Direct producer and user interaction without the involvement of an intermediary. Involves multiple users where each develops their own schema and integrates supplied data into the structure at the end point.
- **Centralised**: Centralised production of data by a single organisation that is tightly coupled application specific. A single organisation develops the schema and is responsible for collecting all the data itself or has private arrangements with the collectors.
- **Aggregated**: Aggregation and integration of data by a single intermediary. The intermediary develops the schema based on identified business needs, and aggregates heterogeneous data from multiple providers, each of which publishes data according to a different structure and format, and publication method, to create a coherent national product.
- **Brokered**: Centralised broker transforms data to a common form from data supplied using separate models. A centralised broker service transforms heterogeneous data supplied by data providers as services in real-time to a common structure based on a community application schema.
- **Federated**: Federated data supply using common community models. In this pattern the data providers provide a view of their data according to a community agreed model. Mapping of the storage data structure to the community schema is performed in a feature service hosted by each provider. This way end users can access services from multiple data providers.

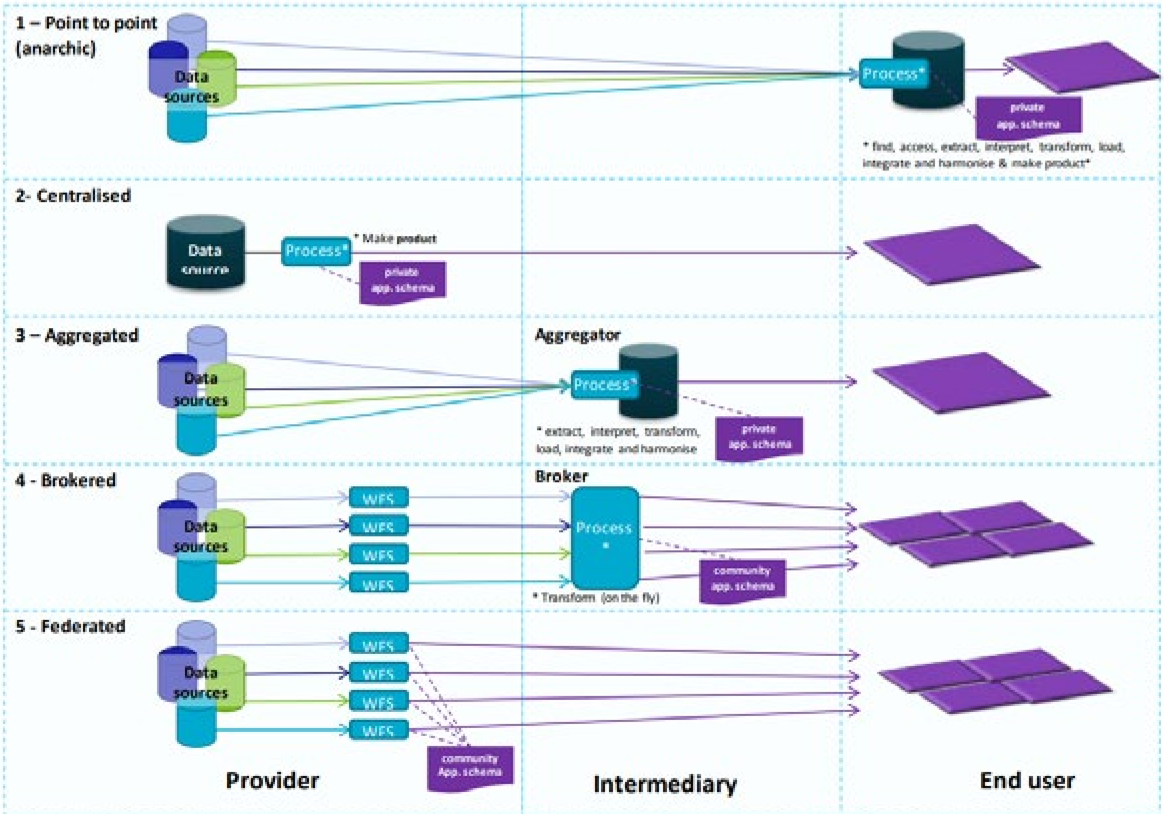


Figure 2: Geospatial Information Supply Chain Patterns. Source: FSDF Data Specifications Framework.

Successful implementation of a data specification framework will culminate in the development of standards-based data specifications such as the Inspire Data Specifications that can be promulgated within a specific community, or the community centred around a particular Geospatial Fundamental Data Theme. Enabling full interoperability between data of the same theme from different providers, as well as some level of interoperability between data across different themes.

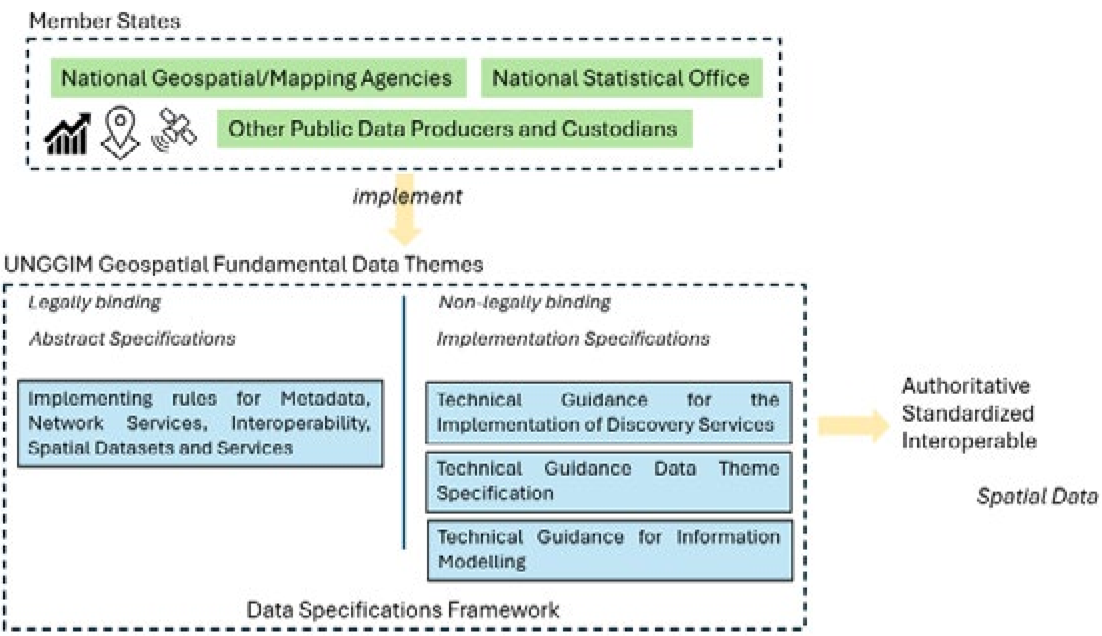


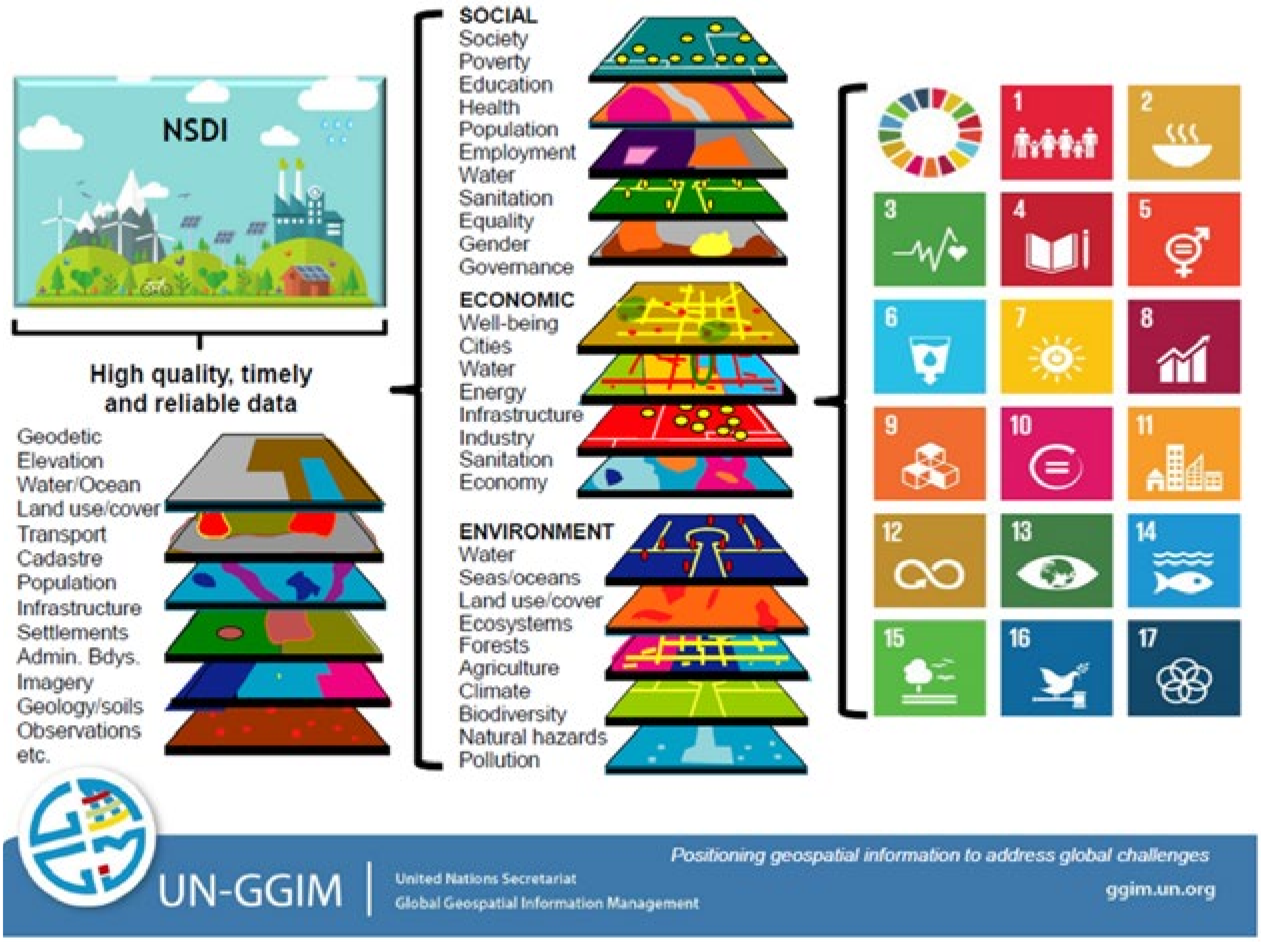
Figure 3: Implementing the Geospatial Fundamental Data Themes.

8.4 How each theme supports Sustainable Development Goals (SDGs)

At a high level, Geospatial Fundamental Data Themes, by serving as an authoritative source of geospatial data, support Sustainable Development Goals in the following manner:

- Baseline measurement
- Evidence-based policymaking
- Implementation of actions
- Monitoring and reporting

The specific correlation between each Geospatial Fundamental Data Theme and the Sustainable Development Goals is depicted in the table below. The depiction is based on Dr Greg Scott Presentation matrix correlating fundamental themes with SDG Goals to the UN-GGIM.



Geospatial Fundamental Data Themes	Sustainable Development Goals Supported
Geodetic Reference Frame	
Addresses	
Buildings and Settlements	

<table>
<tr><td>Elevation and Depth</td><td>
</td></tr>
<tr><td>Functional Areas</td><td>
</td></tr>
</table>

Geographical Names	
Geology and Soils	

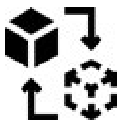
<table>
<tr><td>Land Cover and Land Use</td><td>
</td></tr>
<tr><td>Land Parcels</td><td>
</td></tr>
<tr><td>Physical Infrastructure</td><td>
</td></tr>
<tr><td>Population Distribution</td><td>
</td></tr>
</table>

	
Orthoimagery	
Transport Networks	

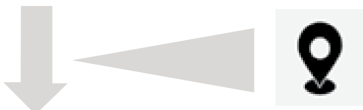
Transport Networks	

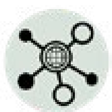
8.4 Role of Fundamental Data Themes in 3D Mapping / National Digital Twins

The infographic below introduces the three concepts of Digital Twins, National Digital Twins, and Geospatial Fundamental Data Themes in relation to one another.



- **Digital Twins** can be described as highly advanced digital representations of the real world.
- Represent real-world phenomenon or objects digitally, in the form of data and data products such as analytics or computer-based visualization.
- Towards perceiving the latter from a 360-degree perspective and being able to manipulate them through their digital avatars.
- Connecting with the physical 'twin' they model, changing alongside the physical system via real-time sensors and actuators.
- Advanced analytics processes the data to deliver insights and model future scenarios.





- A National Digital Twin is to consist of a federation of digital twins.
- An ecosystem of digital twins connected via securely shared data.
- This requires that we relate Digital Twins to one another and position them as they are in the real world.
- NDT would allow interdependencies across sectors to be understood in a way that organisation level or sector-based digital twins could not satisfy.

- Authoritative Spatial (location) data will underpin Digital Twins, providing the essential elements to position digital twins relative to each other to reflect the real world.
- Fundamental Geospatial Data Themes enable the creation, maintenance, and dissemination of Authoritative Spatial Data.
- Digital Twins will be integrated with accurate location information (e.g. satellite positioning data, elevation data, land parcel and property data (cadastre6)) overlayed with 3D reality mesh data to generate spatially enabled digital twins.
- A spatially enabled digital twin combines a digital twin with spatial and positioning information, covering a defined geographic space above and below ground.
- Achieving spatially enabled digital twins will require modernised 3D and 4D (temporal) spatial data, particularly land parcel and property (cadastral) data and land use data.

The Centre for Digital Built Britain (CDBB) proposes the concept of a National Digital Twin (NDT) to connect standalone digital twins. Envisaged to consist of 'federations' of digital twins; the NDT will not be a huge singular model of the entire built environment. Rather, it will be an ecosystem of digital twins connected via securely shared data. By doing so, the NDT would allow interdependencies across sectors to be understood in a way that organisation-level or sector-based digital twins could not satisfy. By enabling cross-sectoral integration, a National Digital Twin (NDT) would be capable of successfully representing real-world socio-economic dynamics digitally in the form of data, and thus enable their manipulation using data analytical tools, generating knowledge and information that empowers us to holistically understand, moderate and manage seemingly disparate real-world phenomenon by placing them within the common context of space and time. Facilitating a futuristic version of Spatial Digital Twins that enables the intelligent and increasingly automated generation of planned and unplanned knowledge horizontally across diverse use cases, using integrated, networked, and open data as a form of universal currency.

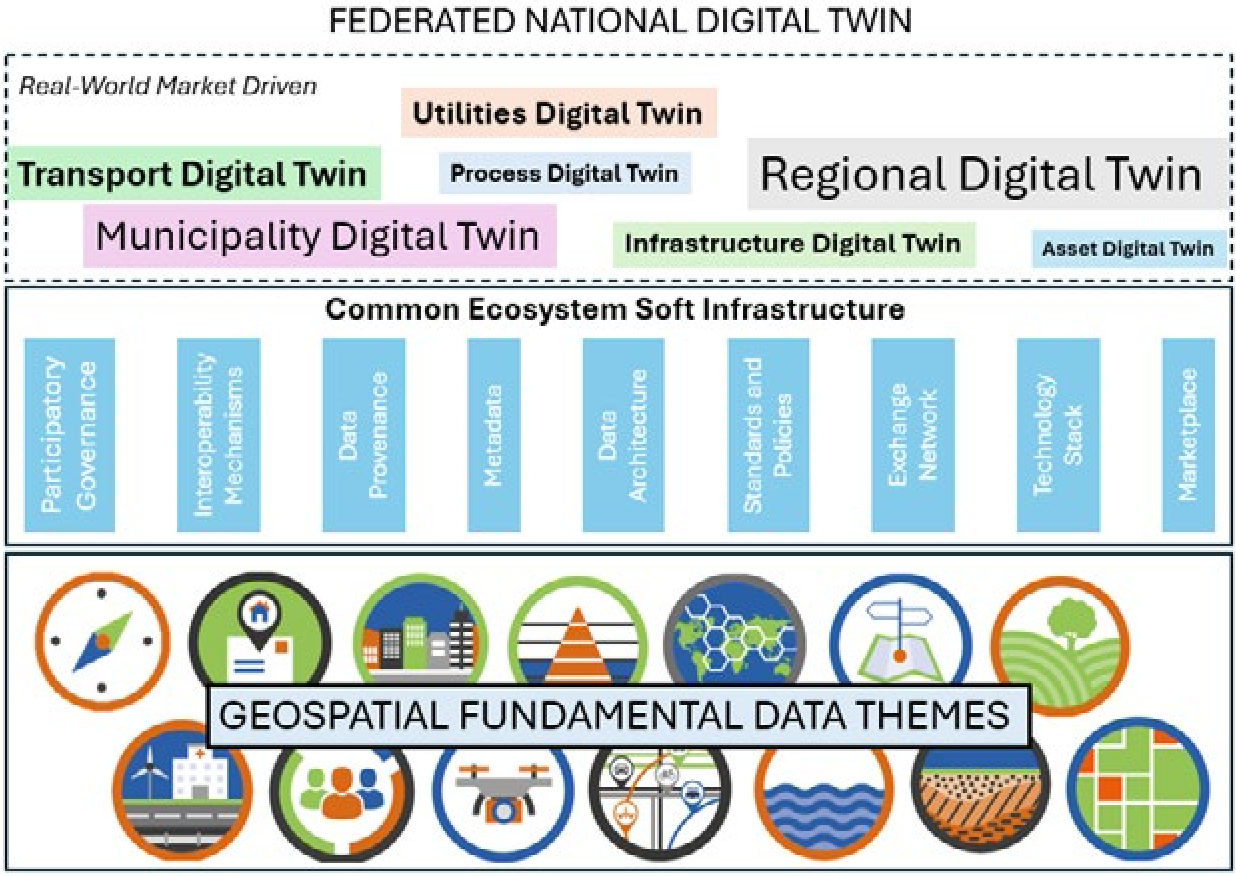


Figure 4: Federated National Digital Twin

A Federated National Digital Twin will be underpinned by authoritative spatial data as enabled by the Geospatial Fundamental Data Themes, followed by a suite of Soft Infrastructure elements, a concept proposed within the Common European Data Spaces. An intangible soft infrastructure that would comprise of the functional, legal, technical, and operational frameworks, which would enable the cross-sectoral exchange, integration, and use of geospatial data and information. This would enable the unrestricted development of the hard infrastructure, largely taking the form of end-use-case-oriented digital twin-like application instances. A development that will be driven by the open market, making use of the foundational soft infrastructure and the commoditization of geospatial data and technology that it facilitates. The figure below represents the concept of a Federated National Digital Twin underpinned by the Geospatial Fundamental Data Themes.

8.5 References

United Nations Committee of Experts on Global Geospatial Information Management, UNGGIM, The Global Fundamental Geospatial Data Themes. URL: https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/Fundamental_Data_Publication.pdf.

1. OPEN DEI, Design Principles for Data Spaces. URL: file:///C:/Users/Dell/OneDrive%20-%20Woolpert/Repository/Common%20European%20Data%20Spaces/Position-Paper-Design-Principles-for-Data-Spaces.pdf.
2. ANZLIC the Spatial Information Council, Principles for Spatially Enabled Digital Twins of the Built and Natural Environment in Australia. URL: https://www.anzlic.gov.au/sites/default/files/files/principles_for_spatially_enabled_digital_twins_of_the_built_and_natural_.pdf .
3. Box P, Simons B, Cox S, Maguire S, (2015) A Data Specification Framework for the Foundation Spatial Data Framework. CSIRO, Australia. CSIRO. URL: https://www.anzlic.gov.au/sites/default/files/files/FSDF-Data_Specification_Framework.pdf

LESSON – 9

Cloud Computing

Trainer's Biodata: Dr. Siva Ravada

Dr. Siva Ravada currently leads development of Oracle Spatial and Graph technologies at Oracle. And Oracle spatial database server is used across a wide range of applications: from traditional GIS/LIS type applications to business geographic applications providing location based services.

He regularly works with many national agencies (US Census, Ordnance Survey of UK, Dutch Cadastre, NGA, EPA, USDA, NOAA, Survey of India, etc.) to help them define the architecture for their spatial data systems and help them guide in their implementation of spatial data platform. He actively participates in the OGC standards process, reviewing and providing feedback on the early drafts of relevant standards. He also participates on the program committees of several international conference to serve the broader communities of spatial databases, Geographic Information Science and spatial computing.

He is also a member of the National Geo-Spatial Advisory Committee (NGAC). NGAC's charter is to advice the US government and its many stakeholders to help inform and improve policy and geospatial programs that are critical to achieving and maintaining access to useful geospatial information and resources through a national spatial data infrastructure.

9.1 Overview of cloud computing, including its definition, characteristics and service models

In simple terms, cloud computing allows you to rent instead of buying your IT infrastructure. Rather than investing heavily in software, databases, and hardware, companies opt to access their compute power via the internet, or the cloud, and pay for it as they use it. These cloud services now include, but are not limited to, servers, storage, databases, networking, software, analytics, and business intelligence.

Cloud computing provides the speed, scalability, and flexibility that enables businesses to develop, innovate, and support business IT solutions. When a company chooses to "move to the cloud," it means that its IT infrastructure is stored offsite, at a data centre that is maintained by the cloud computing provider. An industry-leading cloud provider has the responsibility for managing the customer's IT infrastructure, integrating applications, and developing new capabilities and functionality to keep pace with market demands.

For customers, cloud computing offers more agility, scale, and flexibility. Instead of spending money and resources on legacy IT systems, customers are able to focus on more strategic tasks. Without making a large upfront investment, they can quickly access the computing resources they need—and pay only for what they use.

Cloud computing benefits

There are several trends pushing businesses—across all industries—toward the cloud. For most organizations, the current way of doing business might not deliver the agility to grow or may not provide the platform or flexibility to compete. The explosion of data created by an increasing number of digital businesses is pushing the cost and complexity of data center storage to new levels—demanding new skills and analytics tools from IT.

Modern cloud solutions help companies meet the challenges of the digital age. Instead of managing their IT, organizations have the ability to respond quickly to a more fast-paced and complex business landscape. With modern cloud economics, the cloud delivers business value and reduces cost, helping enterprises achieve their full business potential with their cloud spend.

Cloud computing provides a superior alternative to traditional information technology, including these areas:

- Cost—eliminate capital expenses
- Speed—instantly provision space for development and testing
- Global scale—scale elastically
- Productivity—increased collaboration, predictable performance, and customer isolation
- Performance—better price/performance for cloud native workloads
- Reliability—fault-tolerant, scalable, distributed systems across all services

Cloud computing fosters innovation

Cloud customers benefit from automatically having the latest innovations and emerging technologies built into their IT systems, because the cloud provider takes on the work of developing new capabilities and features. It's about the speed of innovation. With the right cloud provider, customers can leverage a modern cloud computing architecture to innovate faster, increase productivity, and lower costs. Better yet, choosing a cloud provider that offers an integrated cloud (SaaS, PaaS, and IaaS) architecture gives businesses the ability to move from operations to innovation and deliver new apps and services, including the use of innovative technologies such as artificial intelligence (AI), chatbots, blockchain, and the Internet of Things (IoT). Companies can harness the abundance of data to gain predictive insights into their businesses and ultimately drive better outcomes for their customers.

Cloud security builds trust

Moving to the cloud removes the headaches and costs of maintaining IT security. An experienced cloud provider continually invests in the latest security technology—not only to respond to potential threats, but also to enable customers to better meet their regulatory requirements. The best cloud providers invest in every layer of cloud security as part of their overall design across global data center regions. Such a multilayer secure cloud approach offers security at the level the customer's business requires.

Cloud connects business processes

Business processes describe how work is done from beginning to end. They are a good way to describe how people are working together today, how they would like to work together ideally, and how their work will be shaped with the introduction of new cloud technology. With an integrated cloud solution, organizations are better equipped to manage and assess the costs and benefits of technology projects. For organizations that struggle with disconnected business processes and data silos, the cloud offers a way to transform their business operations. With the cloud, there's no need reinvent the process wheel. Complete cloud application suites are not modular but instead are connected, eliminating data silos, and enabling integration and intelligent business decisions. Finding a cloud provider that provides all three layers of the cloud—in connected and unprecedented ways—is essential. Moreover, businesses need a complete, integrated platform with intelligent solutions at every layer.

9.2 Cloud deployment models

There are three types of clouds: public, private, and hybrid. Each type requires a different level of management from the customer and provides a different level of security.

Public cloud

In a public cloud, the entire computing infrastructure is located on the premises of the cloud provider, and the provider delivers services to the customer over the internet. Customers do not have to maintain their own IT and can quickly add more users or computing power as needed. In this model, multiple tenants share the cloud provider's IT infrastructure.

Private cloud

A private cloud is used exclusively by one organization. It could be hosted at the organization's location or at the cloud provider's data centre. A private cloud provides the highest level of security and control.

Hybrid cloud

As the name suggests, a hybrid cloud is a combination of both public and private clouds. Generally, hybrid cloud customers host their business-critical applications on their own servers for more security and control, and store their secondary applications at the cloud provider's location.

Multicloud

The main difference between hybrid cloud and multicloud is the use of multiple cloud computing and storage devices in a single architecture.

There are three main types of cloud services: software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). There's no one-size-fits-all approach to cloud; it's more about finding the right solution to support your business requirements.

SaaS

SaaS is a software delivery model in which the cloud provider hosts the customer's applications at the cloud provider's location. The customer accesses those applications over the internet. Rather than paying for and maintaining their own computing infrastructure, SaaS customers take advantage of subscription to the service on a pay-as-you-go basis. Many businesses find SaaS to be the ideal solution because it enables them to get up and running quickly with the most innovative technology available. Automatic updates reduce the burden on in-house resources. Customers can scale services to support fluctuating workloads, adding more services or features they grow. A modern cloud suite provides complete software for every business need, including customer experience, customer relationship management, customer service, enterprise resource planning, procurement, financial management, human capital management, supply chain management, and more.

PaaS

PaaS gives customers the advantage of accessing the developer tools they need to build and manage mobile and web applications without investing in—or maintaining—the underlying infrastructure. The provider hosts the infrastructure and middleware components, and the customer accesses those services via a web browser. To aid productivity, PaaS solutions need to have ready-to-use programming components that allow developers to build new capabilities into their applications, including innovative technologies such as artificial intelligence (AI), chatbots, blockchain, and the Internet of Things (IoT). The right PaaS offering also should include solutions for analysts, end users, and professional IT administrators, including big data analytics, content management, database management, systems management, and security.

IaaS

IaaS enables customers to access infrastructure services on an on-demand basis via the internet. The key advantage is that the cloud provider hosts the infrastructure components that provide compute, storage, and network capacity so that subscribers can run their workloads in the cloud. The cloud subscriber is usually responsible for installing, configuring, securing, and maintaining any software on the cloud native solutions, such as database, middleware, and application software.

9.3 Cloud storage and data management

Data Management, Defined

Data management is the practice of collecting, keeping, and using data securely, efficiently, and cost-effectively. The goal of data management is to help people, organizations, and connected things optimize the use of data within the bounds of policy and regulation so that they can make decisions and take actions that maximize the benefit to the organization. A robust data management strategy is becoming more important than ever as organizations increasingly rely on intangible assets to create value.

Managing digital data in an organization involves a broad range of tasks, policies, procedures, and practices. The work of data management has a wide scope, covering factors such as how to:

- Create, access, and update data across a diverse data tier
- Store data across multiple clouds and on premises
- Provide high availability and disaster recovery
- Use data in a growing variety of apps, analytics, and algorithms
- Ensure data privacy and security
- Archive and destroy data in accordance with retention schedules and compliance requirements

A formal data management strategy addresses the activity of users and administrators, the capabilities of data management technologies, the demands of regulatory requirements, and the needs of the organization to obtain value from its data.

A data management platform is the foundational system for collecting and analysing large volumes of data across an organization. Commercial data platforms typically include software tools for management, developed by the database vendor or by third-party vendors. These data management solutions help IT teams and DBAs perform typical tasks such as:

- Identifying, alerting, diagnosing, and resolving faults in the database system or underlying infrastructure
- Allocating database memory and storage resources
- Making changes in the database design
- Optimizing responses to database queries for faster application performance

Cloud Storage

The term **cloud storage** encompasses several storage capabilities available to cloud customers that run on a cloud provider's hardware. Each of these capabilities meets a different need, but all provide the flexibility to pay only for what you use. The provider is responsible for maintaining the underlying hardware and ensuring that data remains available, resilient, and protected. The most common types of cloud storage are object, file, and block.

Types of cloud storage

Storage types differ primarily in how they are accessed and the level of performance they provide. The application using the storage and its location determine storage requirements.

Object storage

Object storage is accessed differently than the other storage types used in on-premise solutions. Software applications must be intentionally written to use object storage by calling web APIs. Object storage is maintained remotely from the application and is used in two similar but different situations. First, it's often accessed via the internet by applications running on individual computers, mobile devices, and Internet of Things devices, among others. Second, it can be used by applications running in the cloud.

Applications that use object storage can store and retrieve unstructured data from object storage in a remote location without using a file system. The stored items are merely abstract "objects" in the cloud. This means the application developer maintains maximum flexibility and has an essentially bottomless, free-form datastore in the cloud while being charged only for the amount of data stored and transferred.

The downsides of object storage are twofold: It involves a bit more work for the application authors who must manage their own object formats, and there are performance limitations. Object storage is accessed by software making API calls, typically over the internet, so what might take direct-attached storage microseconds and block storage or file storage milliseconds may sometimes take object storage a second or more. For many use cases, such as end users running applications connected to cloud storage on their phones, this performance is acceptable, especially in return for the "anywhere access" these applications provide. And in cases where an application using object storage is running in the same cloud as the objects are stored, performance is considerably higher because all the resources are in the same region on the cloud provider's own local network.

Archive Storage

Archive storage is the ideal solution for storing seldom-accessed data that requires long retention periods. Archive storage is more cost-effective than object storage for preserving cold data. However, unlike other storage options, archive storage data retrieval is not immediate.

Both object storage and archive storage use buckets as logical containers for storing objects. A bucket is a single compartment with policies that determine the actions that can be performed on objects in the bucket—and who can perform them.

When buckets are created to hold data as objects, organizations can decide which default storage tier—archive or standard—is appropriate for their data. Object storage can also automatically move objects to archive storage.

Cloud file storage

Most traditional applications that run on a physical server and leverage physical drives in your data centre use file storage. Operating systems such as Linux or Microsoft Windows Server present the applications that run on them with a file system—a single consistent set of rules and methods for storing and retrieving data. The operating system handles the details behind the scenes—for example, is the physical disk a solid-state drive (SSD)? A traditional spinning disk hard drive? An optical disk? Or a remote network file share? While the operating system takes care of these details, applications simply open, read from, write to, and save files with standardized API calls.

Cloud file storage presents a standard network file share—similar to the network file shares that might run in your own data centre—to the operating systems you're running on servers in the cloud. Those operating systems present that file system to applications running on that virtual machine. Applications don't need to be modified or changed to run in the cloud; they continue to run with the file storage they've always used.

The cloud provider manages the hardware, including physical disks and network hardware, and ensures the data is protected. Additional incremental capacity is available to customers as their needs grow. The inherent advantages of cloud file storage are clear when compared to a traditional approach that involves scheduled bulk purchases of network file systems to meet the needs of future growth and then requires you to manage your hardware and ensure the protection of your data yourself.

Block storage

Block volumes are like cloud file storage in that they represent an enhanced version of a type of network storage you may already be running in your data centre. Using block volumes results in less network overhead and offers higher performance, but it requires more configuration and management within the operating systems in return.

Block volumes typically use a single volume type, which can be configured on the fly with different settings to increase performance or reduce costs. Unlike cloud file storage, block volumes must be configured with a specific size, but that size can be increased at any time while the volume remains online and available to the applications using it.

As with any cloud service, the provider manages the hardware and capacity planning and ensures the data is replicated and protected.

9.4 Cloud security

IT and security leaders are faced with security challenges that continue to evolve as cyberattacks increase and IT infrastructure grows more complex. Multi-cloud adoption, remote work, digital customer engagement, and the growing amount of data and devices have left organizations with expanding attack surfaces and an excess of tools that can create costly inefficiencies and unnecessary risk. As organizations evaluate how to address these challenges, they are turning to the cloud for a simpler and more efficient way to strengthen security posture. For many organizations, cloud infrastructure can provide a more secure platform than an on-premises data centre and help simplify cybersecurity with services that automate system security tasks and help reduce risk.

Unified identity and access management

As cloud footprints expand to include more applications and services, new silos can emerge. Because apps and services each have their own provisioning mechanisms and systems for managing identities, new adoption can create opportunities for inconsistencies in how access and governance policies are applied. Without a centralized view of security across the organization, this "identity sprawl" can result in overprivileged accounts going unnoticed. Identity sprawl can make rapid user provisioning or deprovisioning more challenging, leading to inconsistent entitlements and ghost accounts that can increase the risk of data loss and compromise. It can also result in inconsistent user experiences. According to Gartner, the top method for breaches involves misused credentials, which indicates that organizations without a strategy to address identity sprawl could increase their risk of a breach.

Identity and Access Management (IAM) is at the centre of a company's interaction with its users and devices, acting as the "front door" for an organization's data and applications. But the expanded use of identity-driven policies across cloud and on-premises environments makes it difficult to manage identities and achieve end-to-end governance. Organizations are turning to unified IAM platforms that position identity as the security control for expanding IT architectures. These platforms offer a centrally managed approach to security and help organizations prevent identity sprawl by managing entitlement across cloud and on-premises applications. Cloud based IAM platform provides the following advantages:

- Address common use cases for human and nonhuman identities with IAM solutions and capabilities that include adaptive multifactor authentication, access management, single signon, and identity lifecycle management.
- Integrate identities and systems to secure access from anywhere, at any time, and by any method by delivering risk aware, end-to-end user authentication and single sign-on.
- Enable a unified approach to identity and access management with cloud-based workflows that help address identity sprawl by consolidating enterprise wide identity silos into a unified platform.
- Simplify lifecycle management and reduce the risk of a compromise with IAM services that improve visibility across the organization and provide advanced authentication.
- Improve efficiency with a centralized view for managing policies and replace manual provisioning processes with a single manageable workflow to easily onboard new applications, users, and devices faster and with greater consistency.
- Use your existing entitlements from traditional software delivered IAM solutions or core business applications such as ERP or human resources management systems to quickly onboard new applications, users, and devices.

9.5 Cloud based GIS applications

Many organizations have historically managed geospatial data in specialized proprietary data stores and data structures, outside of their central enterprise data stores and accessible only through a specific toolset. With cloud based GIS, geospatial data like geocoded addresses and region shapes are 1st class data elements, accessed with standard APIs to perform native spatial operations such as proximity search and distance calculations.

Traditional GIS tools and applications suffer from the "data silos" that make it hard to share data across different departments and users. Cloud based GIS moves away from this problem as all the data is natively available on the same data platform. By making it easy to define fine grained access controls in the cloud systems, data sharing across users and departments can be achieved by defining access controls for each layer of data.

Cloud GIS also provides low code data preparation and application development environments that support natively managed geospatial data. The ecosystem of specialized software (i.e., GIS software and developer toolkits) supported on cloud GIS platform provides the flexibility to mix and match other tools to best fit your organization's needs. By maintaining geospatial data natively in cloud GIS, a wide array of analytics and data warehousing, transaction processing, and mixed workload use cases are enabled, and efficiencies are gained while preserving existing investments in other tools and development.

Cloud GIS typically provides the following components:

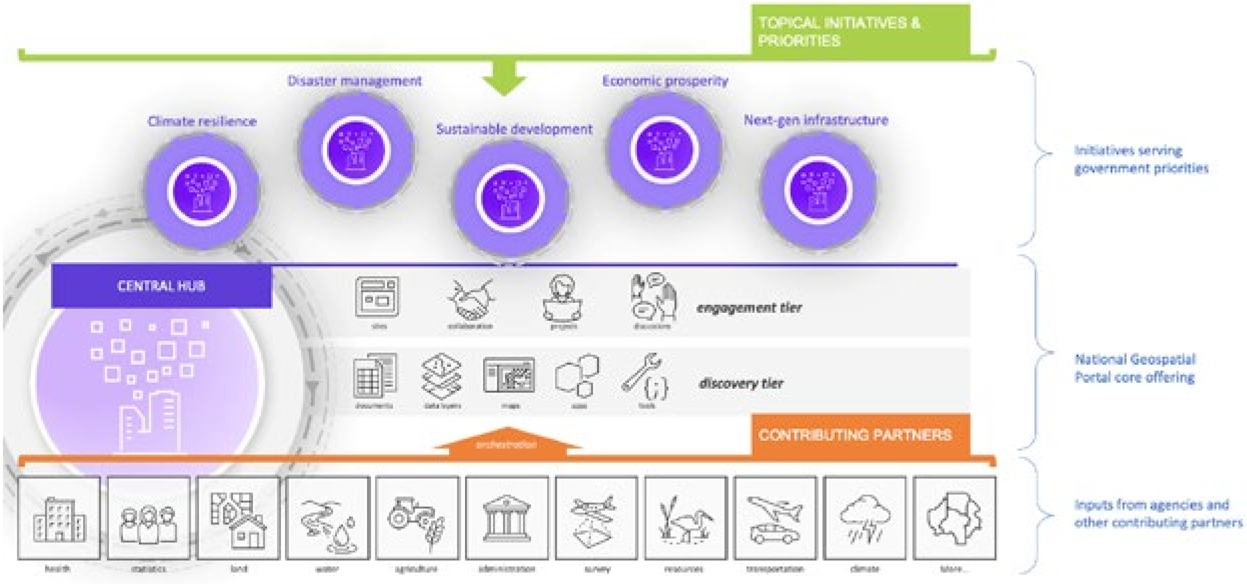


Figure 1

- Data Refinery

    Ingests and refines data for use in each of the data layers in the architecture. The shape is intended to illustrate differences in processing costs for storing and refining data at each level and for moving data between them.

    For advanced use cases, cloud GIS provides native cloud service that helps customers with common extract, load, and transform (ETL) tasks such as ingesting data from different sources, cleansing, transforming, and reshaping that data, and then efficiently loading it to target data sources on some form of cloud storage.

- Data Persistence Platform (curated information layer)

    Facilitates access and navigation of data to show the current business view.

    > Curated information: For relational technologies, data may be logically derived or physically structured in simple relational, longitudinal, dimensional or OLAP forms. For non-relational data, this layer contains one or more pools of data, either output from an analytical process or data optimized for a specific analytical task.
    > Raw information: Stored, granular data used as input to produce curated data. Data is stored in the format and schema derived from the source.

- Access and Interpretation

    Abstracts the logical business view of the data for consumers. This abstraction facilitates agile approaches for development, migration to the target architecture, and the provision of a single reporting layer from multiple federated sources.

9.6 Collaboration and sharing

Today's digital businesses must move beyond simple "file sync and share" solutions as they empower the workforce with 24/7 access not only to business content, but also applications and processes, to drive meaningful collaboration and mobility. Business managers need cloud-based services that can seamlessly blend content, people, processes, and communications as part of a cohesive workflow. Sharing documents is important, but to be truly valuable the content must be contextually relevant to the task at hand.

As organizations adopt software-as-a-service (SaaS) applications to solve business challenges, fulfilling this collaborative vision requires a new generation of connected online tools. Cloud providers often provide comprehensive content sharing and collaboration services that enables users to easily and securely share content, effectively collaborate using social conversations, and access work content anytime, anywhere via web, desktop, or mobile interfaces.

Compared to the traditional model of the sharing data or resources, cloud platforms offer secure and easy to manage sharing models that allow users to share not just files, but any data resource (including data sets, maps, and applications). Cloud platforms provide the unique ability to generate a URL for any content a user wants to share with other users. This can be a data set like an admin boundary data set, or the result of an analysis in the form of a GeoJSON data set. Once the URL is generated, it can be shared with any other user and the receiver just needs a web browser to look at that data. The URL itself can be made public or can be associated with authentication. This gives the data owner the fine grained access controls and the access can be revoked anytime by the owner of the content.

9.7 Cloud based AI and machine learning

AI has become a catchall term for applications that perform complex tasks that once required human input, such as communicating with customers online or playing chess. The term is often used interchangeably with its subfields, which include machine learning (ML) and deep learning.

There are differences, however. For example, machine learning is focused on building systems that learn or improve their performance based on the data they consume. It's important to note that although all machine learning is AI, not all AI is machine learning.

Three factors are driving the development of AI across industries.

- **Affordable, high-performance computing capability is readily available.** The abundance of commodity compute power in the cloud enables easy access to affordable, high-performance computing power. Before this development, the only computing environments available for AI were non-cloud-based and cost prohibitive.
- **Large volumes of data are available for training.** AI needs to be trained on lots of data to make the right predictions. Ease of data labeling and affordable storage and processing of structured and unstructured data is enabling more algorithm building and training.
- **Applied AI delivers a competitive advantage.** Enterprises are increasingly recognizing the competitive advantage of applying AI insights to business objectives and are making it a businesswide priority. For example, targeted recommendations provided by AI can help businesses make better decisions faster.
- Many of the features and capabilities of AI can lead to lower costs, reduced risks, faster time to market, and much more.

In ML models, large sets of data are used to train the algorithm. This data could be structured, unstructured, or raw and needs powerful CPUs and GPUs to process. Only an ideal combination of public, private, or hybrid cloud systems (based on security and compliance requirements) can provide such huge amounts of compute power today. Further, AI cloud computing also enables services that are used in ML, such as serverless computing, batch processing, and container orchestration.

By leveraging the cloud, businesses can access a virtually limitless pool of resources to train AI models, analyse vast datasets, and extract actionable insights. The cloud's scalability ensures that AI applications can handle varying workloads, from small-scale experimentation to large-scale production deployments, all without incurring exorbitant infrastructure costs.

Cloud providers also offer services that abstract complex infrastructure management from valuable services such as machine learning and AI modelling. Cloud-based AI is built on a pay-as-you-go model, which allows businesses to optimize resource utilization and scale up or down as needed. Replicating this sort of capacity management on-premises would be a monumental task with limited returns on investment, rendering AI's effect on organizations nothing short of transformative and definitive.

With public cloud services, developers do not need to build and manage a separate infrastructure for hosting AI platforms. They can use ready configurations and models to test and deploy AI applications.

Further, generic services based on AI but not necessarily requiring an ML model – such as speech-to-text, analytics, and visualization – can be improved by running them from the cloud using first-party data generated by the organization.

Cost Savings

Traditionally ML-based models ran on expensive machines with multiple GPUs in enterprise datacentres. With advances in virtualization in both public and private clouds, the cost of building, testing, and deploying these models has come down drastically. This has levelled the playing field for many small-to-medium businesses.

AI-based algorithms required significant admin time and effort in terms of building testing and production environments, software management, and provisioning hardware resources for compute operations and storage. A centrally managed hybrid cloud or a public cloud does away with this, leaving IT staff to focus on non-repetitive tasks.

9.8 Cloud computing strategy and governance

Governance is a set of well-planned rules and processes that you create, monitor, and amend as necessary. It enables organizations to control costs, improve efficiency, and eliminate security risks.

As organizations move to the cloud and support multi-cloud environments, the network is no longer the perimeter. Organizations need a comprehensive plan to onboard and implement security controls including protective and detective controls. The governance model provides a structure to design and improve cloud security, ensure compliance, and reduce costs.

Using a governance model enables organizations to define a structured process for enforcing security and compliance that can be mirrored, modified, or changed based on your organizational requirements.

Creating an effective governance model includes decisions around the provisioning, administration, stewardship, and offboarding of cloud resources. You can use a manual process to configure the workflow. As the governance model matures, you can configure automated workflows and infrastructure-as-code Terraform stack to deploy the governance model.

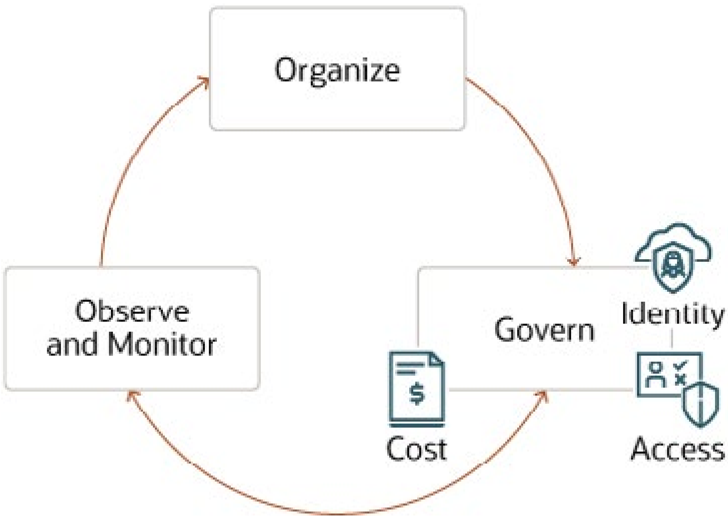


Figure 2

The process to implement a governance model is cyclic and iterative. You begin with organizing resources followed by implementing governance that includes identity, access, and cost governance. Finally, you set up processes to observe and monitor security controls. As you detect issues with the model, you feed the data back into the model to improve it.

Organize

You organize your resources and workloads to set up the foundation of an effective governance model.

Govern

You govern your resources and workloads by managing who gets access to which resources. You can manage access to resources using an identity and access provisioning system.

Observe and Monitor

You continuously observe and monitor your resources to ensure the governance model is working to meet your organizational goals. You will require certain essential services such as Cloud Guard, and Logging and Audit services to observe and detect any issues that need to be fixed.

Establishing a governance model provides your organization with a starting point and helps identify improvements needed. It also helps identify the consequences and the risks that will be associated with the model. You apply the governance model at various stages of your governance implementation.

To enforce cloud governance, security architects must do the following:

- Define and enforce access to resources such as compartments
- Define Identity and Access Management (IAM) policies and tagging
- Observe, monitor, and detect issues using Log Analytics servcies
- Manage costs by defining service limits, budgets, and usage quotas
- Monitor performance to control costs and prevent costs from going over budget
- Use the Audit service to ensure compliance to defined governance policies

9.9 References

1. https://www.oracle.com/cloud/what-is-cloud-computing/
2. https://www.ibm.com/topics/cloud-computing
3. https://www.citrix.com/glossary/what-is-a-cloud-service.html
4. https://www.nutanix.com/info/what-is-hybrid-cloud
5. https://docs.oracle.com/en/solutions/foundational-oci-governance-model/index.html
6. https://gis.usc.edu/blog/gis-and-the-cloud-how-they-work-together/
7. https://www.techtarget.com/searchstorage/definition/cloud-data-management
8. https://www.oracle.com/database/what-is-data-management/
9. https://www.oracle.com/security/cloud-security/

LESSON – 9

Geospatial Infrastructure for AEC Industry

Trainer's Biodata: Prof. Dr. John P. Wilson

Dr. John P. Wilson is a Professor of Spatial Sciences, Architecture, Civil and Environmental Engineering, Computer Science, Population and Public Health Sciences, and Sociology and founding director of the Spatial Sciences Institute at the University of Southern California (USC). He earned a PhD in Geography from the University of Toronto in 1986 and he was elected a fellow in the University Consortium for Geographic Information Science (UCGIS) and the American Association of Geographers in 2014 and 2021, respectively. He teaches courses in hydrology, geodesign, geographic information science, geospatial leadership, spatial data acquisition, and population, health and place. His research focuses on the modeling of coupled human and environmental systems and makes extensive use of GIS, spatial analysis and computer models, and to date, he has published four books (Terrain Analysis: Principles and Applications, 2000; The Handbook of Geographic Information Science, 2008; Environmental Applications of Digital Terrain Modeling, 2018; Spatial Data Science, forthcoming) and 130 book chapters and refereed journal articles on these topics. He also serves as the founding editor of Transactions in GIS and the editor of the UCGIS's Geographic Information Science and Technology Body of Knowledge project.

10.1 Introduction

The infrastructure industry is the backbone of a country's economic goals. It is a defining parameter based on which other sectors can grow, by building and constructing the physical facilities for the distribution and manufacturing of all of the goods and services that are required to support humanity in the twenty-first century. This sector, in general, is very competitive, and the prime focus of the contractors is on enhancing productivity and efficiency to ensure minimum time and cost overruns.

9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

Sustainable Development Goal 9, "Industry, Innovation and Infrastructure", Target 9.1, by 2030 develop quality, reliable, sustainable, and resilient infrastructure to support economic development and human wellbeing. Geospatial knowledge services can help understand patterns of road usage, pockets of congestion, and determine mobility patterns of population, key factors in the development of infrastructure.

Sustainable Development Goal 11, "Sustainable Cities and Communities", Target 11.2, by 2030 provide access to safe, affordable, accessible and sustainable transport systems for all. Contextual data and knowledge regarding the building and transport infrastructure can impart improved road safety and basis for enhanced planning and management of transport systems.

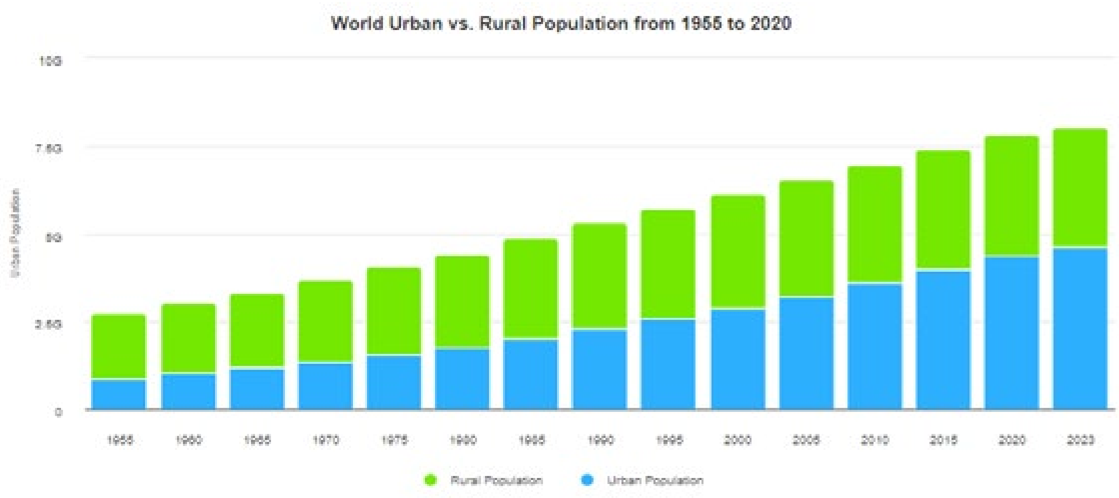
The remainder of this session is organized around three elements:

- The role of infrastructure in the architecture, engineering, construction (AEC) sector, the economy and everyday life.
- The role of the geospatial knowledge infrastructure (GKI) and technology in the AEC sector.
- Five aspirational case studies illustrating the role of the GKI in today's AEC workflows.

10.2 Role of Infrastructure in the AEC Sector, the Economy, and Everyday Life

Building and Infrastructure across the world has undergone a significant transformation over the last 30 years. Countries that had often ignored the relationship of economic growth, resiliency and population growth with infrastructure development, are now realizing that infrastructure is a key engine of growth and the most tangible hallmark of economic development and progress.

There are two complementary sets of trends and ideas driving this transformation. The first is the continuously increasing population and shift from rural of urban life. There was just less than 3 billion people when I was born and nearly two-thirds were rural residents which are very different than today's numbers which point to nearly 8 billion people and six out of every 10 residents residing in cities (Figure 1).



Source: Worldometer

Figure 1: The world's urban population over time

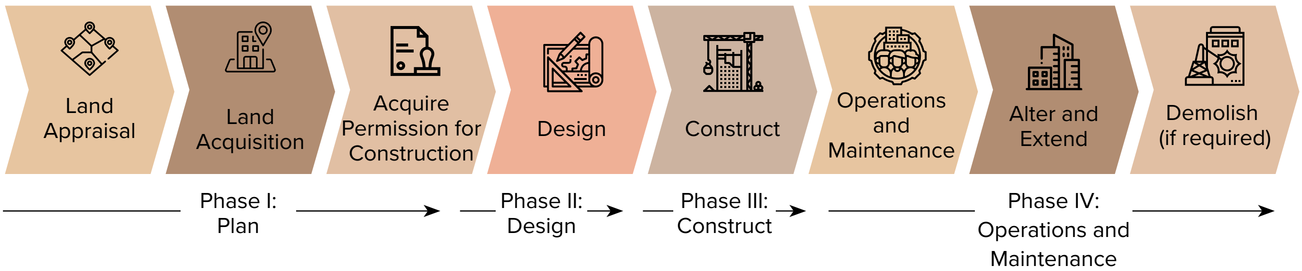
The second set of drivers are organized around the concept of a circular economy that requires fewer raw materials and produces less waste and fewer emissions. The transition to a circular economy will likely slow global warming, protect biodiversity, assure freshwater supplies, and accelerate the path to sustainable and resilient communities (Figure 2).

Source: European Parliament Research Service

Figure 2: The circular economy model

10.2.1 Phases of Infrastructure Lifecycle

The infrastructure industry broadly goes through four life cycle phases — defined as planning, design, construction, and operations and maintenance — for overall project management (Figure 3). Proper coordination across the construction lifecycle ensures cost savings and reduced time overruns, leading to reduced wastage.



Source: Geospatial World Analysis

Figure 3: The four phases that define the infrastructure sector workflow

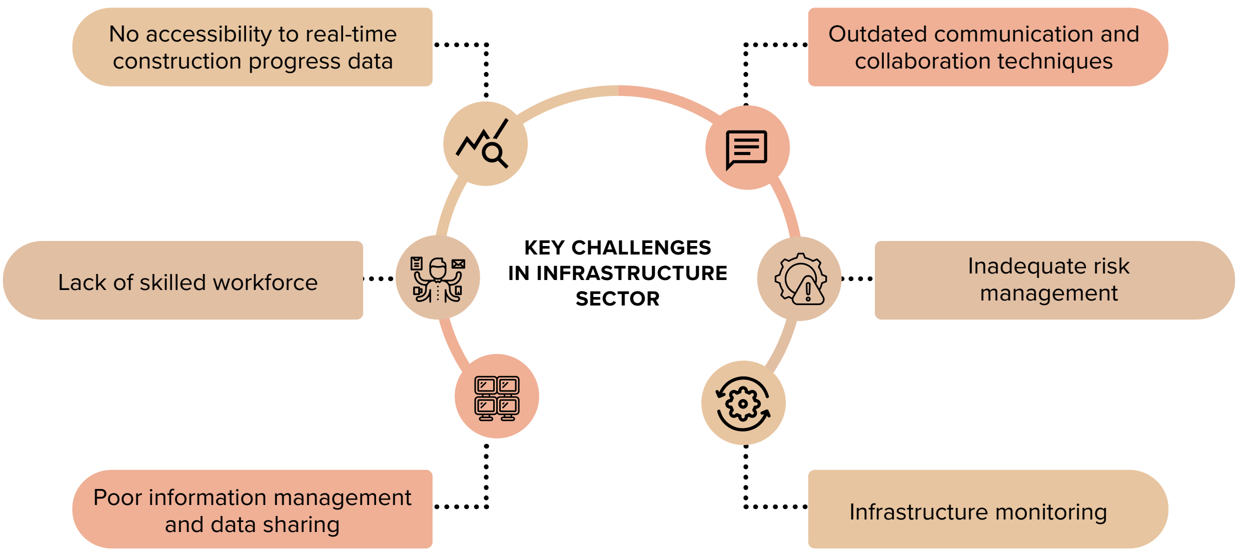
The first planning phase starts with land appraisal and land acquisition. The project details are identified and once the strategic plans are laid out for the commencement of the project, a feasibility study is carried out to gain understanding regarding its practicality and to prepare for possible hurdles and scenarios. Following the feasibility study, permission is granted for construction and major deliverables with timelines are identified. A detailed project report is prepared, outlining the project activities, dependencies and timeframes, after which the design work for the project begins.

The designing of the project is initiated in the second phase with the help of project scheduling and 3D modeling software using CAD/BIM and GIS solutions in a Common Data Environment (CDE). The design model, when approved, is sent for construction in the third phase, wherein material procurement and machinery allocation is carried out in an efficient manner to avoid cost overruns

Following construction, the project enters the fourth and final phase of its lifecycle, which is operations and maintenance. During this phase, the constructed building is monitored for completion, minor defects, (non-) compliance, and other services. At this stage, a decision might be taken to alter or extend the project as decided by the project owner and at a later date, a decision might be taken to demolish the infrastructure associated with the project at hand altogether.

10.2.2 Challenges in the AEC Sector

The infrastructure industry has been growing for centuries but has also faced many challenges during this time. Some of these challenges are relatively new, but others are as old as the industry itself. A few of the key challenges faced by the infrastructure sector are summarized in Figure 4 below.



Source: Geospatial World Analysis

Figure 4: Key challenges in the construction and infrastructure sector

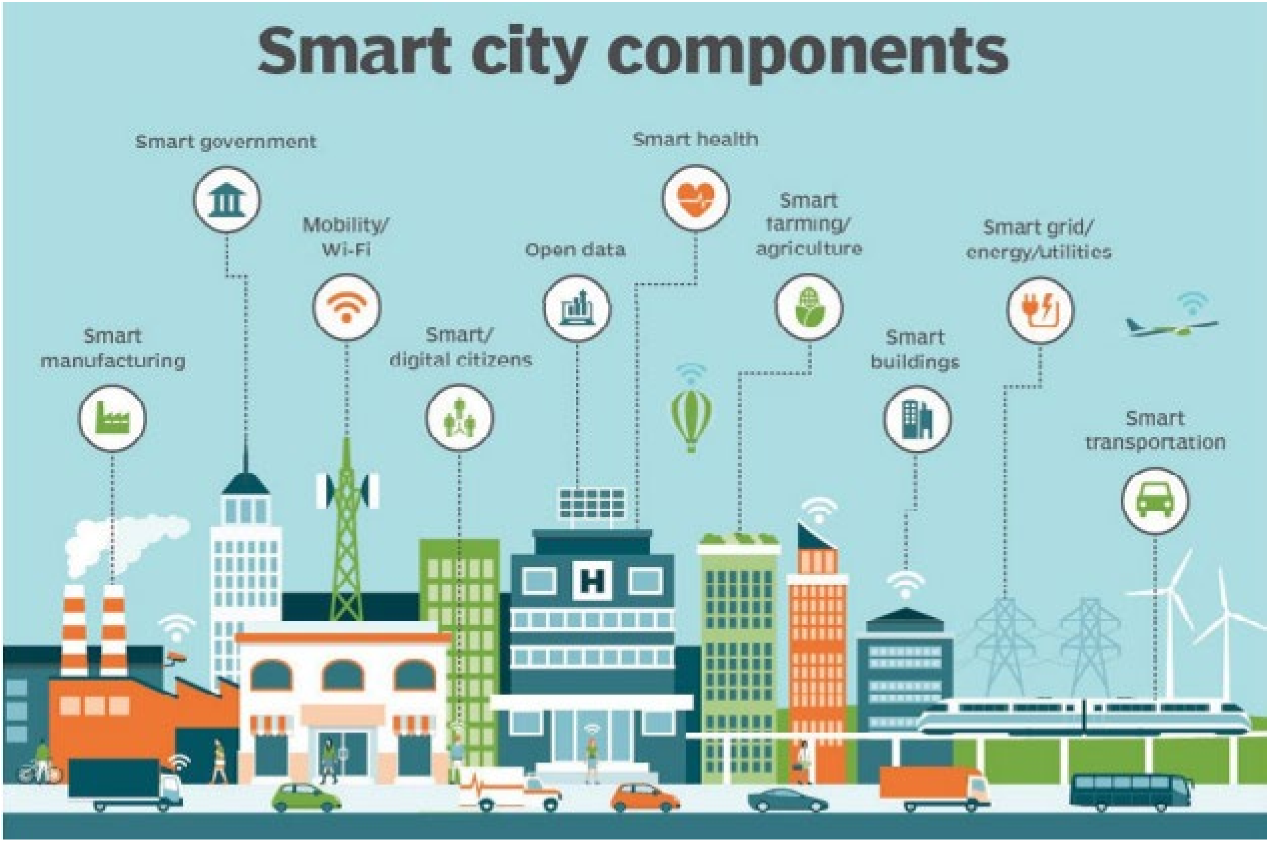
The infrastructure takes many different forms, including satellite systems that have dual civilian and military use, buildings and overhead utilities, parks and other public spaces, and the television or communication cables, gas pipes, metal and concrete pipes, sewer lines, water pipes, and electric lines often located under the Earth's surface.

Rising population and wealth coupled with our growing adoption and reliance on computation and digital data offer new needs and opportunities to deploy sophisticated systems to support life and protect the planet, as encapsulated in the various visions that have been offered for smart cities. Figure 5 shows the 10 components that Shea (2020) used to describe smart cities.

Reading this graphic from left to right, this vision starts with smart manufacturing, which provides a technology-driven approach that utilizes Internet-connected machinery to monitor the production process. The goal of smart manufacturing is to identify opportunities for automating operations and use data analytics to improve manufacturing performance.

The second component, smart government, uses technology and innovation to facilitate and support enhanced decision-making and planning within governing bodies.

The third component, which supports mobility with Wifi, relies on a portable hotspot device or a mobile hotspot on your phone that lets you share your high-speed data to connect more devices to the internet in more places — all powered by a nationwide network.



Source: Shea (2020)
Figure 5: Smart city components

The fourth component, smart and digital citizens, supports digital citizenship which, in turn, entails the responsible, ethical and respectful use of technology to engage online, find reliable sources, and protect and promote human rights. It teaches skills to communicate, collaborate, and act positively on any online platform.

The fifth component, open data, covers data that is available for everyone to access, use and share. It is generally published by governments on freely accessible portals and might include information about local areas, or statistics on topics such as the economy, health, and the environment.

The sixth component, smart health, imagines what a hospital and health system can be in the digital health age and then builds it using a blend of physical and virtual care, connecting patients via sensors and smartphones to clinicians, connected buildings and their health data.

The seventh component, smart farming and agriculture, is about using the new technologies which have arisen at the dawn of the Fourth Industrial Revolution in the areas of agriculture and aquaculture to increase production quantity and quality, by making maximum use of resources and minimising the environmental impact.

The eighth component, smart buildings, focuses on buildings which communicate using subsystems – such as heating and cooling, energy, lighting, plumbing, access control, and security – that can interact with one another through a network and can also be controlled remotely. An intelligent building in the same way gives building owners the flexibility to integrate individual building systems to deliver an enhanced working or living environment for occupants, and optimal equipment functionality and better management of resources for owners.

The ninth component, smart grids for energy and utilities, is an enhancement of the twentieth century electrical grid, using two-way communications and distributed intelligent devices. They use connected sensors across their grids in order to deliver services more efficiently and to analyze their operations.

The tenth and final component, smart transportation, includes the basic management systems we are already accustomed to, such as car navigation that, of course, uses satellite navigation to get position data which is then correlated to a vehicle's position on a road and when directions are needed, calculates routes from the current position to the desired destination.

10.3 The Role of the GKI and Technology in the AEC Sector

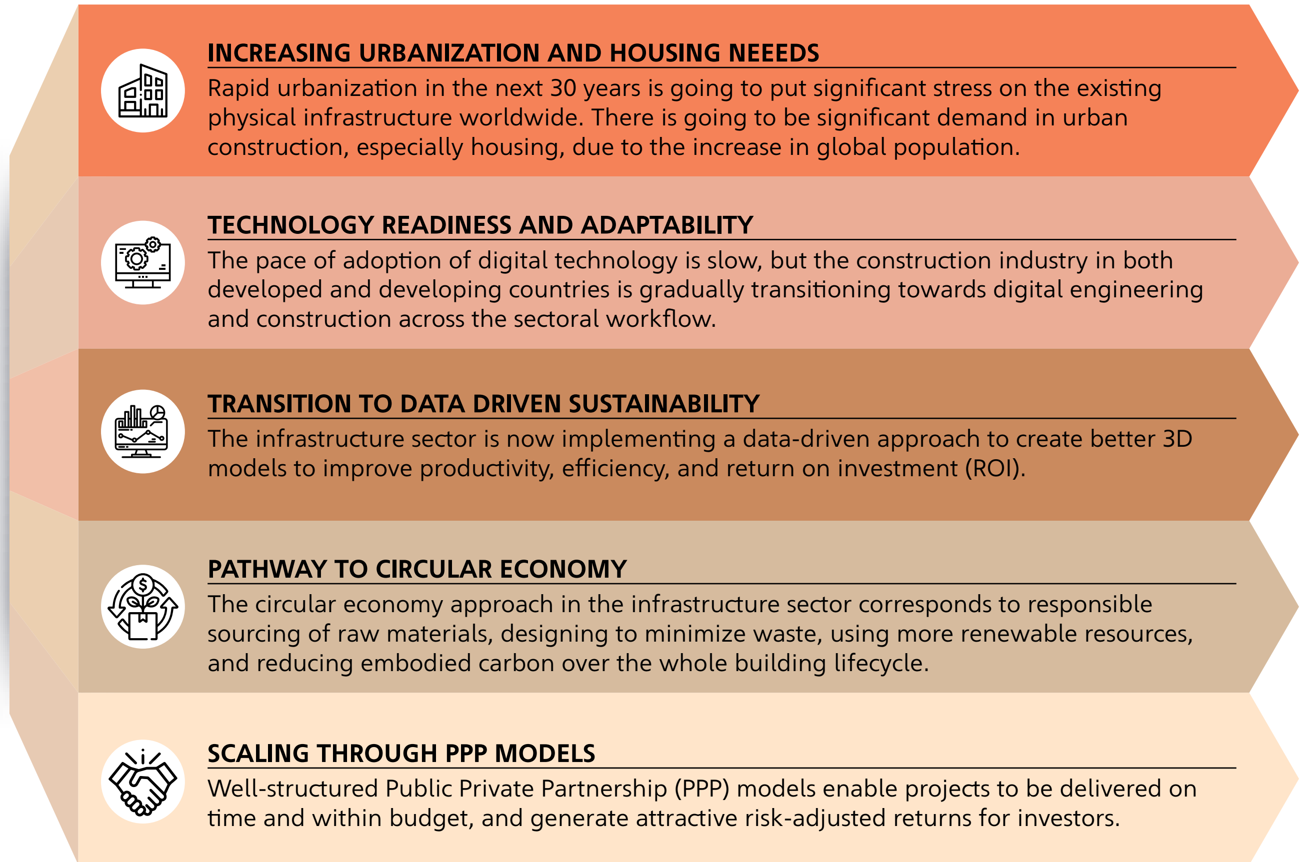
Technologies elevate the aforementioned construction and infrastructure sector workflows with data for decision-making, resource management, waste reduction, and scheduling and project management. The majority of construction organizations, including small and medium-size enterprise (SME) organizations, have yet to adopt enterprise-level applications of digital technologies and the GKI elements described in Figure 6 below.

GKI ELEMENTS	INTEGRATION IN INFRASTRUCTURE SECTOR WORKFLOW
Integrated Policy Framework	• Making the industry more innovative, efficient and sustainable through implementation of integrated policies. • Supporting the digitalization of the infrastructure sector through the implementation of policies and instruments meant to incentivize the uptake of digital technologies, practices and business models. • Strengthen the security and privacy framework to ensure secure and trustworthy data and knowledge exchange within the built environment
FOUNDATION DATA	• Allows more precise and complete designs, which results in safer and more efficient execution of construction activities, and also the efficient operation and maintenance of the overall project. • Provides greater control over the project delivery process, prevents duplication, and allows complementary data from different sources.
PARTNERSHIPS & COLLABORATIONS	• Quick and efficient delivery of projects. • Better value for money. • Facilitates co-creation, which capitalizes on collaborative opportunities to exchange data and knowledge between organizations.
INDUSTRY LEADERSHIP	• Brings more innovation, and autonomous and knowledge services, in the sector at a much faster rate. • Appropriate skill development within the infrastructure sector, which includes geospatial knowledge.
APPLICATIONS, ANALYTICS AND MODELLING	• Effective digitalization of the construction sector. • Deliver tangible improvements and benefits. • Generate and update planning and databases with geospatial data and within a single integrated platform, which improves efficiency. • Provide real-time monitoring of the project. • Efficient resource management and better understanding of risk factors.
GEOSPATIAL DIMENSION TO DIGITAL INFRASTRUCTURE	• Better understanding of interoperability beyond the standardized semantic interoperability of integrated systems. • Easy transition from information to knowledge management. • Increased reusability of knowledge and information within the sector outside the scope of a single project. • Provides technological neutrality and data portability. • Allows transparency in the process. • Improves information management and data sharing.

Figure 6: The roles of GKI elements in fostering integration of GKI elements in infrastructure sector workflows

10.3.1 Trends

The infrastructure industry is leveraging technology to make construction management and site operations more efficient and sustainable. Building Information Modeling (BIM) continues to be a significant technology driver in the AEC industry and is being used extensively across the design, build and the overall project management phases of construction projects. However, recently, a more integrated approach has come to the forefront, wherein different geospatial technologies, along with BIM, are used in a handful of projects all over the world (as illustrated by the case studies featured in Section 10.5). They provide greater efficiency and productivity, improved accuracy, better integration of work processes, accurate visualization, and easy collective decision-making. Some key trends in the sector are mentioned in Figure 7 below.



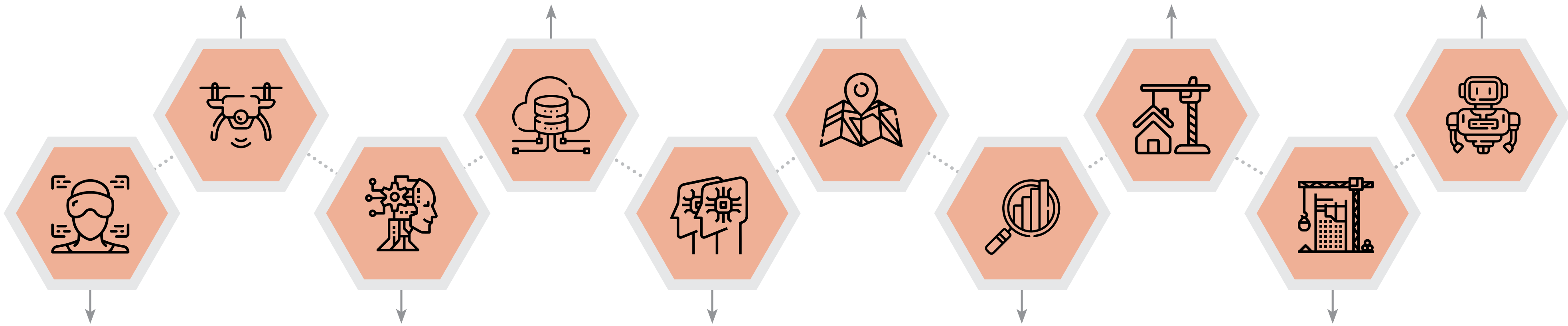
Source: GKI Readiness Index Report

Figure 7: Trends in the infrastructure sector

10.3.2 The Future

The full suite of modern technologies used to advance the AEC sector as we understand it today are captured in Figure 8. The full suite of technologies highlighted in Figure 8 spans drones, cloud collaboration, the global navigation satellite system (GNSS), prefabrication and modulization, construction robotics and automation, immersive augmented and virtual reality solutions, generative design, digital twins, predictive analytics, and BIMs.

Figure 8: The technologies that play prominent roles in modern, state-of-the-art infrastructure workflows



IMMERSIVE SOLUTIONS — AUGMENTED REALITY/ VIRTUAL REALITY (AR/VR)

Enable real-time rendering of objects, streamline projects, prevent expensive re-working, keep projects on schedule, and improve collaboration.

DRONES/UNMANNED AERIAL VEHICLES (UAVS)

Provide highly reliable and precise information to boost productivity and efficiency, especially when combined with BIM and GIS in a common data environment.

GENERATIVE DESIGN — ARTIFICIAL INTELLIGENCE/ MACHINE LEARNING (AI/ ML)

Pushes the boundaries of design automation by enabling designers to reduce costs, improve productivity, and facilitate reuse of successful design processes.

CLOUD COLLABORATION

Improves processes through ease of access, document control, and better communication and checklist. It empowers stakeholders to measure efficiency and anticipate and predict project risks.

DIGITAL TWINS

Provide analytical capabilities through integration of data from multiple sources, and synergize this information into 3D models. This leads to sustainable resource management, value-added services, and improved connectivity across stakeholders in the value chain.

GLOBAL NAVIGATION SATELLITE SYSTEM (GNSS) AND/OR GLOBAL POSITIONING SYSTEM (GPS) EMBEDDED MACHINE CONTROL

Allows for flexible communication among the contractors and project managers in real-time and receives as-built data to ensure effective monitoring.

PREDICTIVE ANALYTICS

Enables risk determination in construction processes. It provides accurate insights into the failure behind the operation of construction equipment, and helps take required action early on . It also plays a critical role in managing resources and financials.

PREFABRICATION AND MODULARIZATION

Accounts for green construction measures due to offsite production, which results in less waste. It leads to significant reduction in raw material demand and minimizes energy consumption.

BUILDING INFORMATION MODELLING (BIM)

Integrating the BIM tool with energy simulation software revolutionizes conceptual design with real-time building performance data to gather instant energy performance feedback on early design decisions.

CONSTRCTION ROBOTICS (COBOTS) & AUTOMATION

Plays an important role in detection of errors and deviations, which enables companies to avoid delays in projects. It ensures quality of work with higher accuracy and reduces human tasks in dangerous environments.

The next part of the narrative dives more deeply into the individual roles play by GIS and BIMs and the emergence of the GeoBIM concept as a way to describe their use as an integrated data framework to support all four phases of the infrastructure portrayed in Figure 3.

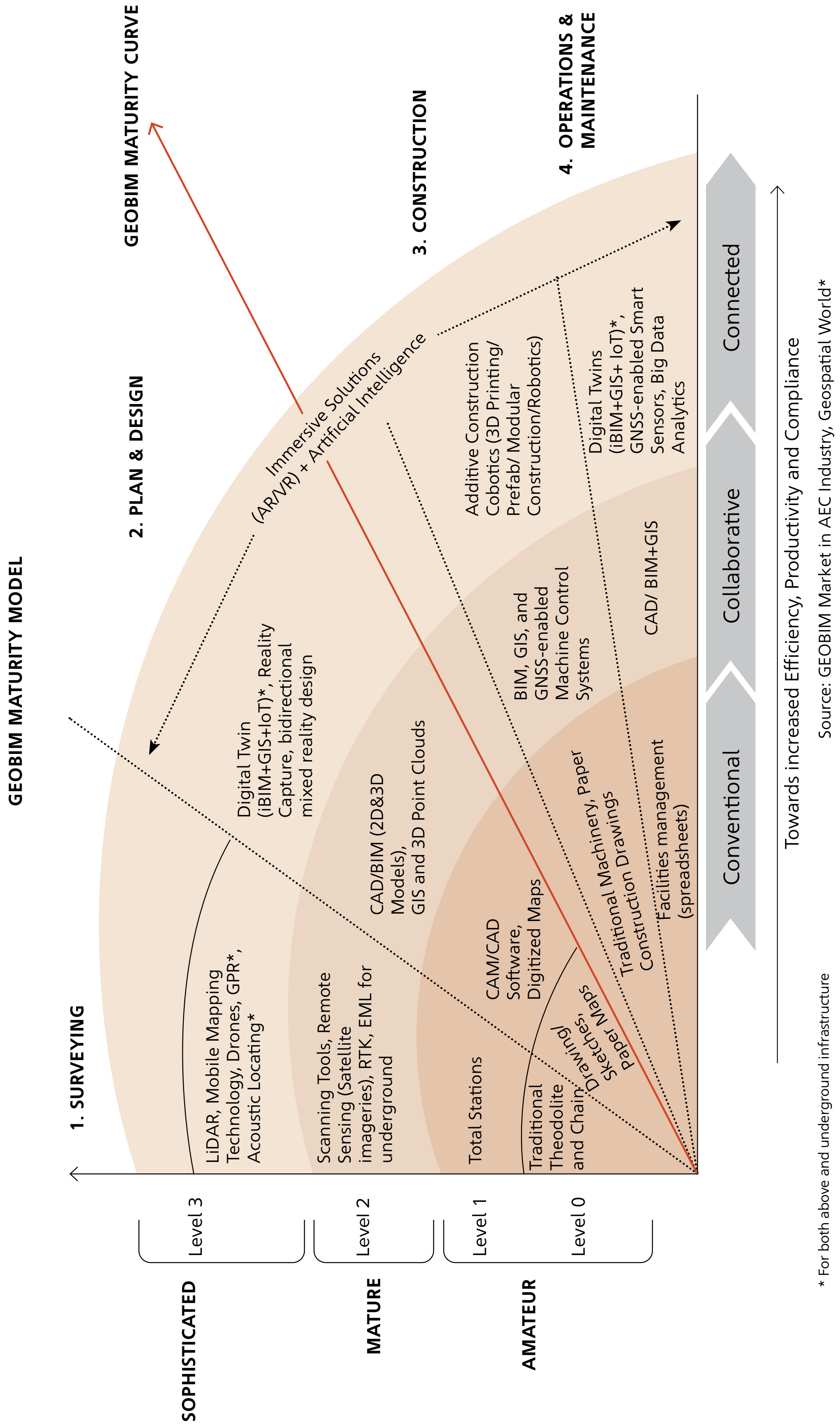


Figure 9: The GeoBIM maturity model

GIS generally operates at a larger scale than BIM and locates data across a topography, so it is particularly suited for urban planning and the design of smart cities.

Geospatial data from GIS can describe the people of the place as well as the places themselves, such as the urban built environment, public spaces, green spaces, water sources and quality, electric grid and transport infrastructure. The merger of GIS and BIM models provides improved data accuracy, reduced project risk, access to project information for all the stakeholders, and integration of project information from multiple data platforms. Further, it also provides urban planners with the opportunity to develop urban digital models, which visualize and analyze the ability of the urban environment to support human wellbeing along with sustainable and resilient communities.

GeoBIM, which describes the use of GIS and BIMs together, is not a new concept. However, an assessment of its maturity is critical to evaluate and assess the ability of an AEC firm to operate in a collaborative and connected data environment.

Geospatial World has developed the GeoBIM Maturity Model to understand the maturity level of adoption of geospatial technology and integration with BIM in the sectoral workflow, ranging from Level 0 to 3 (Figure 9). As more innovation takes place in the GeoBIM technology, and the integration of geospatial and BIM technologies becomes an accepted definition across the construction lifecycle, the maturity model will also progress beyond Level 3.

Note that in the GeoBIM Maturity Model portrayed in Figure 9, surveying is kept as a separate process under the construction lifecycle as its use cuts across the planning, design, build, and operations and maintenance phases.

The challenges we are facing today are large and growing in magnitude. To solve these problems, we need systematic solutions, which are beyond the capacity of any single human brain. Geodesign, by invoking geospatial technologies to focus on what the world can become and upgrading the professions of design, gives us a way to extend our capability, with the vast information technologies, increasingly powerful computation tools and big data, multiple shareholders, and more and more varied participants, to conceive appropriate solutions to address our present challenges that taken as a whole will impact the survival and wellbeing of humanity and for that matter, Planet Earth as well.

We now turn next to the last part of this session, which uses five case studies to show what becomes possible when the geospatial knowledge infrastructure and the accompanying technologies (BIMs, GIS, GNSS, and in situ and remote sensing) are deployed across a variety of important infrastructure projects.

10.4 Five Case Studies Illustrating the Role of the GKI in Today's AEC Workflows

The five case studies highlighted below illustrate the use of nature's services to build more sustainable and resilient urban centres, urban redevelopment projects that seek to create more liveable and sustainable urban centres, new urban environments that focus on human wellbeing, sustainability and resilience from their conception, the use of geospatial and accompanying technologies to support the redevelopment and modernization of one of the world's largest international airports, and the remediation of one of the world's largest abandoned mine sites.

10.4.1 Sponge cities (as one example of nature-based services)

The concept of sponge cities is most often attributed to Dr. Kongjian Yu who serves as the Dean of the School of Architecture and Landscape Architecture at Peking University and as Principal Designer at the internationally acclaimed Turenscape urban design practice based in Beijing, China.

The role filled by sponge cities as an example of how nature-based services can change the ways in which we think of critical infrastructure in large metropolitan areas and is perhaps best captured by a series of quotes reproduced below.

VISIONARY, NATURE-BASED URBAN DESIGN FROM CHINA

What if cities were designed so that they could absorb excess rainfall, neutralize floods, and turn their streets green and beautiful in the process? Kongjian Yu is doing just that, as he reports from China. This award-winning leader in ecological urbanism and landscape architecture, and founder of the planning and design firm, Turenscape in Beijing, has become world renowned for his "sponge cities" and other revolutionary nature-based solutions. These approaches are being implemented in well over 200 cities in China and beyond.

Yu's extraordinary city-wide systems of stormwater-retaining ponds, wetlands, and parks draw from both ancient Chinese hydrological wisdom and cutting-edge design and engineering to offer the whole world a model of inspired climate adaptation in an era of rising seas and extreme rainfall events.

Kim Harrisburg wrote an article in 2023 that asked whether sponge cities could curb climate-fueled floods. She noted how researchers at Arup had measured how much of seven major cities were covered by 'green and blue infrastructure' including grass, trees, ponds and lakes, and how much was covered by 'grey infrastructure' such as concrete, pavement and buildings. Auckland in New Zealand finished first (35%) followed by Nairobi in Kenya (34%), New York in the USA, Mumbai in India and Singapore (30%), Shanghai in China (28%), and London in the UK (22%).

Her article went on to describe how many cities are adding green spaces to increase sponginess and deliver other benefits, from cleaner air to wildlife habitat and places to escape the summer heat. Landside-hit Freetown, Sierra Leone's capital for example, is planting trees to help prevent future disasters, while Tirana in Albania is creating a ring forest to clean the air and halt urban sprawl. This article also notes how digital mapping tools allow cities to quickly gauge the best use of their available space – from rainwater harvesting to ponds and inner-city gardens – as well as the risks of not doing so.

The extraordinary city-wide systems of stormwater-retaining ponds, wetlands, and parks envisaged by Yu provides a model of inspired climate adaptation in an era of rising seas and extreme rainfall events and the various projects he has led illustrate how modern digital tools and workflows can be used to complete the first three phases of these projects. Those interested in learning more about Yu's work should consult a book on his work edited by William Saunders and published in 2015 and the website built by Yu and his team in 2021 detailing the theory and methods supporting the sponge city concept.

10.4.2 The Meridian Water Redevelopment Project in Enfield, United Kingdom

The second case study is a $7.75 billion, 20-year London regeneration program led by Enfield Council, bringing 10,000 homes and 6,000 jobs to the Borough of Enfield in north London. Alongside beautiful homes and world-class public spaces and community facilities, the 200-acre development will support a new railway station, linking commuters to the broader London region. The aim is to build something that is sustainable and resilient.

Jacobs Solutions Inc. is partnering with Enfield Council to provide program and master plan leadership, technical design and planning advisory services for the development of a new masterplan for the Meridian Water redevelopment. They also provide client technical advisory service across several disciplines during the design and delivery of over $150 million of strategic infrastructure works to help unlock the Meridian Water site, and the data visualization platform and KPI-driven geodesign processes required to synthesize data coming from multiple sources into a common and connected framework to support decision-making.

Jacobs Solutions Inc. is a large American international technical professional services firm with ~$15bn in annual revenue and a talent force of ~60,000, provides engineering, technical, professional and construction services, as well as scientific and specialty consulting for a broad range of clients globally, including companies, organizations, and government agencies.

The Meridian Water redevelopment results produced to date include a set of live, interactive 2D and 3D dashboards and models to help optimize the master planning process to ensure that it delivers on the vision and meets the sustainability and resilience goals set forward by the council, including:

- A vibrant mixed-use development with access to parks and green infrastructure for flood control combined with multiple well-being benefits.
- Access to multi-modal transit and pedestrian-friendly streets.
- A vibrant setting for 'making and creating, with workshops, shops, restaurants, businesses, and with ample affordable housing.

The geodesign approach used in this project provides a powerful planning method that uses stakeholder input and spatial analytics to simulate the possible impact of alternative design scenarios.

The geodesign value proposition noted by Shannon McElvaney and Mary-Jane Wood, the lead designers of this urban redevelopment, included:

- Increased design efficiency saving time and money.
- Improved communications using visually compelling content.
- Enabling innovative solutions through realized synergies.
- Increased speeds of data sharing, reporting, and the surfacing of issues.
- Improved transparency and trust.
- The use of a documented and data-driven approach to validate and back the decisions that are made during the redevelopment process.

Geodesign is a way to invoke and use geospatial methods and data to describe not only the current state of the world but also to tackle the creation and monitoring of the performance of what various locations could become in the future and as such, can support all four of the infrastructure phases noted at the start of this session.

10.4.3 The Construction of Neom and the LINE City project in Tabuk Province, Saudi Arabia

The third case study focuses on the construction of Neom in Tabuk Province, Saudi Arabia. The site is north of the Red Sea, east of Egypt across the Gulf of Aqaba and south of Jordan. The total planned area of Neom is 26,500 km^2 and this new city will stretch across 170 km, from the mountains of Neom across desert valleys to the Red Sea. The plans completed in 2017 envisage a mirrored architectural masterpiece towering 500 m above sea level, but just 200 m wide.

Work is ongoing and on March 20, 2024, Neom's Board of Directors unveiled Treyam, a premier resort positioned where the desert landscape meets the sea. Boasting a bold design with minimal land intervention, Treyam offers guests an elevated site in which to unwind, re-energize and enjoy stunning natural surroundings. Strategically located across the opening of one of the most beautiful azure lagoons at the southern end of the Gulf of Aqaba, Treyam stands as a luxury gateway inviting guests to pursue numerous indoor and outdoor activities. Its striking bridge-like architecture connects the northern and southern shores, hosting a 250-room luxury resort. The visual splendour of Treyam extends further, with a 450-m long rooftop infinity pool elevated 36 m above the sea, providing guests with stunning views.

The LINE city project, once built, will eventually accommodate 9 million people and will be built on a footprint of just 34 km^2. This will mean a reduced infrastructure footprint, creating never-before-seen efficiencies in city functions. The ideal climate all-year-round will ensure that residents can enjoy the surrounding nature and they will also have access to all daily essentials within a 5-minute walk, in addition to high-speed rail – with an end-to-end transit time of just 20 minutes.

10.4.4 The Los Angeles International Airport Development Program

The fourth case study describes how geospatial methods and data are being used to reimage and modernize the Los Angeles International Airport (LAX) ahead of the 2028 Olympic Games whist the airport continues operating. As part of the LAX Development Program, Los Angeles World Airports (LAWA) is strategically leveraging GIS to streamline the workflows of multiple construction projects. It is also using BIM applications to ensure that operations of the airport proceed with minimal disruptions.

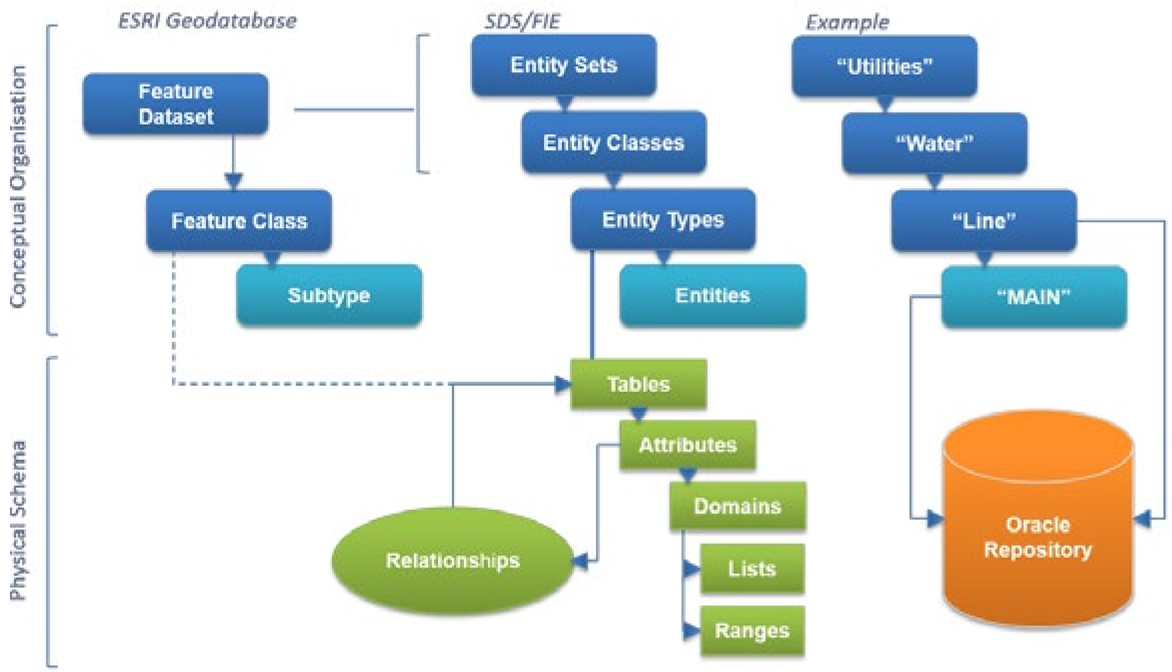
A GIS has been used to create a platform called Coordination and Logistics Management (CALM), which combines the Airport Enterprise GIS (AEGIS) and the document management system to help managers oversee the airport's multiple construction projects. Staff and stakeholders use the system to access all construction records and drawings. The GIS displays project locations and provides a geospatial interface for data. Level of Development (LOD) 200-400 LiDAR data have been integrated in a 3D model of the airport to help with the visualization of project information and integrated GIS data to identify workflow conflicts. The benefits of standing up and deploying these geospatial assets include lower project contingencies, risks and costs due in part to the workflows described next.

The LAX CALM Program was primarily developed to manage the complexity of coordinating on-going construction projects while minimizing airport disruptions. The CALM team reviews and approves logistics and phasing plans for all planned construction projects to minimize the impact to on-going airport operations.

KDG, a horizontal construction management firm with more than 40 years' experience and over $8 billion worth of capital improvement projects in aviation, transportation, and infrastructure, provides support to the CALM team, with resident managers tracking each project from start to finish including the individual project phases. The specific services include:

- Analysis of adjacent projects to ensure they do not have conflicting activities scheduled at the same time.
- Meeting with project stakeholders to discuss potential consequences of a project and to develop mitigation strategies for reducing these impacts on the traveling public, airport tenants and operations.

The current work has also benefited from earlier investments, including the development and adoption of GIS Standards, by LAWA in 2014. These standards covered CAD and GIS data and Figure 10 shows the spatial data hierarchy that underpins the standards. This part of their geospatial knowledge infrastructure is required for all five of the case studies highlighted in this section and speaks to the need to employ geospatial specialists to help with the planning, design, operation, and maintenance of infrastructure projects of all sizes.



Source: LAWA Information Technology Management Group (2014)

Figure 10: Los Angeles World Airports spatial data hierarchy

The current LAX work builds on prior successes – the airport was awarded the U.S. Green Building Council's Leadership in Energy and Environmental Design (LEED) Gold Certification in 2015 for the redevelopment of the Tom Bradley International Terminal and international certification for greenhouse gas reductions in 2019 and the current work aims to coordinate the logistics for the 130 projects running in parallel and support uninterrupted airport operations despite work on multiple improvement projects spread over multiple years.

10.4.5 Faro Mine Remediation Project

The fifth and final case study focuses on the Faro Mine remediation project in Canada's Northwest Territories.

The Faro Mine was once the largest open pit lead-zinc mine in the world but today, it is the site of one of the most complex abandoned mine remediation projects in Canada. The mine spans 25 km^2 and is located in south-central Yukon, near the town of Faro, on the traditional territory of the Kaska Nations, and upstream from Selkirk First Nation. The processing the valuable minerals at the mine left behind 70 million tonnes of tailings and 320 million tonnes of waste rock. This waste has the potential to leach metals and acid into the surrounding land and water.

The Canadian Government chose AECOM, a multinational infrastructure consulting firm with 57,000 employees and revenues of ~$14.4 billion in 2023, to coordinate and lead this large and complicated remediation project. AECOM is working with the Government of Canada, affected First Nations communities and additional consultants to conduct remediation plan design and quality assurance support for the Faro Mine Remediation Project. Their work will help minimize the mine's impact on the environment and the health and safety of community members, including the Kaska Nations (Ross River Dena Council, Liard First Nation and Kaska Dena Council) and Selkirk First Nation.

This final case study shows the need to advance plans to build and operate a circular economy and the five case studies, taken as a whole, demonstrate the importance of adopting and using the geospatial knowledge infrastructure and accompanying geospatial tools and data to plan, design, operate, and manage new projects (Neom and the LINE City Project), redevelopment projects (the LAX and Meridian Water Redevelopment Projects and those using the sponge cities concept) along with impacts of human activities that have been discontinued (i.e., the Faro Mine Remediation Project).

10.5 References

E | 2. (n.d.). Los Angeles International Airport Development Program. https://e2.com/experience/los-angeles-international-airport-development-program-support/

Enfield Council. (n.d.). Meridian Water. https://www.meridianwater.co.uk/#:~:text=Meridian%20Water%20is%20Enfield%20Council's,brings%20nature%20onto%20your%20doorstep

European Parliament Research Service. (2015). https://www.europarl.europa.eu/topics/en/article/20151201STO05603/circular-economy-definition-importance-and-benefits

FMRP. (n.d.). Faro Mine was once the largest open pit lead-zinc mine in the world: Today it is the site of one of the most complex abandoned mine remediation projects in Canada. https://faromineproject.ca/

Geospatial World. (2020). The Power of Where: A Geospatial Knowledge Infrastructure to Enhance the World Economy, Society and Environment. https://geospatialmedia.net/pdf/GKI-White-Paper.pdf

Geospatial World. (2022). Geospatial Knowledge Infrastructure Readiness Index and Value Proposition in World Economy, Society, and Environment. https://www.geospatialworld.net/consulting/gki-phase-2/gw-assets/pdf/GKI-Report.pdf

Government of Canada. (n.d.). Faro Mine Remediation Project: Yukon. https://www.rcaanc-cirnac.gc.ca/eng/1480019546952/1537554989037

Harrisberg, K. (2023). Are 'sponge cities' enough to curb climate-fuelled floods? https://www.context.news/green-cities/are-sponge-cities-enough-to-curb-climate-fuelled-floods

Karakusevic Carson Architects. (n.d.). Meridian Water Masterplan. https://www.karakusevic-carson.com/projects/meridian-water-masterplan

Los Angeles World Airports. (n.d.). LAX: Change is in the Air(port). https://www.lawa.org/transforminglax/projects

McElvaney, S., and Wood, M-J. (2021). Meridian Water: Sustainable and Resilient by Design. https://www.youtube.com/watch?v=NTaz1wtRP7I

Neom. (n.d.). NEOM | Changing the Future of Design and Construction. https://www.neom.com/en-us/our-business/sectors/design-and-construction

Neom. (n.d.). NEOM | What is THE LINE? https://www.youtube.com/watch?v=0kz5vEqdaSc

Neom. (n.d.). NEOM Made to Change | Treyam. https://www.neom.com/en-us/regions/treyam

Paslay Group. (n.d.). Los Angeles International Airport. https://paslaygroup.com/?gad_source=1&gclid=EAIaIQobChMI8pnur_SihQMVNaNmAh2KVAykEAAYASAAEgIo5fD_BwE

Saunders, W.S. (2015). Designed Ecologies: The Landscape Architecture of Kongjian Yu. https://www.amazon.com/Designed-Ecologies-Landscape-Architecture-Kongjian/dp/3034607385

Shea, S. (2020). Smart city. https://www.techtarget.com/iotagenda/definition/smart-city

Worldometer. (n.d.). World urban population. https://www.worldometers.info/demographics/world-demographics/#urb

Yu, K. (2023). "Sponge Cities": Visionary, Nature-Based Urban Design from China. https://www.youtube.com/watch?v=6s8LtyCtpBwYu, K. et al. (2021) Sponge City: Theory and Practice. https://www.turenscape.com/topic/en/spongecity/index.html

LESSON – 11

Strategy for integrated geospatial knowledge for digital economy

Trainer's Biodata: Dr. Mark C. Abrams

Dr. Abrams specializes in calibration and independent verification and validation of complex information systems. Examples include establishing absolute spectral calibration systems in the 1980s, establishing 3D geodetic networks with DGPS class stability and accuracy, and demonstrating and certifying 3D remote sensing systems with sub-meter absolute accuracy.

In 2004, Dr. Abrams demonstrated 11 cm (3DRMS) absolute geodetic calibration in the remote third world. In 2012, he led a USG team that certified a country-scale LIDAR survey with an absolute accuracy of 2 m CE90 and 1 m LE90, from which he developed a concept of "vendor self-certification." Most recently, he supported the certification of high-resolution 3D digital surface models and co-registered imagery covering 26 globally distributed international sites. He has deployed on four continents, most recently in Africa, and has delivered national-scale reference databases across the last three decades on each continent.

Dr. Abrams believes in hands-on science and engineering and actively mentors the next generation of technicals and field officers, sponsoring and teaching hundreds of government employees and deployers. He is the father of four adopted boys, each from a different continent. His oldest is now a cybersecurity engineer, number two has spent three years learning field operations and logistics, number three is fluent in multiple languages, and his youngest is both learning disabled and training to be a physicist.

11.1 Introduction

Government policies are often developed in silos, with sectoral policies and cross-cutting policies such as data policies often not aligned, leading to a fundamental challenge of "how do you geospatially enable a digital economy at marginal cost at the local level."

Section 11 comprises four fundamental concepts: partnerships, standards, integration, and capacity development. Each concept shall be illustrated with an implementation approach that starts from essentially nothing and builds something nationally and internationally relevant to GKI and participation in the global location-based digital economy.

Three of the nine strategic thrusts of the UN GGIM will be used as anchors for the implementation approach: partnerships, legal and policy, and standards. The premise of this strategy is that a common GKI framework underlies what all 200+ countries could implement if they chose to work together and focus on interoperability, standards, and cooperation.

11.1.1 Examples

Illustrative examples of highly accurate global databases that link the 201 countries of the planet: (a) global airfields, (b) global shipping, and (c) global air traffic will be used as foundational data in this discussion (Figure 1).

Figure 1 A worldwide map of 45,000 airports with no contextual reference map.

11.1.2 The Digital Ecosystem (extract from Abrams and DePalacios, 2012)

In a peculiar fashion, the explosion of location-enabled cellular devices and the public availability of high-resolution base maps provide a new foundation for measuring and defining human geolocation and the space-time dynamics of human movement on the planet. Current GPS and precision navigation technologies dramatically extend the definition of location from the 1st world into the remote third world. With the deployment of Differential Global Positioning System (DGPS, now Differential Global Navigation Satellite Systems (DGNSS)) control networks into the third world and the use of Satellite-Based Augmented System DGPS (SBAS DGPS), it is possible to investigate the potential for terrestrial and airborne mapping at human scales in real time. Combining geodetic networks with commercially available location-enabled devices challenges the notion that human-scale geolocation is only available to governments. It presages a convergence between location-enabled services and mapping the planet at human scales and human-relevant timescales.

The existence of accessible control networks in remote areas allows for the characterization of the absolute and relative geolocation performance of location-enabled devices – cellphones, tablets, cameras, GPS units and trackers – any of the modern devices that have a GPS chip in them. In contrast to geodetic GPS and DGPS devices, these commercial-off-the-shelf GPS devices make commercial compromises to balance their geolocation and geolocation accuracy with every other engineering constraint (primarily power) that is imposed on a low-cost device that needs to provide multiple functions in your pocket all day long. The commercial navigation community realized that an acceptable solution most of the time was better than a perfect solution some of the time and developed a family of approaches to multi-path mitigation – from which evolved the current generations of GPS-enabled devices that work under canopy inside cities, inside buildings, and underneath vehicles.

These devices yield surprisingly good geolocation results in their normal low power modes – sub-3-meter class results in some cases and they can generate trajectories that are often much more precise (sub-meter) with a residual bias of 1-3 meters in absolute coordinates[1]. In an integration mode (high power), these devices display a traditional random walk pattern (the so-called "happy dance") that collapses to a

1 See 2012 examples from Kandahar, Afghanistan, and San Diego, U.S.A., two extremes of "smart" environments in the charts, Abrams and DePalacios, NAP, 2012.

remarkably accurate value with sufficient averaging[2]. Used as logging navigation and geolocation devices, the results provide a simultaneous map of position and RF interference – given sufficient time, these devices will map out the RF permeability of the canopy overhead or the multipath inside a building (Istanbul). In addition, these devices map out the RF voids – the negative spaces in urban canyons and remote forests where GPS quad-lateration cannot be achieved and a geolocation solution is not available.

In 2012, we predicted (NAP 2012) that the proliferation of cell phones and smart devices would create a digital ecosystem with more than 600 GPS geolocation events per square meter per year. Technology has far outpaced this prediction. In the subsequent decade, with the advent of ad ID and device ID, many smart devices are tracked 86,400 times per day (once per second), generating a nearly USD 90B digital ecosystem leveraging self-volunteered geolocation data. As was demonstrated in 2013, four unique geolocation tags are sufficient to identify 95% of smart device users based on their pattern of life activity (de Montjoye, 2013). This is precisely the nexus of geospatial knowledge infrastructure and the digital economy.

11.1.3 Growth in Space Vehicles (and space observations)

In 1966, space was the domain of a few select countries, including the United States, the USSR, Canada, England, France, and Italy, and only two countries had active space launch capabilities. In 2023, there were 7560 satellites in operations (Figure 2), approved licenses had been issued for 65,000 satellites, and applications were pending for a total of 136,000 satellites.

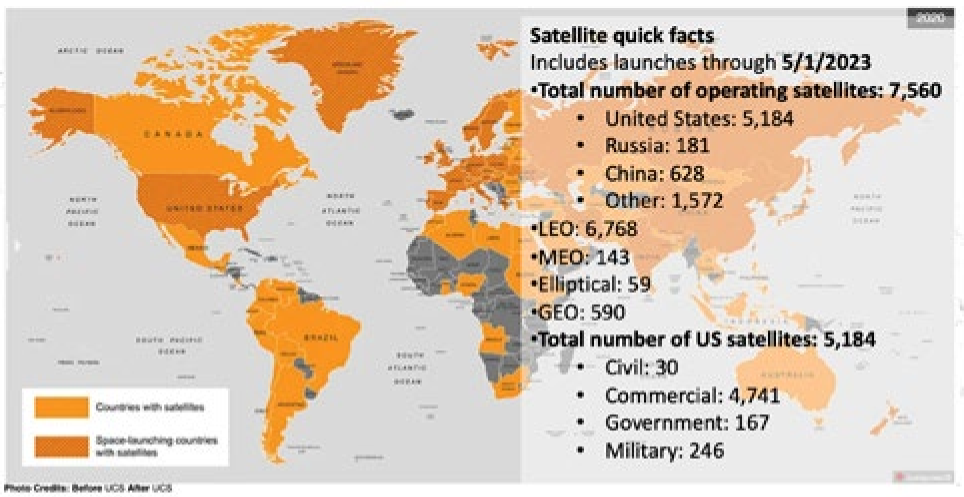


Figure 2 Countries in 2020 with satellite programs and summary of satellites in 2023 (after UCS).

At the same time, the development of commercial launch capabilities has dramatically reduced the cost of a space launch from $10,000 per kilogram to less than $2,000/kg (2022), and with SpaceX's Starship, potentially below $200/kg. The unplanned consequence of this transformation of space access is the creation of significant space debris, which must fundamentally transform space access and operations in the following decades.

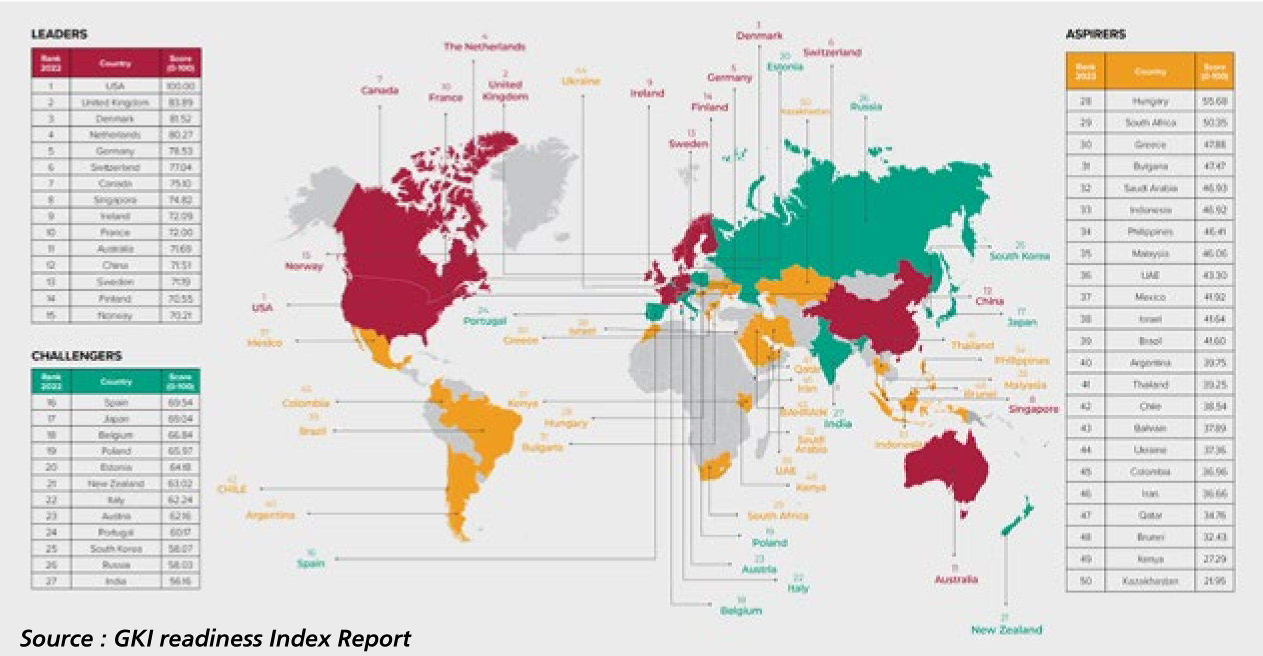
2 The results from 2014/5 of 4.9m are probably indicative of "typical smart device" results: van Diggelen, Frank, Enge, Per, "The World's first GPS MOOC and Worldwide Laboratory using Smartphones," Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015), Tampa, Florida, September 2015, pp. 361-369. https://www.ion.org/publications/abstract.cfm?articleID=13079, or https://www.gps.gov/systems/gps/performance/accuracy/

The most recent (2022) results were 1.38 m (average of 50% and 95% horizontal errors from 36 samples), Van Diggelen, F., et al., Smart ways to improve smartphone location accuracy, GPS World, 2022.

https://www.gpsworld.com/smart-ways-to-improve-smartphone-location-accuracy/

11.1.4 Taxonomy and Demography of Geospatial Knowledge Infrastructure (UN GGIM)

In this initial short course on strategies for geospatial knowledge infrastructure, we will focus on the "bottom" 151 countries that did not rate in the 2023 UN GGIM top-50 assessment (Figure 3). We will discuss three elements of the UN GGIM strategy: partnerships, legal and policy, and standards, and use them to examine multiple case studies that illustrate the potential for doing something useful without significant investment capital (Figure 4).



Source : GKI readiness Index Report

Figure 3 UN GGIM performance index assessment (2023).

Partnerships: Self-validation and certification will be introduced as an approach to building partnerships between GKI leaders and the "Next One Hundred and Fifty-One" (NOHFO) countries, with the idea that the leaders desire global validation and the NOHFO countries have sharable local geolocation knowledge.

Standards: With international collaboration, we can aspire to a common, interoperable set of GKI standards that allow the global GKI community to actively participate in the digital economy enabled by location-based services. Absent this collaboration, vendors and location-based service providers will exploit each country's inexperience and maximize their profitability[3].

Policy and Legal: Ultimately, each country will contract services to support the development of its GKI. A critical opportunity arises from the realization that none of these contracts will be technically "unique." Across the GKI ecosystem, a country has contracted for a "like/as" service, product, or capability. Creating a community of practice that shares both technical and policy best practices allows the NOHFO countries to leverage the experience of the leaders, challengers, and aspirants in the GGIM ecosystem.

Teaching acquisition best practices may seem relatively "non-technical," but a non-technical contracts officer will authorize most contracts; providing best practices and training for the "contracting officer's technical representative" is the only way to ensure that an acquisition delivers the intended product or service with the expected performance, on schedule, and cost. Building a library of sharable acquisition documents minimizes the risk for all countries. It shares the good and bad experiences when single words such as "interoperable" are omitted from a contract's terms and conditions.

3 A reviewer noted the following: "Think of the lost opportunities … for everyone … absent this. Only a few benefit financially, and only for the near term." Acquisitions remain legal processes and documents with a tendency towards obscurity and complexity, often hidden behind proprietary terms and conditions. Ultimately, contract terms and conditions are best created in public, in an open, shared, peer-reviewed environment where the whole can benefit from the experience of the few.

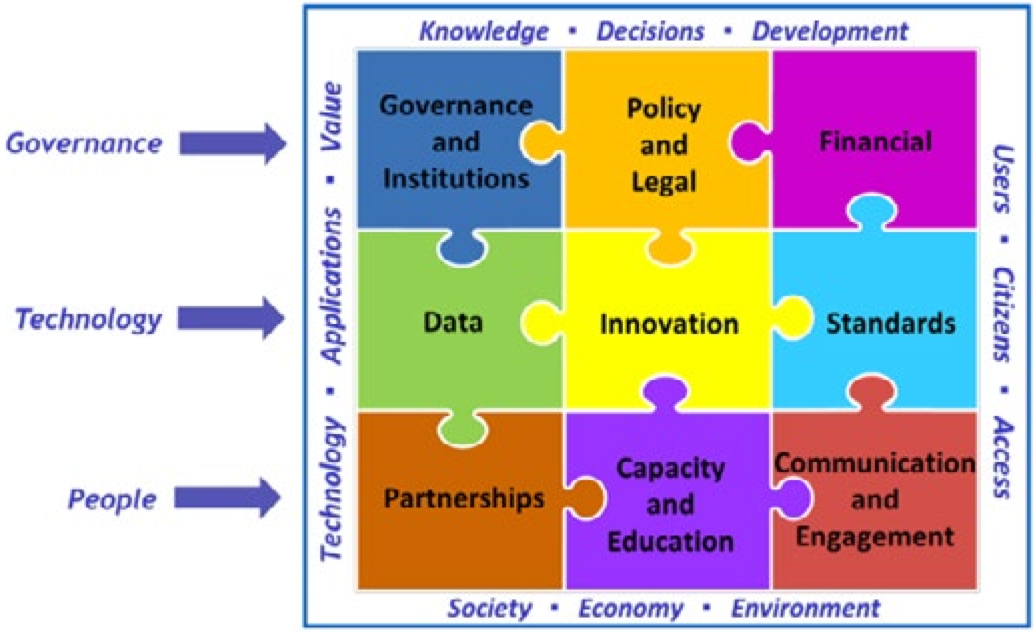


Figure 4 UN GGIM Nine Strategic Pathways; we will focus on "Policy and Legal" to reimagine acquisition policy, "Standards" to aspire to less than 201 unique standards, and "Partnerships" to examine what the NOHFO countries can do with marginal resources.

11.1.5 Methodology

We will examine the GKI strategy with three themes:

- Interoperability and breaking silos within and between countries;
- Self-validation and establishing a local basis of absolute geolocation knowledge; and
- Cost-benefit analyses that can permit the TOP-50 countries to share their best practices with the NOHFO countries.

We will share a struggle with two conflicting metaphors:

- If we have to do something two hundred and one times, how do we minimize the duplication of effort?
- Can global localization and local globalization inform one another effectively?

Finally, we will build around one concept, self-certification, four things the NOHFO countries can learn and do with marginal or no real investment.

- Building calibration targets from gravel yards to create new partnerships (11.2).
- Leveraging boundary regions where the TOP-50 and the NOHFO counties meet, namely the 45,000 airports distributed across all 199 countries with airports, and
- self-certification using existing aviation standards for instrument landing systems (ILS). (11.3)
- Data integration through registration and the ability to fuse global and local data. (11.4)
- Capacity development through local and global self-certification (case studies) (11.5)

We will conclude section 11 with a question-and-answer discussion focusing on two foundational questions and individual exercises:

11.1.6 Questions and Answers

Individual national strategies inherently struggle with international collaboration and a definition of a national strategy in an international context. Simple questions become difficult.

- How do we define requirements?
    - > Coverage, currency, and completeness?
    - > Resolution (in space, time, and spectrum), precision, and accuracy?
    - > Legacy?
    - > Interoperability?
    - > Timeliness?
    - > Resilience?
- How do we develop and grow a community of practice?
    - > What can we learn from Picasso's admonitions that:
        - » "Good artists copy, and great artists steal," and
        - » "the chief enemy of creativity is 'good' sense."

11.1.7 Exercises

Individual exercises are introduced upfront to posit possible challenges that any national strategy must address. They are used to bring the participants into a discussion of the local implementation of strategy within their own local digital and geospatial ecosystems with a global perspective.

Standards should be the most trivial item of common agreement, yet there is a natural tendency towards 201 unique standards.

In the U.S.A., we have at least three standards for geolocation accuracy (MIL-STD 600001 (1990), Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy, 1998, and the American Society of Photogrammetry and Remote Sensing (ASPRS), Positional Accuracy Standards for Digital Geospatial Data (EDITION 1, VERSION 1.0. - NOVEMBER 2014), and a fourth is in development and approval. While incremental improvements have been made, none of the three fundamentally change where a measurement "is" or "how accurate" it is; rather, they simply redefine the calculations and the necessary number of decimal places.

Is it possible to converge on a community of practice with a common set of interoperable standards?

Likewise, **policy** - the mechanisms by which we acquire, share, compare, enhance, and use geospatial knowledge - could be shared, but they are rarely shared. Can we create a sharable library of policy and acquisition documents? Can we identify relevant past performance references for common services? Can we leverage existing common services? Ironically, a "library" of past failures and lessons learned would be an invaluable starting place for the next set of acquisitions. There is a tendency for failures to be forgotten in the absence of documentation and for the same "failure" to recur decades later as a new generation tries to solve an old problem with new methods.

Partnerships start from a fundamental realization about geospatial knowledge strategy: while the methodology is inherently global, each national implementation is intrinsically **local. Locality provides unique knowledge (and value in a geospatially enabled digital economy).** Beyond locality, some relevant questions are: Is there something inherently unique in a particular implementation strategy? Why is it unique? Is there a legacy best practice that serves as a point of reference? How is an individual national strategy integrated into a partner network of peers, mentors, and mentees with whom a strategy is evolved and refined?

We should also be mindful of the dynamic nature of the local environment. Some localities are comparatively timeless… think a static desert… while others change on an almost minute-by-minute basis. These very dynamic localities will be challenging but also a source of invaluable lessons.

11.1.8 Implementation Guide: Linkage to the UN GGIM Nine Strategic Pathways

We will focus on three of the nine strategic elements: partnerships, policy, legal, and standards and use them as solutions to three fundamental problems that crosscut the challenge of effective GKI strategy:

Problem 1: All GKI implementers will realize (sooner or later) that 70-80% of GKI exploitation time is spent getting and aligning data. Interestingly, the last forty years have demonstrated that "one format" is not the answer.

Problem 2: Acquisition best practices and standards are poorly defined, and sharing acquisition experiences is rarely accomplished, despite the fact that each of the 201 GKI implementations will allow some nearly identical contracts for products and services from a finite number of technical providers. If technical standards are difficult to accomplish, acquisition standards are even rarer.

Problem 3: Standards, interoperability, and effective contracting remain technically challenging.

The solution space consists of building and sustaining a community that understands:

- the critical role that standards play in acquisition, integration, and knowledge extraction;
- data integration, fusion, content and context extraction, and understanding;
- 100% of GKI development consists of activities done at least once, most successfully!

11.2 Productive partnership between stakeholders for the development of a national geospatial knowledge strategy

Calibration sites have a long and storied history in both the remote sensing and geospatial knowledge community. In the 1950s and 1960s, the US Army and Air Force maintained multiple large calibration targets for the qualification and certification of airborne and satellite-based remote sensing systems (Byars, 1975). In the intervening decades, these sites have dropped out of standard usage, with the NASA Stennis Calibration Site providing spatial and radiometric calibration for visible wavelengths and Railroad Valley Playa providing radiometric calibration at visible and infrared wavelengths. In contrast, the Chinese Academy of Sciences has built the Baotou Comprehensive Calibration and Validation Site in Inner Mongolia (northern China) to address spatial and radiometric calibration of airborne and satellite data across the visible, infrared, and RF spectrum (Li, 2015, Figure 5).

Figure 5 Baotou Comprehensive Calibration Test Site, Inner Mongolia (40.854 N 109.628 E).

It is unique in using gravels to provide a common target for visible, near-infrared, mid-infrared, and C, S, and X band wavelengths. It can be considered a "best-in-class" site for geospatial knowledge test, evaluation, IV&V, and data qualification worthy of global replication. We will revisit the Baotou site as a potential international standard for next-generation calibration sites.

Baotou was established in 2013 for airborne and satellite payload performance, calibration, evaluation, and validation. In 2014, the site was selected as one of the first four international demonstration sites for the Radiometric Calibration Network (RadCalNet) for the Committee on Earth Observation Satellites (CEOS), and in 2015, it was adopted into the global Modulation Transfer Function (MTF, which measures image quality) by the United States Geological Survey (USGS). It was prominently featured in the USGS's annual

Joint Agency Commercial Imagery Evaluation (JACIE) conference in 2024. The site consists of permanent optical targets with broad dynamic range, uniformity, and high stability due to natural gravel construction. Since 2013, the site has been used for payload radiometric calibration and on-orbit performance assessments of a variety of domestic and international satellites, including Sentinel, Kompsat, WorldView, Gaofen (GF - high resolution), Ziyuan/CBERS (ZY), Yaogan (YG), TH 1/2 (InSAR), and GaoJing/SuperView (GJ) as illustrated in Figure 6.

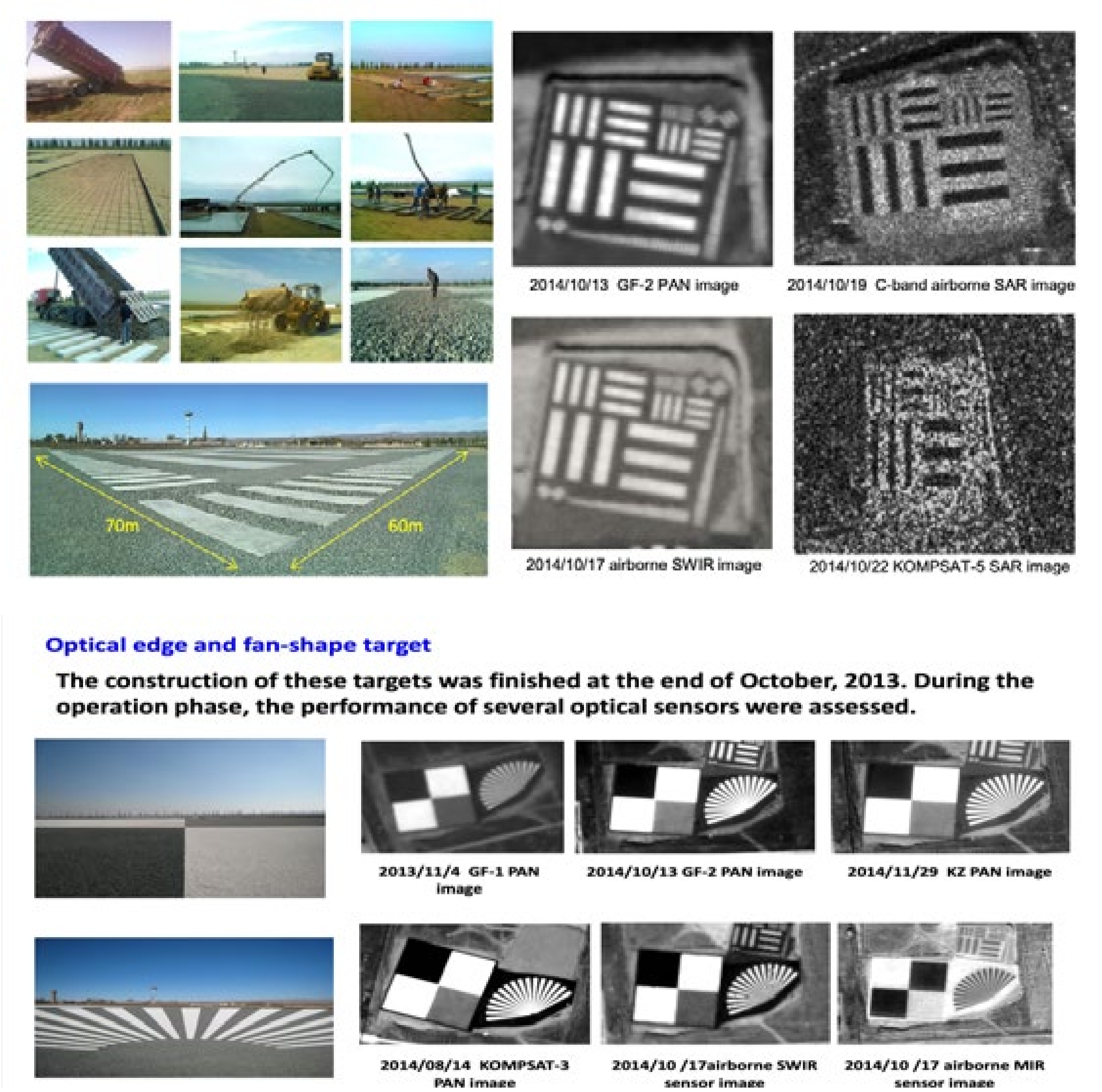


Figure 6 Optical, SWIR, MIR, C-band, and S-band Imagery of Baotou Test Site (after Li, 2015).

11.2.1 Resolution

Any discussion of requirements gets caught "betwixt and between" the simultaneous desires of resolution and coverage, with the obvious answer of "both" - that we want microscopic resolution with macroscopic coverage. A layer deeper lies the challenge of coverage, currency, and completeness; beneath that lies the challenge of legacy. And then there is the issue of precision and accuracy – random and systematic errors and ultimately the challenge of interoperability – both with the local historical record (the real legacy) and with the 200 other global GKI practitioners with whom we desire to share our locality and local knowledge. Notice, for the record, that we have not gotten to the data source – and it does not matter – regardless of the source, we need to think of the basic granule of data – the minimal sampling in x, y, z, and t that is necessary to describe the observation being made. Typically, we will work in seconds and meters, with GNSS (nee GPS) time setting a basic clock for activity on a per-second cadence, and for human activity, this makes the meter a sensible scale of human activity and motion.

One of the first questions that will be asked by a manager, scientist, or technician concerns the image quality:

- What is its resolution?
- What is the size of an object that can be seen in an image?
- What cannot be seen in an image?

Figure 7 Declassified U.S. satellite image (Brumfiel, NPR, 2022).

Resolution is frequently misunderstood, or worse, improperly applied by system designers, operators, and users, typically with an ambition to describe "the best capability" of a system without providing a concrete example. In the U.S., there is an imagery rating scale, the National Image Interpretability Recognition Scale (NIIRS), that was developed years after overhead imagery systems were designed and imagery collected. In a completely simplistic form, ground sample distance (GSD), which is the diffraction limit of optics, gives us a simple formula to relate the wavelength λ, aperture diameter(D), and distance (H) from the target of GSD = 1.22λH/D which yields a range of interesting image resolutions between 6.4 cm and 90 cm for typical apertures in low Earth orbit. GSD starts the conversation about data quality that allows the designer, engineer, interpreter, and user to have a sensible conversation about product quality; however, in practice, there is much more to the conversation. Two images are used to illustrate resolution: (a) a declassified U.S. satellite image with unspecified resolution (Figure 7), and (b) a demonstration radar image with 15 cm sampling (Figure 8).

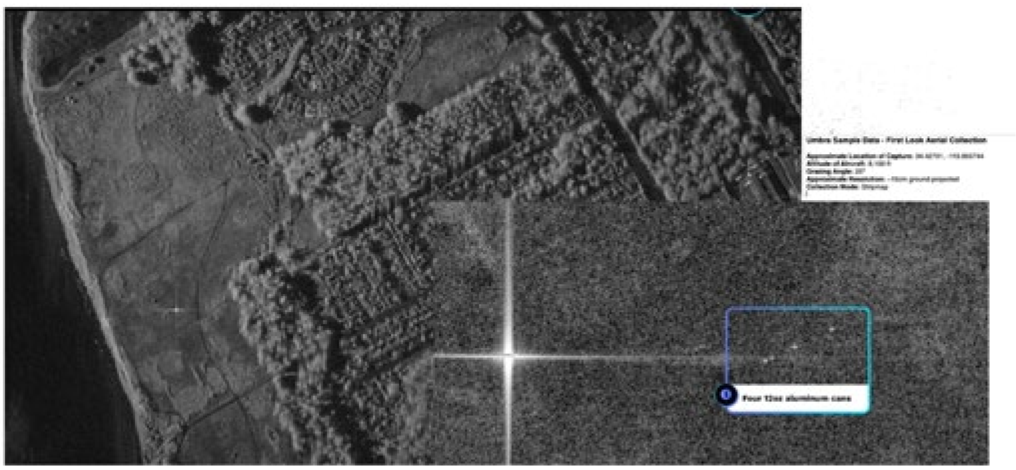


Figure 8 Umbra airborne 15 cm demonstration SAR imagery, 2021 (https://umbra.box.com/s/9x8oe8li6b5smz2p5k5k4kgmva6zxifa).

Both are spectacular images; however, the optical image spawned an immediate family of technical analyses, "reverse engineering" the possible platform and imaging system necessary to enable such an image. SAR images remain more esoteric and difficult to interpret, although they were used recently to demonstrate damage at an Iranian airbase on 04.19.2024. Finally, in an attempt to illustrate the challenges of cross-platform and cross-system resolution, Figure 9 illustrates a single site in both optical and SAR imagery.

Figure 9 Commercial satellite imagery of the Dole Pineapple Maze (Oahu, HI): 50 cm resolution optical image (Maxar/Google - left) and 15 cm SAR imagery (Umbra, right); the hedgerows are approximately 1 m in width and 2 meters in height (http://umbra-open-data-catalog.s3-website.us-west-2.amazonaws.com).

11.3 Strategy for implementation of standards for enhanced data acquisition, integration, and exploitation

Airfields constitute a unique convergence between the exquisite geospatial knowledge required for instrument landing systems (ILS), which utilize precisely surveyed runway markers to enable routine commercial air travel regardless of most weather conditions, and the digital economy, which flourishes due to global travel.

A summary discussion of data fusion of satellite imagery with airfield survey data illustrates the particular challenges (and amusements) of data fusion and conflation (Figures 10 and 11).

Figure 10 Two examples of a runway in 2020 (left) and 2016 (right) illustrating typical maintenance differences at airfields that confound data fusion and registration.

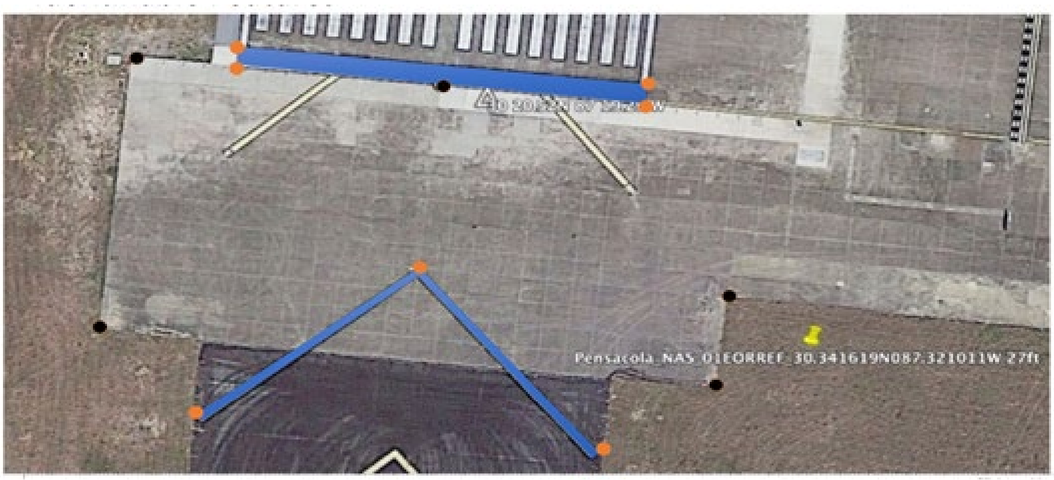


Figure 11 Recommended "photo-recognizable" target markers (orange) in addition to non-visual (black) traditional survey markers (blue coloration is for object identification only in this graphic.)

A recommended strategy for migrating from "invisible" survey markers to photo-recognizable survey markers that are AI-ready has been in discussion for more than a decade, but as with any discipline, such as cartography that has more than four hundred years of history to rationalize with modern technologies such as GNSS (30 years) and ADS-B (newly mandatory aircraft tracking beacons and systems). The solution expands airfield survey markers to include "photo-recognizable" features (runway markers) that are in proximity to the ILS survey markers. Ironically, most runways have multiple "arrow features" pointing at the runway to assist with visual landing in the event that ILS is not available (Figure 11).

Finally, existing, and emerging GKI accuracy standards are briefly discussed, including the legacy 1991 MIL-STD 600001 standard, the ASPRS 2014 accuracy standard, and a newly proposed NGA standard (NGA.SIG.0026).

11.4 Strategy for Integration of geospatial and digital ecosystems

Invariably, as requirements and capabilities evolve, every GKI ecosystem will have to address data registration and fusion, which will precipitate an extended analysis of calibration and geolocation standards. Recent work with 3D photogrammetric drones and terrestrial lidar scanners has pushed the sampling towards 5 cm resolution and 3-5 cm/axis absolute accuracy, challenging the legacy approach of registering the drone data to the medium resolution (50 cm) terrain and imagery data derived from satellites. With best practices in drone collection, ground control can achieve 20-50 cm RMSE absolute accuracy using SBAS/GNSS technology, enabling the high-resolution inset (HRI) drone data to be taken as reference data for the registration of the area data to the HRI data (Figure 12).

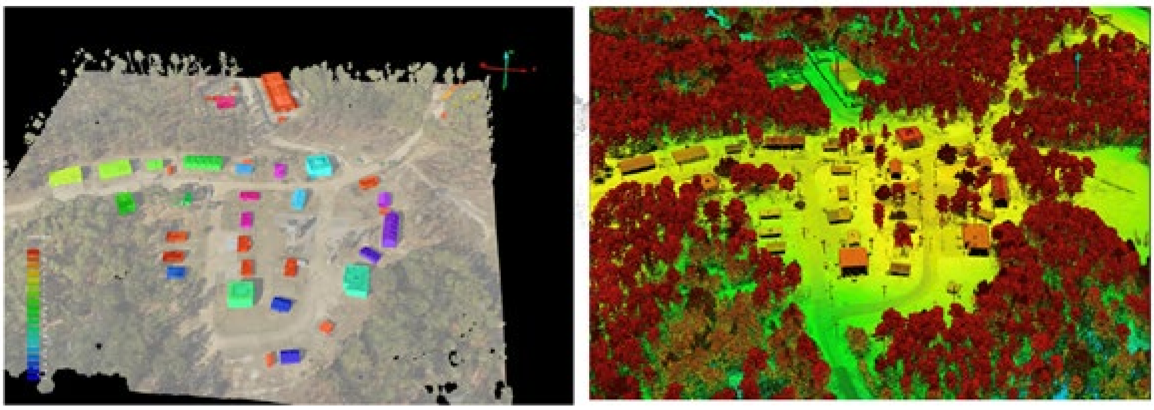


Figure 12 UAV-based photogrammetric point cloud with individual buildings identified and ready for mensuration (left) and USGS commercial (3DEP) airborne lidar data for reference (Neuenschwander et al., 2023)

The particular novelty of this strategy is that we are seeing the emergence and convergence of two independent commercial technologies: 3D photogrammetric drone systems and commercially available medium and high-resolution 3D commercial data from satellite providers, both with human-scale (or better) absolute accuracies of 3m (SE90) and ~ 30 cm RMSE_H as demonstrated in Table 1.

			3D Conformal Transformation Parameters						Errors (m)	
site	Acquisition date	Reference Data	DX	DY	DZ	θX	θY	θZ	RMSE baseline	RMSE after 3DC applied
JRTC Sangari	01/23	USGS lidar	0.34	-1.37	3.27	-0.019	-0.032	0.094	6.023	1.62
JRTC Sangari	01/23	TLS	0.89	-0.77	2.49	-0.037	-0.260	0.107	2.759	0.284

Table 1 A comparison of 3D roof GCP fitting of 3D photogrammetric drone point cloud data to airborne lidar and terrestrial lidar scan (TLS) demonstrating the ability to achieve less than 30 cm RMSE after registration (Neuenschwander, 2023).

11.5 Strategy for capacity development of the geospatial ecosystem

Returning to the initial premise of capacity development for the NOHFO countries, consider the following "hypothetical" considerations: purchase some 0.5m terrain and imagery data for the national airport and combine with existing ILS navigation data to establish a 3D control site (Figure 13).

Figure 13 Example 3D geolocation testing at Taipei International Airport, Taiwan using Maxar's Precision3D DSM and TrueOrtho imagery and airport runway markers for DGPS reference with a measured absolute accuracy of CE90 of 0.7m and LE90 of 1.1m (left and lower). The inset image (upper right) is the knife edge calibration target at Punghu Airport (23.519989 N 119.583726 E).

With that 3D context established, begin a conversation with existing and emerging data providers and value-added resellers about including your calibration site in their collection strategies and working with them to establish a local standard for exploiting their data in our locality.

Find an adjacent space at the airfield where you can clone the Baotou site (or the next-generation TxCALIPER supersites; more on that in a bit) and move from geodetic calibration to instrument performance characterization. This exercise was performed at the Taipei International Airport in Taiwan, with a derived accuracy of 0.7 m (CE90) and 1.1m (LE90). Combined with the knife-edge target at Punghu Taiwan, it is possible to address image quality and geolocation accuracy with indigenous resources.

In practice, self-certification is a multi-year (translation, slow and painful) process. Examples from Afghanistan (Figure 14) and Maxar's Precision3d (Figure 15) took multiple years but led to new standards in 3D geolocation. A final example is an emerging high-altitude LIDAR firm providing high-resolution, wide-area 3D mapping with remarkable accuracy (19 cm CE90 and 5 cm LE90) from 11,000-18,000 ft AGL (Figure 16).

Remote sensing is crucial in many aspects of society, including environmental monitoring, disaster management, navigation, and mapping. As geospatial data are used to inform decisions, data must answer three fundamental questions: what the spatial resolution is, how

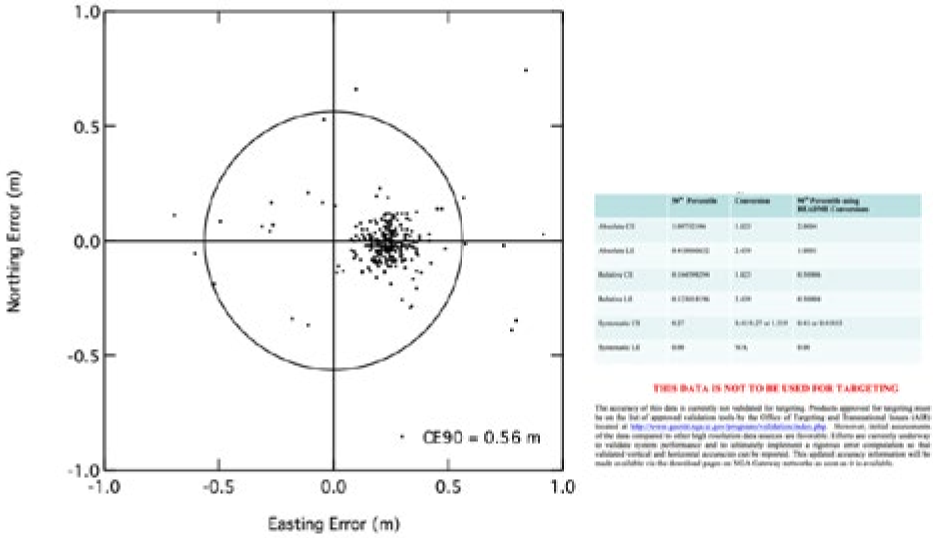


Figure 14 Example of self-certification to MIL-STD 600001 for the Airborne Ladar Imaging Testbed (Knowlton, 2011) after NGA certification in 2012) with a CE90 of 56 cm. The caveats (right) are necessary and appropriate and part of the policy and legal considerations as we approach and exceed human-scale absolute accuracies.

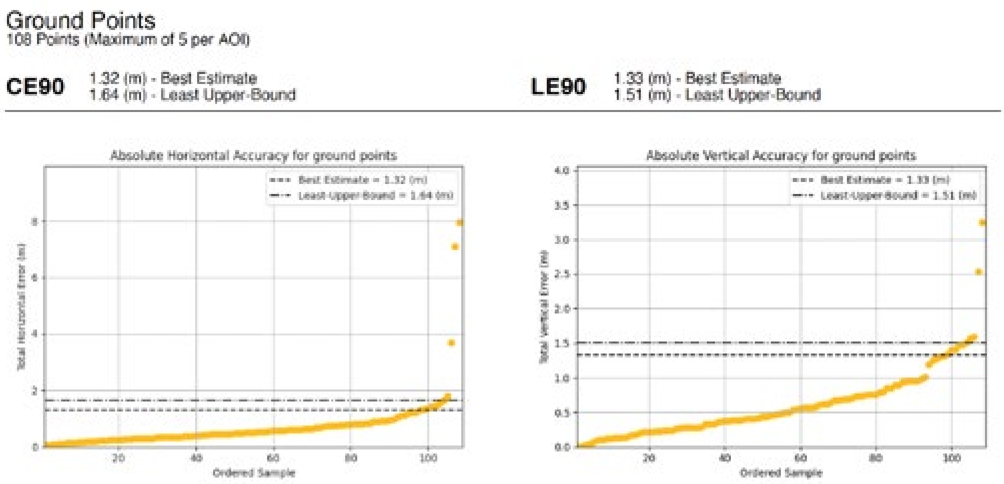


Figure 15 shows a global self-certification demonstration for Maxar's Precision3D data using 26 globally distributed airfields with NGA surveys as DGPS reference data following NGA's proposed SIG.0026 accuracy standard. Results are sub-sampled to avoid bias error, with five random samples combined to generate a statistical analysis of more than 100 individual measurements (Abrahamsen, Kozicki, and Kondas, private communication, 2023).

accurate are the data, and how current the data is[4]? The spatial and temporal resolution of remote sensing data and its accuracy are crucial considerations for decision-making, and thus, accurate data inspires confidence in decision-making. The Texas Calibration Network for Sensor Performance (TxCALIPER) will consist of a network of post-launch characterization and calibration facilities for optical imaging, synthetic aperture radar (SAR), laser altimetry data (lidar), Radio Frequency emissions, and other sensor modalities. TxCaliber will establish

4 Nothing is more useless than an out-of-date data point. In fact, it is potentially more dangerous.

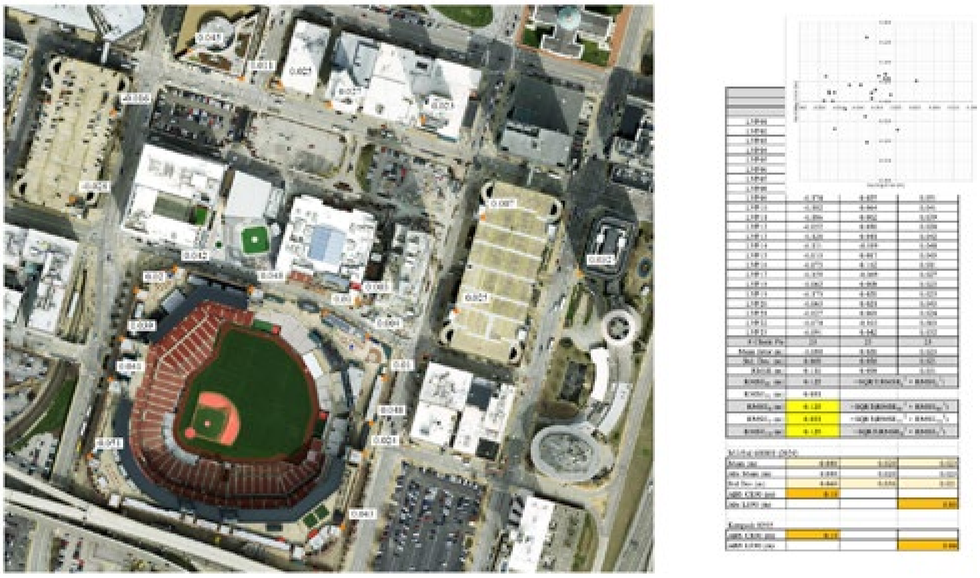


Figure 16 Self-certification and independent verification and validation of a commercial airborne Geiger-mode LIDAR system with a CE90 of 18 cm and a LE90 of 5 cm (Vecciho and Gillis, VeriDaaS, private communications, 2024).

performance benchmarks in spectral and spatial remote sensing c apabilities (accuracy and precision) across a broad reach of active and passive sensor technologies and for any platform vantage point. The TxCALIPER network will allow users/agencies/providers to answer four questions: what is the data's spatial, spectral, and temporal resolution, and how accurate are the measurements? The TxCALIPER Network will consist of multiple calibration sites distributed across the state of Texas, including a supersite, as well as smaller field sites that are well-instrumented and measured for biologic, geologic, and topographic applications. What makes this calibration supersite different from other calibration/validation sites around the world is (a) the inclusion of color targets for spectral calibration of optical instruments, (b) the inclusion of 3D targets for the validation of relative heights for validation of heights derived from photogrammetry, INSAR, or laser altimetry, and (c) absolute geolocation positions of targets. The ability to extract meaningful geospatial data from multiple sensors depends upon knowing the geolocation accuracy, as many remote sensing applications rely on both precise and accurate location data at human scales (Figures 17 and 18).

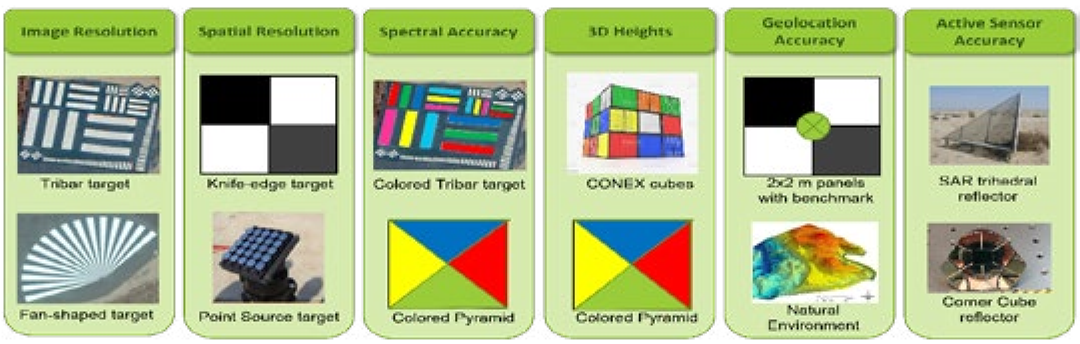


Figure 17 Conceptual targets for the TxCALIPER sites including traditional image calibration targets such as tri-bar, Siemens star, knife edge, and point source targets, including color, 3D, and SAR targets (Magruder et al., 2024).

Ultimately, calibration—the ability to characterize the spatial, temporal, and spectral resolution of sensors and data sets—is the cornerstone of unifying GKI and the location-based digital economy. Anyone who has used a smart device for navigation over the last two decades will recall its progression from useless to essential. As indicated above, the location-based advertising market is estimated to be USD 96B in 2022 and will likely grow to more than USD 110B by 2030. Returning to the theme of "what can you do with essentially nothing" there is a new concept of "trees as ground calibration points" that marries geodetic mapping with biomass conservation and global sustainability as illustrated in Figure 19.

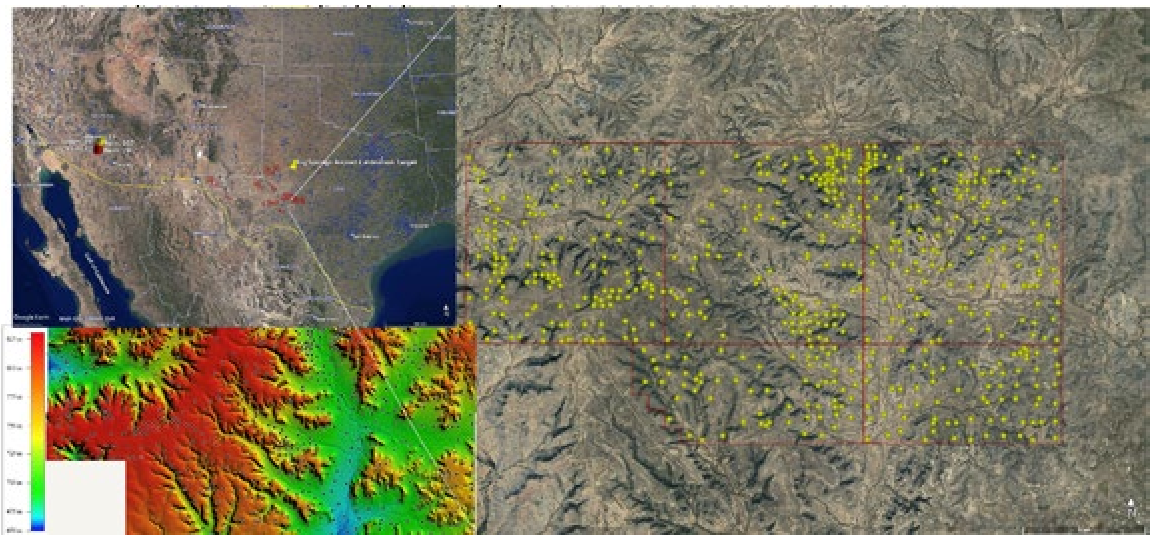


Figure 18 TxCALIPER will expand the concept of randomized sampling by utilizing existing Univ. of Texas properties with existing infrastructure, such as this 355 km2 area that contains 682 spatially and topographically distributed sites (Magruder et al., 2024).

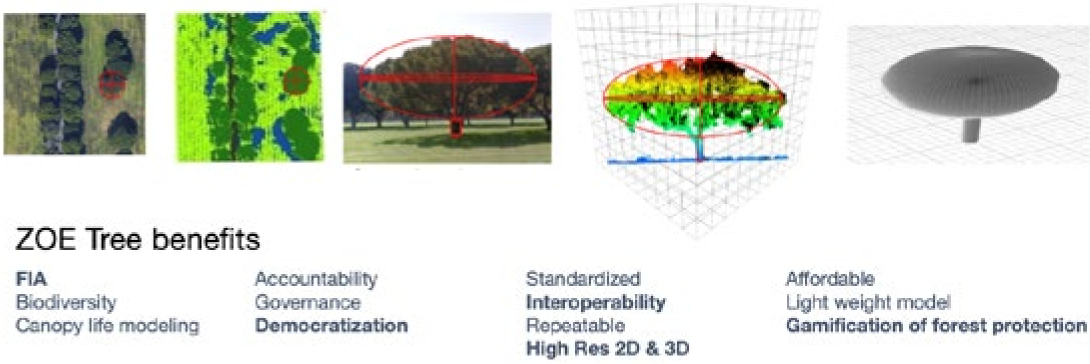


Figure 19 Trees as GCPs address multiple challenging problems with infrastructure for GKI integration for the NOHFO countries where infrastructure may be difficult to afford, but trees are plentiful. If we can establish trees as sustainable biomass and geodetic control points, we have another globally distributed resource to include in the marriage of GKI and the digital economy. (Kabachnik, Systems & Methods for Ground Truthing Forest Inventory Analysis, JACIE 2024)

11.5.1 Data Privacy, Data Ethics, and Cyber Security

As we open the can of worms, both crowd-sourcing and self-volunteered tracking data raise an inherent set of privacy concerns that warrant noting.

In 1914, Edmond Locard demonstrated that twelve points in a fingerprint were sufficient to identify an individual (Locard, 1914) uniquely. In 2013, a study of human mobility data established that human mobility traces are highly unique and that only four spatiotemporal points are sufficient to uniquely identify roughly 95% of individuals who move within an average radius of fewer than 100 kilometers (de Montjoye et al., 2013).

The concern is that even within anonymized datasets (without name, address, or phone number), it is possible to reconstruct human patterns of life across space and time and associate them with unique individuals. A candid discussion about "privacy" is appropriate, as commercial collectors are not inhibited by national laws and can "broker" data as the market desires and permits.[5] Figure 20 is typical of human trafficability data collected worldwide in 2024, enabling continuous surveillance and entity custody.

5 Tau, B., U.S. Spy Agencies Know Your Secrets. They Bought Them, WSJ, 8 March 2024.

Returning to the concept of resolution, different communities use different words to describe the process of turning a "picture into a story". The British define this function as image interpretation. To interpret is to curate the meaning of the image as story. This 4th dimension requires a degree of logical speculation on the "what" and the "why" - to test the premise - to fill out a chosen interpretive "story" of the activity so observed.

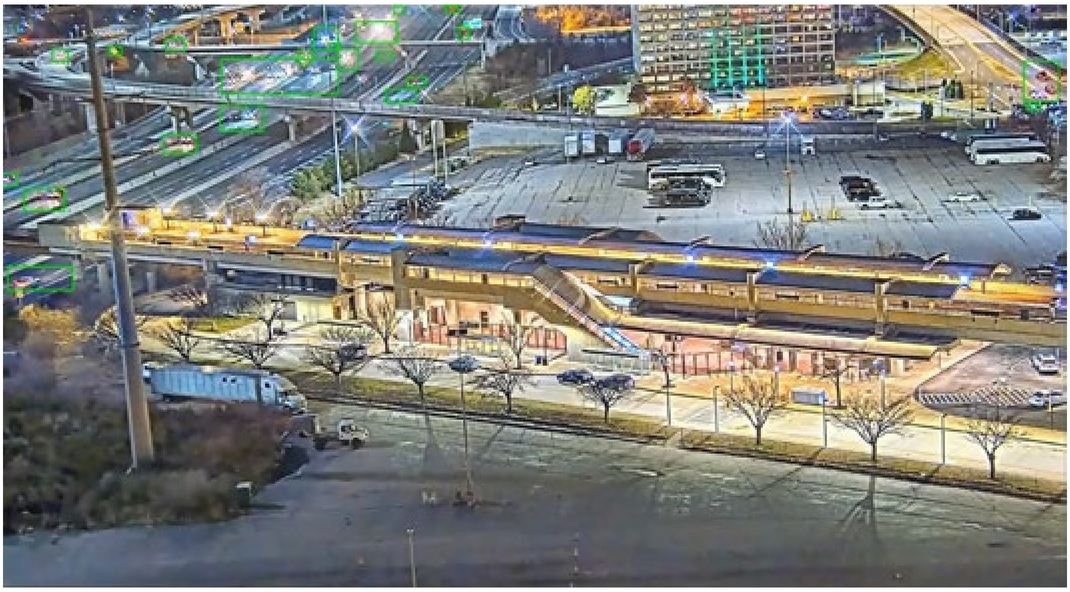


Figure 20 Typical high-definition video surveillance of urban areas used for machine learning, persistent surveillance, and entity custody.

11.5.2 Parting Observations

In this section, we have largely avoided a discussion of artificial intelligence (AI), focusing instead on finding common ground between the TOP-50 GKI countries and the NOHFO countries. With that as context, the technical members of the GKI community will have followed the explosive growth of Large Language Models (LLMs) and the proliferation of AI applications into the metaverse and digital economy that surrounds us all. At least three observations are appropriate.

As for ChatGPT and its counterparts, we routinely use ChatGPT for rapid code development, and the results are typically practical and useful at $1/10^{th}$ of the time it would take an average programmer to generate, test, debug, and document the equivalent code. In explaining Microsoft Excel macros and writing generic contract vehicles, ChatGPT can be efficient, effective, and delightfully valuable for a world where Google searches increasingly produce circular references.

In contrast, two far brighter community experts offer the following observations.

In March 2024, Gary Marcus, and Damion Hankejh "bet" Elon Musk $1M and $10M, respectively, that "humans will remain smarter than AI, at least in 2025." Marcus cites several things that AI probably will not accomplish in 2025, including:

- Reliably drive a car in a novel location;
- Drive off-road vehicles without maps, across streams, or around obstacles;
- Learn to ride a mountain bike off-road through forest trails; and
- Walk into random kitchens and prepare a cup of coffee or make breakfast.

Recently, in Nature Biotechnology, Jennifer Listgarten, in a correspondence entitled "The perpetual motion machine of AI-generated data and the distraction of ChatGPT as a 'scientist'", suggests the following:

- ChatGPT and its relatives will not themselves, anytime soon, be virtual scientists.

- AI can help us to understand the data we've collected and, with enough of it, to generalize from them, within reason. It can also help us to decide what to measure, initially or iteratively.
- In order to probe the limits of current scientific knowledge, however, we need data that we don't already have[6].

The author will cling to the observation that for many GKI experts in both the NOHFO countries and in the TOP-50 countries, using LLMs to accelerate coding and data integration may be a valid use of LLMs without discussing synthetic data and AI training on synthetic data.

In conclusion, the author offers the following metaphor: each new mapping project, much like a visit to a new country, starts with a first step. There is always a moment of hesitation and discomfort when arriving somewhere new, which is rapidly overtaken by the excitement of operating in a new environment. In each of these adventures, I am immersed in the human detail (and detritus) of the human imprint on the landform as the local community introduces their place – their digital ecosystem – into the global geospatial knowledge infrastructure. Today, there is no place on the planet where smart devices have not arrived and brought a digital economy and ecosystem; how GKI supports and integrates this digital economy is often hidden from the local community, although the citizens are the only ones who understand the validity of the local information. A colleague reduced this chapter to four maxims that span the process and partnership:

1. "Know what you don't yet know, but could & should find out next - but always measure and calibrate well" ...
2. "Understand the limits - gaps - in current sensor data interpretable as information in light of how it applies to useful automation" ...
3. "Only expect to apply what is proven repeatably under diverse - even controlled - environments (make your AI/ML robust)" - and ...
4. "Do the above with a "Do No Harm" feature built-in with a degree of error correction by automation (enable improved & fully robotic learning)."

Hence, the foundational concept of "self-calibration" establishes a geodetic "ruler" to ground everything and provide a local point of origin. Around this, we can integrate standards, policy, and partnerships, applying global best practices in GKI to essential local knowledge thereby creating the best possible final product.

As we establish a technical basis for implementing GKI strategy, we should not forget the legal and policy aspects of implementation – everything will boil down to a contract with terms and conditions that must be mutually understood. Hence, every technical requirement must be legally binding, and every contract term and condition must be measurable and accountable. Even before 1857[7], certain scientific communities could be divided into "lumpers and splitters" – people who aggregate and those who separate things, often with excessive attachment to the details. Ideally, as a global community of practice and 201 localities, we can find a common ground that places each location in a global context and reference frame at a human scale and respects the unique locality of that place and the people who have chosen to live there.

Subject to your questions!

11.5.3 Acknowledgments

The author acknowledges the continued collaboration with a group of peer reviewers with whom I have had the privilege to have an extended, multi-decade conversation about science, engineering, policy, and the universe. Their contribution is acknowledged and appreciated.

11.5.4 Competing interests

The author declares no competing interests.

6 This insight is as old as Archimedes, and it's likely older. It is why we have science and why it works. In a world of data-hungry large language models (LLMs), the value of new, unique, and provenanced data will remain critical and essential to machine learning.

7 Endersby, J., Lumpers and Splitters: Darwin, Hooker, and the Search for Order. Science 326,1496-1499(2009). DOI:10.1126/science.1165915

11.6 References

Abrams, M. C. and DePalacios, Absolute and Relative Positioning in the Third World, National Academy, 2012.

Abstract. In a peculiar fashion, the explosion of location-enabled cellular devices and the public availability of high-resolution base maps provide a new foundation for measuring and defining human geolocation and the space-time dynamics of human movement on the landform. Current GPS and precision navigation technologies dramatically extend the definition of location from the 1st world into the remote third world. With the deployment of DGPS control networks into the third world and the use of Satellite-Based Augmented System DGPS (SBAS DGPS), it is possible to investigate the potential for terrestrial and airborne mapping at human scales in real time. Combining geodetic networks with commercially available location-enabled devices challenges the notion that human-scale geolocation is only available to governments. It presages a convergence between location-enabled services and mapping the planet at human scales.

Location-based Advertising Market Size, Share & Trends Analysis Report by Component (Geofencing, Geotargeting), By Advertisement Type, By Promotion Type, By Application, By Industrial Vertical, And Segment Forecasts, 2023 – 2030.

Abstract. The global **location-based advertising market size** was estimated at **USD 96.39 billion in 2022** and is expected to grow at a compound annual growth rate (CAGR) of 15.1% from 2023 to 2030. The market is leveraging geographical data to deliver tailored content and offers to consumers. It has witnessed exponential growth due to several critical factors including the widespread adoption of mobile devices, advancements in geospatial technology, and the increasing demand for personalized marketing experiences. The ubiquity of smartphones has fundamentally changed how consumers interact with information and brands. The potential reach of advertising is extensive, as mobile devices serve as a primary conduit for accessing location data, enabling advertisers to deliver contextually relevant messages in real time. The continuous evolution of technology has been a major driving force behind the market's growth.

https://www.grandviewresearch.com/industry-analysis/location-based-advertising-market

Byars, David, D., THE CASAGRANDE-PHOTOGRAMMETRIC TEST RANGE, Proceedings of the American Photogrammetric Society, pp. 162–176, 1975.

Abstract. The Casa Grande Photogrammetric Test Range was established· in the mid-sixties to test the dynamic performance of aerial cartographic cameras. The test range design and the field surveys required to establish it are discussed. In addition, the photogrammetric procedures normally used to test aerial mapping cameras using the range are presented.

https://ia904602.us.archive.org/12/items/the-casagrande-photogrammetric-test-range/THE%20CASAGRANDE-PHOTOGRAMMETRIC%20TEST%20RANGE.pdf

Li, Chengmei & Tang, L. & Ma, Lingling & Zhou, Yongsheng & Gao, Caixia & Wang, Ning & Li, X. & Wang, X. & Zhu, X. (2015). A comprehensive calibration and validation site for information remote sensing. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-7/W3. 1233-1240. 10.5194/isprsarchives-XL-7-W3-1233-2015.

Abstract. As a naturally part of information technology, Remote Sensing (RS) is strongly required to provide very precise and accurate information product to serve industry, academy, and the public at this information economic era. To meet the needs of high-quality RS product, building a fully functional and advanced calibration system, including measuring instruments, measuring approaches and target site become extremely important. Supported by MOST of China via national plan, great progress has been made to construct a comprehensive calibration and validation (Cal&Val) site, which integrates most functions of RS sensor aviation testing, EO satellite on-orbit caration and performance assessment and RS product validation at this site located in Baotou, 600 km west of Beijing. The site is equipped with various artificial standard targets, including portable and permanent targets, which supports for long-term calibration and validation. A number of fine-designed ground measuring instruments and airborne standard sensors are developed for realizing high-accuracy stepwise validation, an approach in avoiding or reducing uncertainties caused from nonsynchronized measurement. As part of contribution to worldwide Cal&Val study coordinated by CEOS-WGCV, Baotou site is offering its support to Radiometric Calibration Network of Automated Instruments

(RadCalNet), with an aim of providing demonstrated global standard automated radiometric calibration service in cooperation with ESA, NASA, CNES and NPL. Furthermore, several Cal&Val campaigns have been performed during the past years to calibrate and validate the spaceborne/airborne optical and SAR sensors, and the results of some typical demonstration are discussed in this study.

https://www.researchgate.net/publication/277348236_A_comprehensive_calibration_and_validation_site_for_information_remote_sensing/citation/download

https://ceos.org/document_management/Working_Groups/WGCV/Meetings/WGCV-36/WGCV36_Presentations/CalVal_Sites-v2-submit.pdf

http://english.aircas.cn/ne/international/202002/t20200219_230041.html

Brumfiel, G., NPR, 2022, https://www.npr.org/2022/11/18/1137474748/trump-tweeted-an-image-from-a-spy-satellite-declassified-document-shows

Brumfiel, G., NPR, 2019, https://www.npr.org/2019/09/02/756673481/amateurs-identify-u-s-spy-satellite-behind-president-trumps-tweet

Langbroek, M., Twitter/X, 2019, https://twitter.com/Marco_Langbroek/status/1167809039872475136

Umbra images:

Santa Barbara, CA, https://umbra.box.com/s/9x8oe8li6b5smz2p5k5k4kgmva6zxifa, 34.42791, -119.895744

Dole Pineapple Maze, Oahu, HI, http://umbra-open-data-catalog.s3-website.us-west-2.amazonaws.com, 21.52502 -158.03756

Isfahan, IR, https://www.bbc.com/news/world-middle-east-68866548

F. Viallefont-Robinet, D. Helder, R. Fraisse, A. Newbury, F. van den Bergh, D. Lee, and S. Saunier, "Comparison of MTF measurements using edge method: towards reference data set," Opt. Express 26, 33625–33648 (2018).

Abstract. A sensor's spatial resolution has traditionally been a difficult concept to define, but all would agree that it is inextricably linked to the Ground Sampling Distance (GSD) and Instantaneous Field of View (IFOV) of an imaging sensor system. As a measure of the geospatial quality of imagery, the Modulation Transfer Function (MTF) of the system is often used along with the signal-to-noise ratio (SNR). However, their calculation is not fully standardized. Further, consistent measurements and comparisons are often hard to obtain. Therefore, in the Infrared and Visible Optical Sensors (IVOS) subgroup of the Working Group on Calibration Validation (WGCV) of the Committee for Earth Observation Satellites (CEOS), a team from various countries and professional entities who are involved in MTF measurement was established to address the issue of on-orbit MTF measurements and comparisons. As a first step, a blind comparison of MTF measurements based on the slanted edge approach has been undertaken. A set of both artificial and actual satellite edge images was developed, and a first comparison of processing results was generated. In all, seven organizations contributed to the experiment and several significant results were generated in 2016. No single participant produced the best results for all test images as measured by either the closest to the mean result, or closest to the truth for the synthetic test images. In addition, close estimates of the MTF value at Nyquist did not ensure the accuracy of other MTF values at other spatial frequencies. Some algorithm results showed that the accuracy of their estimates depended upon the type of MTF curve that was being analyzed. After the initial analysis, participants were allowed to modify their methodology and reprocess the test images since, in several cases, the results contained errors. Results from the second iteration, in 2017, verified that the anomalies in the experiment's first iteration were due to errors in either coding or methodology, or both. One organization implemented a third trial to fix software errors. This emphasizes the importance of fully understanding both methodology and implementation, in order to ensure accurate and repeatable results. To extend this comparison study, a reference data set, which is composed of edge images and corresponding MTF curves, will be built. A broader audience will be able to access the edge images through the CEOS CalVal Portal (http://calvalportal.ceos.org/). This paper, which is associated with the reference data set, can serve as a new tool to either implement or check, or both, the MTF measurement

that relies on the slanted edge method.

https://opg.optica.org/oe/fulltext.cfm?uri=oe-26-26-33625&id=402992

FAA Airport GIS site:

Abstract. Airports GIS Program. In support of NextGen, the FAA is moving to a Geospatial Information System (GIS). In FY 2007, we issued three Advisory Circulars to provide guidance for the collection and submission of aeronautical data and to identify the FAA's GIS data model for airport-related data. We will use the data to develop satellite-based approach procedures and to better utilize and manage the National Airspace System. More information about Airports GIS and related training can be found on the FAA's Airport Data and Information Portal (ADIP).

https://www.faa.gov/airports/planning_capacity/airports_gis_electronic_alp

FAA data interface:

https://adip.faa.gov/agis/public/#/public

Survey derived Universal Delivery Data Format (UDDF) v 1.07 for the airports:

https://nfdc.faa.gov/nfdcApps/services/publicData/uddfList.jsp

Three letter airport identifiers:

https://www.ngs.noaa.gov/AERO/annof_3.htm#f30

FAA AGIS site:

https://adip.faa.gov/agis/public/#/public

U.S. Bureau of the Budget. National Map Accuracy Standards, 1947.

Abstract. Establishes equivalence of circular error criteria as error standard of maps of various scales.

Greenwalt, C. and Shultz, M., Principles of Error Theory and Cartographic Applications, 1962.

Preface. Optimum utilization of ACIC research and production requires that the accuracy of source material, interim and final products be considered. The accuracy is expressed by an error statement which indicates whether the product is reliable and acceptable or should be used with discretion. Therefore, the error statement must be representative of the product and have a sound statistical basis. The purpose of this paper is to present and explain the theory and procedures for providing a valid and meaningful error statement.

Summary. Provides rigorous treatment of circular error, assuming that error is zero mean, normally distributed, and near circular.

https://apps.dtic.mil/sti/tr/pdf/AD0276978.pdf

Shultz, M., Circular Error Probability of a Quantity Affected by a Bias, 1963.

Abstract. A procedure for determining the radius of the 39.35%, 50%, and 90% probability circles for a biased distribution is presented. Both circular and elliptical, normal bivariate distributions are considered. The elliptical distribution is replaced by an equivalent circular distribution and an approximate radius obtained using the circular distribution procedure. Tables giving an indication of the error resulting from this replacement are included.

https://apps.dtic.mil/sti/tr/pdf/AD0644106.pdf

MIL-STD-600001. Mapping, Charting and Geodesy Accuracy, 1990.
Scope. This standard defines MC&G product accuracy and provides a common

basis for the appropriate application of these definitions. Adopts the 1963 Shultz approach to horizontal bias. Discusses an empirical approach as an alternative estimate.

Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy, 1998.

Objective. The National Standard for Spatial Data Accuracy (NSSDA) implements a statistical and testing methodology for estimating the positional accuracy of points on maps and in digital geospatial data, with respect to georeferenced ground positions of higher accuracy.

https://www.fgdc.gov/standards/projects/FGDC-standards-projects/accuracy/part3/chapter3

Ager, T., An Analysis of Metric Accuracy Definitions and Methods of Computation, 2004.

Abstract. This paper reviews the definitions and methods for calculating the absolute and relative accuracy values that are required in military mapping and targeting. It is stressed that the definitions of these terms are expressed very precisely in equations. Ultimately, it is the language of mathematics that defines what we mean when we say 'absolute LE90' or 'relative CE90'. Words alone are not sufficient, and we must take particular care in our definitions because the military mapping and targeting community uses statistics that are not generally known, and this can cause inconsistency in calculation and confusion in application.

ASPRS Positional Accuracy Standards for Digital Geospatial Data (EDITION 1, VERSION 1.0. - NOVEMBER, 2014)

Abstract. The objective of the ASPRS Positional Accuracy Standards for Digital Geospatial Data is to replace the existing ASPRS Accuracy Standards for Large-Scale Maps (1990), and the ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data (2004) to better address current technologies.

This standard includes positional accuracy standards for digital orthoimagery, digital planimetric data and digital elevation data. Accuracy classes, based on RMSE values, have been revised and upgraded from the 1990 standard to address the higher accuracies achievable with newer technologies. The standard also includes additional accuracy measures, such as orthoimagery seam lines, aerial triangulation ac- curacy, lidar relative swath-to-swath accuracy, recommended minimum Nominal Pulse Density (NPD), horizontal accuracy of elevation data, delineation of low confidence areas for vertical data, and the required number and spatial distribution of checkpoints based on project area.

https://www.asprs.org/a/society/committees/standards/Positional_Accuracy_Standards.pdf

de Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the Crowd: The privacy bounds of human mobility. Sci Rep. 2013;3:1376.

Abstract. We study fifteen months of human mobility data for one and a half million individuals and find that human mobility traces are highly unique. In fact, in a dataset where the location of an individual is specified hourly, and with a spatial resolution equal to that given by the carrier's antennas, four spatio-temporal points are enough to uniquely identify 95% of the individuals. We coarsen the data spatially and temporally to find a formula for the uniqueness of human mobility traces given their resolution and the available outside information. This formula shows that the uniqueness of mobility traces decays approximately as the 1/10 power of their resolution. Hence, even coarse datasets provide little anonymity. These findings represent fundamental constraints to an individual's privacy and have important implications for the design of frameworks and institutions dedicated to protect the privacy of individuals.

https://pubmed.ncbi.nlm.nih.gov/23524645/

Locard, Edmond, La Preuve Judiciaire par les Empreintes Digitales, Lyons. A. Rey. 1914.

https://babel.hathitrust.org/cgi/pt?id=hvd.32044116032152&seq=7

Il y a actuellement quatre-vingt-dix ans que Purkinje' a fait entrer dans le domaine scientifique la question des dessins digi- taux, jusqu'alors connus seulement d'une façon toute empirique et en dehors de la civilisation européenne. Malgré ce long temps écoulé, et malgré les nombreux travaux qui ont mis progressivement au point la technique des empreintes digitales, c'est dans ces toutes dernières années seulement que la preuve dacty- loscopique a été reconnue valable, et non sans lutte, par un certain nombre de Tribunaux. A l'heure où, malgré de surpre- nantes résistances, une jurisprudence commence à se créer en tous pays, j'ai pensé qu'il ne serait pas sans intérêt, pour les magistrats comme pour les criminologistes, de résumer en une brève étude synthétique les raisons sur quoi se fonde la valeur de cette preuve incomparable et les résultats qu'elle a fournis jus- qu'ici dans le domaine judiciaire et policier. Je me propose donc, après avoir rappelé les données physiologiques du problème, de discuter la nature et la portée objective de la preuve indiciale par les empreintes et d'en calculer les chances d'erreur, puis de passer en revue les jurisprudences comparées des divers Etats au point de vue dactyloscopique.

It is now ninety years since Purkinje' brought into the scientific field the question of digital drawings, hitherto known only empirically and outside European civilization. In spite of the long passage of time, and in spite of the numerous studies which have gradually perfected the technique of fingerprinting, it is only in the last few years that fingerprint evidence has been recognized as valid, and not without difficulty, by a number of Tribunals. At a time when, in spite of surprising resistance, a case law is beginning to be created in all countries, I thought that it would not be without interest, for magistrates as well as for criminologists, to summarize in a brief synthetic study the reasons on which the value of this incomparable evidence is based and the results it has provided so far in the judicial and police field. I therefore propose, after recalling the physiological data of the problem, to discuss the nature and objective scope of the indicial evidence by impressions and to calculate the chances of error, and then to review the comparative case law of the various States from the dactyloscopic point of view.

Knowlton, Robert, Airborne Ladar Imaging Research Testbed, MIT Lincoln Laboratory, Tech Notes, 2011.

ALIRT is an airborne three-dimensional (3D) imaging laser radar system that can rapidly collect high-resolution maps of wide-area terrain from altitudes up to 9000 m and decimeter accuracy from altitudes of 3000 m.

https://archive.ll.mit.edu/publications/technotes/TechNote_ALIRT.pdf

Knowlton, Robert, Airborne Ladar Imaging Research Testbed, Engineering, Environmental Science (2011).

Abstract: ALIRT is an airborne three-dimensional (3D) imaging laser radar system that can rapidly collect high-resolution maps of wide-area terrain from altitudes up to 9000 meters and decimeter accuracy from altitudes of 3000 meters. MIT Lincoln Laboratory's Airborne Ladar Imaging Research Testbed (ALIRT) produces images of wide areas quickly -- scanning 2000 sq km per hour -- and has numerous applications for both military and civilian users. ALIRT's ability to collect data from such great altitudes, coupled with its unparalleled data collection rate, enables the system to obtain large amounts of data in a single pass. Maps generated by ALIRT reveal details of both natural structures, such as the Grand Canyon, and man-made structures. These images could not be produced by commercial ladar systems. The Grand Canyon has a depth of nearly 2000 meters, which surpasses the maximum data collection altitude of other ladar systems. In general, data obtained using ALIRT enable answers to geometrical questions. For example, ALIRT enables battlefield situational awareness and reconnaissance, such as identifying landing zones or targets that are obscured; effective emergency response during and after natural disasters (e.g., by assessing viability of helicopter landing zones, migration of survivors, trafficability of routes); and production of detailed topographic maps quickly, further enabling land surveys, flood-plain analyses, earth-science research, and recreational activity such as hiking.

https://www.semanticscholar.org/paper/Airborne-Ladar-Imaging-Research-Testbed-Knowlton/49f67878cb22d4cc3d41d355eceeb42caebfbe1a

Ryan, Dorothy, Technology confronts disasters, MIT Lincoln Laboratory News, 2015.

In 2010, soon after Haiti was devastated by an earthquake, a team from MIT Lincoln Laboratory collected and analyzed information to help the U.S. Southern Command (USSOUTHCOM), the lead military agency responding to the crisis, effectively dispatch vital resources, including food, water, tents, and medical supplies, to the victims of this disaster. The Laboratory's capabilities in advanced imaging also aided relief operations: A laser-radar imaging system, the Airborne Ladar Imaging Research Testbed (ALIRT), produced high-resolution, three-dimensional renderings of terrain and infrastructure that were used to generate maps indicating road trafficability, helicopter landing zones, and the changes in populations at camps for displaced persons.

https://www.ll.mit.edu/news/technology-confronts-disasters

Clifton, William & Steele, Bradley & Nelson, Graham & Truscott, Antony & Itzler, Mark & Entwistle, Mark. (2015). Medium altitude airborne Geiger-mode mapping LIDAR system. Proceedings of SPIE - The International Society for Optical Engineering. 9465. 10.1117/12.2193827.

Over the past 15 years the Massachusetts Institute of Technology, Lincoln Laboratory (MIT/LL), Defense Advanced Research Projects Agency (DARPA) and private industry have been developing airborne LiDAR systems based on arrays of Geiger-mode Avalanche Photodiode (GmAPD) detectors capable of detecting a single photon. The extreme sensitivity of GmAPD detectors allows operation of LiDAR sensors at unprecedented altitudes and area collection rates in excess of 1,000 km2/hr. Up until now the primary emphasis of this technology has been limited to defense applications despite the significant benefits of applying this technology to non-military uses such as mapping, monitoring critical infrastructure and disaster relief. This paper briefly describes the operation of GmAPDs, design and operation of a Geiger-mode LiDAR, a comparison of Geiger-mode and traditional linear mode LiDARs, and a description of the first commercial Geiger-mode LiDAR system, the IntelliEarth™ Geospatial Solutions Geiger-mode LiDAR sensor.

https://www.researchgate.net/publication/283130841_Medium_altitude_airborne_Geiger-mode_mapping_LIDAR_system

Albota, M.A., Gurjar, R., Mangognia, A., Dumanis, D., & Edwards, B. (2017). The Airborne Optical Systems Testbed (AOSTB). https://apps.dtic.mil/sti/tr/pdf/AD1033558.pdf

Abstract. Over the last two decades MIT Lincoln Laboratory (MITLL) has pioneered the development of enabling technologies and systems for high-sensitivity, photon-counting, scanning three-dimensional imaging laser radar (3D ladar). Examples include the ALIRT mapping and MACHETE foliage-penetrating ladars. While these and other systems have been transitioned to operation, there is a need to maintain a testbed for novel phenomenology investigation and validation of new sensor architectures. To that end, MITLL has developed an airborne optical system testbed (AOSTB) that is re-configurable, allows for roll-on roll-off capability, and can accommodate multiple sensors on a low-operating-cost Twin Otter aircraft. AOSTB mission areas include wide-area down- looking high-resolution imaging, side-looking and up-looking laser ranging and tracking, and sensor fusion with EOIR cameras and hyperspectral payloads. We describe the AOSTB ladar system, present recently collected airborne down and side-look 3D data and discuss testbed configurations that can support various defense and non- DoD applications.

Marius A. Albota, Rajan Gurjar, Anthony Mangognia, Daniel Dumanis, Brendan Edwards; Contributed Review: Advanced three-dimensional laser radar imaging with the airborne optical systems testbed. Rev. Sci. Instrum. 1 October 2018; 89 (10): 101502. https://doi.org/10.1063/1.5017855

Abstract. We describe the development of a state-of-the-art airborne 3D laser radar system capable of area coverage rates in excess of 150 km^2/h. The ladar testbed can operate day and night and uses a low-power, eye-safe, and covert near-IR laser transmitter and a large-pixel-format photon-counting detector array coupled with a precision, fast-scanning, beam pointing mechanism. Mission areas include wide-area down-looking high-resolution open terrain and foliage-penetrating imaging, side-looking and up-looking laser ranging and tracking, and sensor fusion with electro-optical infrared cameras and hyperspectral payloads. We describe the airborne optical systems testbed ladar system, present recently collected 3D data, and discuss testbed configurations that can support advanced imaging applications.

Mehta, S., Peach, J., Weinert, A., To Expedite Roadway Identification and Damage Assessment in LiDAR 3D Imagery for Disaster Relief Public Assistance, Infrastructures, 2022, 7, 39, https://doi.org/10.3390/infrastructures/7030039

Abstract. Aerial surveys using LiDAR systems can play a vital role in the quantitative assessment of infrastructure damage caused by hurricanes, floods, and other natural disasters. GmAPD LiDAR provides high-resolution 3D point-cloud data which enables the surveyor to take accurate measurements of damages

to roads, buildings, communication towers, power lines, etc. Due to the high point cloud density, a very large volume of data is generated during an aerial survey. The data collected during the airborne imaging is post-processed with calibration, geo-registration, and segmentation. Albeit very accurate, extracting useful information from this data is a slow and laborious process. For disaster response, methods of automating this process have spurred the development of simple, fast algorithms that can be used to recognize physical structures from the point-cloud data that can later be assessed for structural damage. In this paper, we describe an efficient algorithm to extract roadways from a massive Lidar dataset to assist the Federal Emergency Management Agency (FEMA) in assessing road conditions as a step toward helping surveyors expedite a quantitative assessment of road damages for providing and distributing public assistance for disaster relief.

Marcus, G., Superhuman AGI is not nigh, Marcus on AI, 2024.

A few days ago Elon Musk claimed "My guess is that we'll have AI that is smarter than any one human probably around the end of next year…" Skeptical of this, I offered Musk a $1 million bet that AGI superior to the smartest individual humans would not arrive by the end of 2025, reprinted here for those who missed it. (The PS is important.)

https://garymarcus.substack.com/p/superhuman-agi-is-not-nigh?utm_source=profile&utm_medium=reader2

Listgarten, J., The perpetual motion machine of AI-generated data and the distraction of ChatGPT as a 'scientist', Nature Biotechnology, 42, 371-373, 2024

Correspondence. As a longtime researcher at the intersection of artificial intelligence (AI) and biology, for the past year I have been asked questions about the application of large language models and, more generally, AI in science. For example: "Since ChatGPT works so well, are we on the cusp of solving science with large language models?" or "Isn't AlphaFold2 suggestive that the potential of AI in biology and science is limitless?" And inevitably: "Can we use AI itself to bridge the lack of data in the sciences in order to then train another AI?"

I do believe that AI—equivalently, machine learning—will continue to advance scientific progress at a rate not achievable without it. I don't think major open scientific questions in general are about to go through phase transitions of progress with machine learning alone. The raw ingredients and outputs of science are not found in abundance on the internet, yet the tremendous power of machine learning lies in data—and lots of them.

https://doi.org/10.1038/s41587-023-02103-0

Acknowledgement

The Training Manual of the pre-conference **Training Program on Geospatial Knowledge Infrastructure for National Development** conducted in **Rotterdam, The Netherlands on 10-12 May 2024**, is the compilation and collaborative work of various thoughts and minds, who worked together with Geospatial World in bringing out this informative handbook. The final structure of the manual has been prepared under the thoughtful guidance of **Mr. Sanjay Kumar, CEO, Geospatial World**. A heartfelt gratitude to all the individuals, specifically the trainers, for their expertise, dedication, generosity, and collaboration in putting together this document.

We acknowledge, the trainers, **Prof. Josef Strobl, Mr. Alan Devenish, Dr. Joan Maso Pau, Dr. Rosemarie Mijlhoff – Portier, Gen. Girish Kumar, Dr. Kumar Navulur, Dr. Zaffar Sadiq Mohamed Ghouse, Dr. Siva Ravada, Prof. John Wilson, Dr. Mark C. Abrams and Ms. Ananya Narain** for sharing their insights and knowledge content.

We acknowledges, the unwavering efforts of **Dr. Shivangi Somvanshi and Mr. Ruban Jacob** in crafting the training program and its components.

Finally, we recognise the steadfast combined efforts of the GKI and design team of Geospatial World comprises of **Mr. Subhash Kumar, Mr. Pradeep Chauhan, Mr. Saurabh, Ms. Shikha Sonam, Ms Maitreyee Chauhan and Mr. Yagyik Pant** in bringing all insights together, compiling and neatly packaging the report for the readers.

Produced by

GEOSPATIAL WORLD

ADVANCING KNOWLEDGE FOR SUSTAINABILITY