Application form

National Roadmap for Large-Scale Research Facilities 2011

General information:

Kind of proposal *(please choose only one type of proposal)*

New facility (not in NL Roadmap 2008)

□ inclusion in the roadmap only (no request for funding) **(type 1)**

X inclusion in the roadmap and request for funding **(type 2)**

Facility from NL Roadmap 2008

□ progress report and request for funding **(type 3)**

□ progress report only (no request for funding) **(type 4)**

Main applicant: 1. TUD

Title(s)	Prof. dr. ir.
First name	Peter
Initials	P.J.M.
Surname	van Oosterom ☒male ☐female
Address for correspondence	Technische Universiteit Delft, OTB Jaffalaan 9, 2628 BX Delft
Telephone number	015 2786950
Fax	015 2784422
Email	P.J.M.vanOosterom@tudelft.nl
Website (optional)	www.otb.tudelft.nl

Co-applicants 2. Alterra

Title(s)	Ir.
First name	Jandirk
Initials	J.D.
Surname	Bulens ☒male ☐female
Address for correspondence	Alterra Droevendaalsesteeg 3, 6708 PB Wageningen
Telephone number	0317 480700
Fax	0317 419000
Email	Jandirk.Bulens@wur.nl
Website (optional)	http://www.alterra.wur.nl/

3. DANS

Title(s)	Dr.
First name	Peter
Initials	P.K.
Surname	Doorn ☒male ☐female
Address for correspondence	Data Archiving and Networked Services (DANS) Anna van Saksenlaan 10, 2593 HT Den Haag
Telephone number	070 3446484
Fax	070 3446482
Email	peter.doorn@dans.knaw.nl
Website (optional)	http://www.dans.knaw.nl

4. Geonovum

Title(s)	Drs.
First name	Rob
Initials	R.J.
Surname	van de Velde ☒male ☐female
Address for correspondence	Geonovum Barchman Wuytierslaan 10, 3818 LH Amersfoort
Telephone number	033 4604100
Fax	033 4656457
Email	r.vandevelde@geonovum.nl
Website (optional)	http://www.geonovum.nl

5. NLR

Title(s)	Dr.		
First name	Rob		
Initials	R.W.		
Surname	Van Swol	☒male	☐female
Address for correspondence	NLR Voorsterweg 31, 8316 PR Marknesse		
Telephone number	088 5114252		
Fax			
Email	Rob.van.Swol@nlr.nl		
Website (optional)	http://www.nlr.nl/		

6. UvA

Title(s)	Prof. dr.		
First name	Tom		
Initials	T.M.		
Surname	van Engers	☒male	☐female
Address for correspondence	Universiteit van Amsterdam Vendelstraat 8 (BG13a), Amsterdam		
Telephone number	020 5253494		
Fax			
Email	vanEngers@uva.nl		
Website (optional)	www.LeibnizCenter.org		

7. UT/ITC

Title(s)	Dr. ir.		
First name	Rolf		
Initials	R.A.		
Surname	de By	☒male	☐female
Address for correspondence	Universiteit Twente, ITC Hengelosestraat 99, 7514 AE Enschede		
Telephone number	053 4874553		
Fax	053 4874335		
Email	deby@itc.nl		
Website (optional)	http://www.itc.nl		

8. UU

Title(s)	Prof. dr.		
First name	Steven		
Initials	S.M.		
Surname	de Jong	☒male	☐female
Address for correspondence	Universiteit Utrecht Heidelberglaan 2, 3508 TC Utrecht		
Telephone number	030 2534050		
Fax	030 2531145		
Email	S.M.deJong@uu.nl		
Website (optional)	http://www.geog.uu.nl		

9. VU/EduGIS

Title(s)	Prof. Dr.	
First name	Joop	
Initials	J.A.	
Surname	van der Schee	☒male ☐female
Address for correspondence	Vrije Universiteit Amsterdam / EduGIS, Onderwijscentrum VU De Boelelaan 1105, 1081 HV Amsterdam	
Telephone number	020 5989213	
Fax		
Email	j.a.vander.schee@vu.nl	
Website (optional)	http://www.edugis.nl/ http://www.feweb.vu.nl/gis	

10. WU

Title(s)	Prof. dr. ir.	
First name	Arnold	
Initials	A.K.	
Surname	Bregt	☒male ☐female
Address for correspondence	Wageningen Universiteit Droevendaalsesteeg 3, 6708 PB Wageningen	
Telephone number	0317 481699	
Fax	0317 419000	
Email	Arnold.Bregt@wur.nl	
Website (optional)	http://www.grs.wur.nl	

maps4science

Abstract

Summary

The Netherlands is historically known as one of worlds' best-measured countries. It is continuing this tradition today with unequalled new datasets, such as the nationwide large-scale topographic map, our unique digital height map (nationwide coverage; ten very accurate 3D points for every Dutch m^2) and a range of public and private collections of environmental and socio-economic geo-datasets. Despite the wealth of existing and well-maintained geo-information (GI), the access to GI for our academic community is rather poor. In particular, science domains discovering the enabling power of GI have difficulties in accessing and using it. Initial analysis shows that i) legal and commercial licence restrictions, ii) technical inability to deal with more advanced types of spatial data and iii) lack of awareness are the top three barriers preventing large uptake in science. This is a missed opportunity to achieve scientific breakthroughs in a wide range of scientific disciplines.

Nowadays, science groups gain competitive advantage through using the ICT infrastructure in a creative and productive way. As 80% of all data has a spatial component, the emerging availability of GI and the growing ability for real time monitoring make it a very powerful enabler in many scientific fields. Maps4Science will provide scientists with the required geo-information and associated tools and services for discovery and analysis. It will provide a geo-ICT infrastructure to allow unrestricted experimentation, linking and discovery of information, as much as possible.

Setting up such a facility requires a combined effort of several universities, research organisations and data providers. The development of this large-scale facility and its operation for 5 years requires a total budget of 18,1 M€ from NWO, complemented with at least 25% of own contributions (today already spent on sub-optimal local facilities), possibly much more if the value of the geo-data provided and maintained by partners is considered. Maps4Science will provide the science community with an adequate infrastructure overcoming the main barriers: accessibility, handling of huge spatial datasets and awareness. It will perform research to create and renew the facility in the fields of: Digital Rights Management – in particular related to the issues regarding geo-information; Spatial Cyber Infrastructure (SCI) – related to handling huge datasets and services for processing spatial data; Science-with-GI – addressing the science cases of various academic domains (provide SCI requirements and validate if these are met). In addition, the facility will study the SCI itself as a scientific field in which the Netherlands have achieved a Top-5 position worldwide which will be reinforced. The SCI research results will be implemented in the operational facility.

The facility will need 1,5 year initial stage to start and develop its required SCI. Operational stage will be reached by the end of 2013, marking the start of a 5 year operational stage. In its 5 year operational stage the facility will develop science and talent cases that will exploit the abilities of this facility in creating science that makes a difference. The consortium includes the participation of university libraries, DANS, and main data producers. They will play a role in continuation of the facility after its conclusion. The living lab will consist of a community construction with public and private bodies to recognize and develop potential valorisation aspects. This may include participation in EU calls for FP7 and FP8 or ICT Research. The facility will require a budget of 22,8 M€ of investments, operational costs and additional research. Consortium partners will bring in 6,1 M€ of own contribution.

Summary of the research proposal in layman's terms

Maps4Science provides the Dutch academic community easy, fast, cost-effective access to the wealth of spatial data (also known as geo-information - GI) that is available, both earth and socio-economic data including multi-temporal and level of detail support. GI is an enabler for new research in domains like biology, medicine, architecture and history. The science case is the integral of mono-disciplinary sciences cases in addition to setting up a unique science programme for studying and improving the facility.

Key words

Location, Geo-data, Geo-services, spatial cyberinfrastructure

Research proposal

Detailed description

"Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets. The speed at which any given scientific discipline advances will depend on how well its researchers collaborate with one another, and with technologists, in areas of eScience such as databases, workflow management, visualization, and cloud computing technologies."[1]

1 Science case

1.1 Introduction

Location information is a key enabler for integration and analysis of data in many science disciplines now. Dutch scientists have a world top position and also public and private achievements in geo-information infrastructure and applications are renowned. But the cascade of techno-social developments coming from ICT, its uptake and use in society require us to deal with the impact these developments have on the geo-information infrastructure. The scale and variety of issues demands a joint effort between science partners in the Netherlands. The geo-information sciences in several universities and research institutes want to collaborate in a joint large-scale research facility to keep pace with the external developments confronting us and to maintain and improve our leading position in the world. In addition, the growing use of location information in other science disciplines show how an effective geo-information infrastructure can contribute or enable new scientific breakthroughs in many fields. A joint facility allows the Dutch science community to pick-up these enabling methods and stay or become leading in its application.

It goes both ways. While location data makes many mobile applications possible, the impact of the ever-expanding information society on geo-information sciences is enormous. Recent analysis of the KNAW/NCG (Royal Netherlands Academy of Arts and Sciences/Netherlands Geodetic Commission)[2] identified the following trends with respect to geo-information use and science:

- *From practice to theory:* The early GI-research questions were mainly derived from application domains and rather technical in nature, but over the years an own GI body of knowledge started to develop. Current GI theoretical questions deal with spatial scaling, the space-time description of spatial phenomena and processes, spatial perception of humans, spatial ontologies, etc.
- *From GI application to geo-information infrastructure:* A few decades ago a GI-application was a combination of data and a software system for the support of a specific question. Since the beginning of this century a world-wide paradigm shift took place. Spatial data is increasingly organized in the form of local, national en international infrastructures (GII) that support many actual and potential applications. This has resulted in GII research which focuses on spatial data and technical interoperability, standards, policy, organization issues, assessment frameworks, etc.
- *From spatial data structuring to meaningful spatial data integration:* Data structures and the efficient algorithms for storing and retrieving data was a key research activity in the eighties and the nineties of the previous century. Now the emphasis has shifted towards meaningful exchange and integration of spatial data. The concept of spatial data ontologies as a potential solution to this challenge is emerging research topic world-wide.
- *From mapping to dynamic real-time spatial data collection and visualization:* The 2D static map was and probably still is the dominant way of obtaining and presenting geo-information. Fast developments in sensor technology and visualization techniques induce a shift to dynamic real-time spatial data collection and the direct use of these data in process models and in visualizations. Research on how these new approaches can be used to obtain reliable information on spatial phenomena and how these data can be used by public, governments and business is still in its infancy.
- *From technological to socio-technical:* Originally GI research was quite technical in nature, however with the strong diffusion of the GI technology in society it is also becoming a research area for policy-, organization- and

[1] *(The Fourth Paradigm: Data-intensive Scientific Discovery, http://research.microsoft.com/en-us/collaboration/fourthparadigm/).*
[2] NCG – GII subcommission report 2011 (in prep.).

law researchers. Also the functioning of the GI setting within a society has become a research topic. From a mainly technical research field is has developed into a socio-technical research field. Appropriate scientific methods that can be applied for this new setting still need to be developed and tested.

- *From a few application areas to many disciplines in society:* One of the major developments the last 20 years is probably the strong increase in disciplines where GI is used. The classical domains for GI are agriculture, spatial planning, environment, land registration and transportation. Now, GI has found its way into almost all disciplines in our society, ranging from history to medicine and banking and tourism. This has resulted is a strong demand for "with GI" research. A key question is how can GI concepts and technologies be beneficial to the discipline.

In order to frame these trends we derive two dominant axes, namely the USAGE axis and the DATA SOURCE axis. On the usage axis, the trends indicate the evolution of how maps prepared for dedicated users became interesting information products to other users. At a certain stage, the users were unknown to the developers of geo-information and the usage was sometimes very surprising. On the sources axis, the trends indicate a movement from authoritative single sources, to linked sources of all kinds of origins, from static to dynamic, real-time, authoritative and non-authoritative etc.

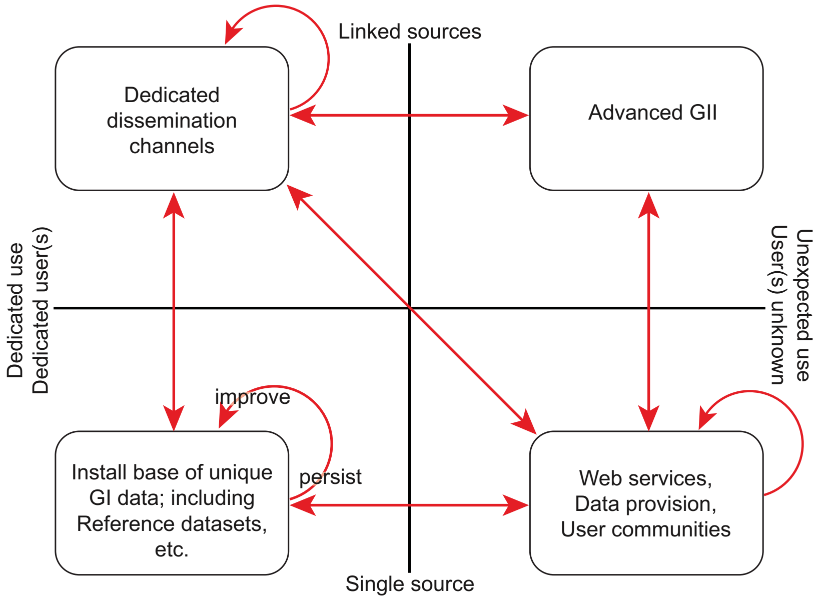


Figure 1: The geo-information source and the geo-information usage dimensions.

We have now come to a point where we have to deal with an alignment and amplification of developments. Our IT architecture and the way we handle data has evolved from site-based to distributed and cloud-based. Our users are evolving from the 'in crowd' of trained professionals to a potentially huge community including professionals in other domains and even the citizen. Our data policies are evolving from proprietary and protective, to open and shared. Location data has become a new enabler for integration of all kinds of datasets. The pay-off of MapInfo company *"Knowing where is just the beginning..."* simply illustrates this principle.

In this changing context where location has become an important enabler for new scientific methods not in the least for integrating sources there is a loud call for new answers. The key to the future lays in 'Linked Sources for Unexpected Usage'. In the Netherlands, being a country with one of the highest densities of geo-information per square meter (due to the dense use of space in our country and the advanced GI sector), this is more eminent then in other countries. Our great install base of data, skills and usage as a result of government will and action to deliver a National Geo Information Infrastructure puts us in the front seat to discover also the potential drawbacks our good initiatives run into in about 5 – 10 years ahead. In order to maintain our world leading position and to lead Europe into the new era of spatial information society we need to cooperate, collaborate and invest in scientific breakthroughs supporting the challenges ahead. Therefore we need a large-scale research facility that allows our universities to join efforts and experiment at the scale of our national infrastructure. We need more, better and integrated maps for scientific breakthroughs. Maps4Science.

1.2 Maps4Science: a large-scale research facility for GI sciences

The mission of Maps4Science

The mission of Maps4Science is to facilitate and enhance GI sciences and the enabling role of GI in other science disciplines. Maps4Science is a virtual facility of collaborating universities and research institutes active in the field of location information and its application in society. Maps4Science is an infrastructure that provides scientists with a wide variety of resources and a set of core services for discovery, exploration, analysis and visualisation of location-based information.

Facility outline

Combining efforts allows the Dutch GI science community to build an cyberinfrastructure for spatial data at a scale, large enough to deal with the complexity and size of the growing amounts of data, to link and integrate these data and to create services for all kinds of processing required. First of all, this will facilitate the GI science community in its pursuit of dealing with the developments of the still expanding information society. The facility itself and its usage provide experiences and measurements relevant for similar initiatives in the public (and private) domain. Setting up such a facility also boosts research in understanding and designing how to deal with distributed, linked sources in terms of rights management, ICT and its particular application in the GI domain. Secondly, making available these large amounts of data to the whole science community will serve as an enabler for new and innovative science practices in many domains. This is supported by a small (non-exhaustive) set of use cases mentioned in this proposal. The enabling role in other science disciplines, and its consequences in terms of requirements and pitfalls is also subject to scientific advancements.

The facility will bring benefits to: many types of science using the facility within their discipline (and maybe even partly changing the way they conduct research and opening avenues to new results), GI scientist will be in the position to develop and further improve the facility up to the level of world leading quality, with the facility potentially growing into a EU-level facility (scientific complement of the INSPIRE directive). In the annex @@ a selection of relevant facilities is described, supporting the potential of the benefits.

Maps4Science is a large-scale research facility consisting of a number of components organized in 3 building blocks. The bottom block is the infrastructure consisting of a poule of data sources, an (additional) own storage body and a set of core services for disclosing and processing purposes. The second block describes the scientific improvements sought, in 4 categories. The third block is the Living Lab approach, where the cyberinfrastructure and the developed science results will thrive in a collaborative innovation setting with public and private partners.

The Infrastructure Block

The dissemination of the data is the heart of Maps4Science. Where possible existing infrastructures will be reused SURFnet, the University libraries / repositories and national and international collaborations. Maps4Science does not aim to establish a redundant portal, but rather make use of existing networks for distributed data management. From day one of the GOF the store is open (based on state-of-the-art propositions within the existing facilities), but this will continuously be improved and expanded based on the experience and needs of users / researchers. The GII Maps4Science part is for the very wide range of researchers who conduct research *with* geo-information (supporting role).

One of the main assets of Maps4Science is its disclosure of existing data to be used in research. To this end, agreements are made with owners and providers of data. Despite the increased move towards Open Data, scientists will be interested in datasets with restricted access or use (consider data under military governance). Therefore one key topic in this is the digital rights management. In addition to disclosing external sources, the facility will make available datasets itself. This may include derived data, model-output, experimental data or archived data elsewhere no longer available. The third component of the infrastructure is the services needed to use the datasets. The infrastructure also involves a dedicated group of staff members, responsible for operating and maintaining the facility.

Unimpeded access to datasets (possibly life, in the form of sensor networks) is not obvious. All over the world there are investigations into various business models and their legal and economic consequences on data use shifts in society and market. The target group of academic researchers may be more demanding than their counterparts in practice (government and industry): access to raw data, unfiltered and (very) large data quantities (eg time series of satellite data, or sensor network data), multi-scale consistency (perhaps even resulting into vario-scale models),

links between interpreted (map) objects and source (survey, measurement) data, feedback and adding data (the facility provides both services to obtain data and add data to the GII, e.g. the results of a simulation / model), extremely massive data sets (point clouds AHN2 accessible), etc., etc.
It is of great importance to connect to the current developments (taking place the research community: the Dutch 'basis registraties', 'Publieke Dienstverlening Op de Kaart' (PDOK), Shared Service Organization geo (SSO geo), international standards / INSPIRE. This requires an appropriate architecture of this infrastructure (centralized / decentralized, connecting). An important role is also provided for the 'information brokers' in the research world, such as DANS, Surfnet, and various university libraries / repositories.

The Science Block
The Maps4Science also consist of a GI-research program to further improve geo-information infrastructures to the benefit of other scientific disciplines that use geo-data in a more advanced manner. The GI-research itself is closely related to this unique large research facility. It offers GI-science researchers within the Netherlands the important opportunity to investigate and improve the structure, its functioning and impact. The benefits are mutual, because after a successful test phase, a specific GII innovation will be transferred to the production environment of the Maps4Science GII.
The Geo-Information Infrastructure (GII) is continuously developing and changing due to new user requirements, technological opportunities, changes in the legal field and commercial considerations of resource owners. On the functioning of a GII is still little known (which legal, organizational and technical controls should we turn, to obtain progress?) The consequences of various choices are barely explored, e.g. what is the impact of privacy legislation using the system of key registers ('basisregistraties')? What are the limitations of the current state of the art GII and which areas to improve: handling very large datasets (terabytes to get users, ICT infrastructure aspects), concepts multi-/vario-scale (up to 5D) geo-information, linking geo-information with the original (and possibly life) sources, semantic techniques for better (more meaningful) searching and translation answers tailored to user interface with (multidisciplinary) spatial process models and simulations, etc. The GI-Science program of Maps4Science is meant for research into the geo-information (infrastructure) itself.

The Living Lab
The scientific achievements are important for innovation in the public and private domain. Maps4Science plays a role in the Golden Triangle by creating a Living Lab in which researchers, in collaboration with government and business transpose the scientific research results in tools and approaches that directly contribute to actual problems in society and markets. Application in society is not an end point, but the starting point for demand. Maps4Science Living Lab can be a game changer in getting the spatial component into the base solving strategy of policy problems, emergency response and the major infrastructural choices to be made in The Netherlands. It is an important step to realize the necessary transformation of the current supply-oriented government to a more demand-driven government, facilitated by tools developed in the Living Lab.

Yet another GII ?

You may think: this is just another GII. What is new? The Maps4Science facility is indeed a GII. It will serve as a scientific exploration platform to deal with the datasets, user and processing volumes and related issues of the coming years. Maps4Science is developing a GII that can handle the massive amounts of data currently not handled by existing GIIs. For instance, the up-to-date height model of the Netherlands (ANH2) consumes 15 TByte of storage. One year of cyclorama images consumes 30 TByte and the intended annual aerial photo flights will result in 160.000 images of 400 MB (approx. 64 TByte) for the stereo photos and another 5 TByte for the orthophotos. And these are the authoritative datasets. In addition, we see enormous amounts of data coming from sensor networks, non-authoritative sources like citizens with mobile phones, mobile applications used by citizens etc. etc. At this moment, we can only guess about the amounts of data. But for sure, the current GII has neither means nor intention to handle it.
Another issue for research that may lead to an improvement of the current GII is the semantic capabilities of the services and datasets. In order to facilitate broad groups of users it is eminent to align to their ontologies and jargons. This will also facilitate unexpected use and allow new, unknown users to deploy location data.

Maps4Science is not an experimental lab. It is first of all a production facility to serve the science community. It will not replace existing infrastructures but it will connect and extent these facilities. The extension will be an open infrastructure, but primarily dedicated to the scientific users and usage.
Needless to say, scientists are not common users. Besides the presented data, scientists can be interested in the original measurements and source data, to study or evaluate the way datasets are produced. Scientists are interested to share experimental data and simulated data and have no means in the authoritative GIIs to accomodate that. And last but not least, the facility allows scientists to study the legal consequences of the opportunity to link in itself innocent data and discover relationships that might infringe privacy issues.
In short, Maps4Science is a facility to gain experience and measurements on future use of next generation cyberinfrastructures. In it is appearance, scale and shape, Maps4Science is a great way to stay on top of developments in the GI sciences and to maintain the leading position of Dutch GI scientists.

1.3 The science cases

Maps4Science has 4 distinct science cases, each of which will be described in further detail below:

1. *Geo-Information Science case*: A Geo Information Infrastructure as will be developed in this facility is an important instrument in GI Science. The facility is instrumental in this science fields;
2. *Science with geo-information case*: location information is a key enabler for innovative research in several science disciplines, in earth sciences, humanities, social sciences and arts;
3. *Spatial Cyberinfrastructure science case*: Through its nature, spatial data and in particular processing it, requires special cyberinfrastructure aspects. The Maps4Science facility will be used to develop these specific cyberinfrastructure;
4. *Digital rights management science case*: Through Maps4Science datasets can be integrated using the location as linking pin. This may cause unexpected infringements of privacy and data use. With the increased use of location data in society this requires dedicated research for legal.

1.4 Science case 1: Geo-Information Science

The Maps4Science facility aims to facilitate a broad spectrum of the Dutch research community in the domain of spatial information, aiming not only at earth science researchers but also to support researchers in other fields of science, such as the life, social and medical sciences, in which location intelligence is or may become important.
GI is a container term for a rather peculiar type of scientific data with the common characteristic of being location-bound, typically through earth coordinates. It encompasses the acquisition of raw data through a wide variety of sensors and survey techniques, the production of data, their data management, and the dissemination to end-users in a wide variety of formats, including of course maps. Given that peculiarity, the use of GI products is often non-trivial, and requires special tools and techniques. This is even more the case when multiple sources need to be recombined into an information end product. Data analysis and data integration tools are similarly complex to operate, and should be viewed as resources themselves. The interoperation of data and computational resources requires them to be properly described and characterized.
From the perspective of use and users, a user needs analysis will have to be conducted. The human interface to the facility will require smart and efficient filtering techniques as well as interactive functionality for visual exploration and analysis, allowing the gradual fine-tuning of a data search. Such an interface explicitly includes the typical components for map production, and may require the development of a map specification language.
The sketched broad spectrum of use, users and resources poses enormous challenges to the resource management. This is typical for scientific data management, but in addition, it carries the characteristics of spatial data and its use. Below, we propose an integrated body of collaborative research themes to address these challenges. The proposed work has potential of scientific breakthroughs for the field of GI-Science itself.

1.4.1 Architecture, resources and standards

Under this theme, we will define and develop an abstract architecture fitting with scientific data architectures already in place and to be used, while at the same time making the fit with architectures developed in the domain of GII. These have matured in recent years. Such is necessary because of the special nature of spatial data, and the need to comply to global standards in this domain, especially those by the Open Geospatial Consortium (OGC). For

example, precise standards are available for spatial feature, map and catalog services. The challenge for this research theme is in the merge of those architectural styles, while retaining the simplicity required to achieve scalability and genericity in user and programmatic interfaces to the resource facility.
We believe that an architecture in the spirit of the resource-oriented architecture (ROA) associated with pure RESTful services to serve up and manage the data and computational resources of the facility would be a good candidate. But this is an untested hypothesis. This research theme directly links to the theme below, and is also strongly related with the work on the Spatial Cyberinfrastructure component, which it will look at as the enabling technological substrate on which the specific GI functionality can be based.

1.4.2 Usability and dissemination modes

The ambition to serve a large and diverse user community necessitates a full needs analysis, including a complete use case run-down, especially aiming to clarify fitness-for-purpose for users less initiated to GI. This in itself requires the application or development of paradigms known to or intuitive for that user community. Such paradigms must be prepared with the various modes of resource dissemination in mind. Dissemination in Maps4Science will come in two flavours: through *human* interfaces, and through *programmatic* interfaces. Orthogonal to this stands the purpose of interaction: search, explore, fetch or push. The first three are especially important in the human interface(s), especially for the less GI-savvy user. In this domain, the paradigm of National Atlas (online, and highly interactive) appears truly applicable.
We propose here the development of a map specification language that is constraint-based, and that will allow underspecified maps This will allow end-users to ‘play’ interactively with such unresolved parameters, and refine their map product in stepwise fashion.
On the programmatic interfaces side, we will aim to develop a highly uniform style of API-based on principles like those of the http protocol. This is needed to allow the standardized development of client applications to the facility, as well as the adaptation and use of already existing applications. The innovation and challenges in this research theme derive from the characteristics of the end-user community, and the long and growing list of spatial data types. In combination, they ask for simple and intuitive mechanisms that can stand the test of time with upcoming, at present unknown data types.

1.4.3 Management of very large data sets

Disk storage has followed Moore's Law (storage capacity doubles every 1-2 years) for the last 30 years. A dataset that was huge 10 year ago now easily fits in your pocket on a memory stick. In some cases however the size of the dataset grows even faster than Moore's law making the management the dataset increasingly challenging. This is the case for some very large datasets in the Netherlands:

- The AHN dataset (a Dutch height map of the Netherlands acquired by laser scanning from the air) was first collected between 1996 and 2003 with a density of one measurement per 16 square meters. In 2008 a new cycle of measurements was started (AHN2) with a density of about 10 accurate measurements per square meter. With a dataset of this density it is possible to reconstruct buildings. Management of the AHN2 dataset of 15 TB is already quite a task with current technology and the integration of this data with ground-based laser scanning will be a real challenge. A similar situation occurs under water with the huge multi-beam echo’s that produce large bathymetric point cloud data [NCG [3]]
- With aerial imagery the data size not only grows because of the increasing image quality (7 cm pixels) but also because pictures are taken more and more often. In addition this, also historic archives are created and often very relevant in a research context.
- Another area where the size of the data is growing beyond Moore's Law is with crowd-sourced data: Millions of people have mobile phones, or GPS devices that track their location. In this case the challenge is not in the size of an individual track (which can be significant in itself by producing a point every 1 or 2 points over a larger time interval), but in the sheer number of tracks that is collected.

The challenging size of these datasets together with the unpredictable way data is used in a research environment results in the following research topics:

- Integrated storage model;

[3] http://www.ncg.knaw.nl/eng/publications/Green/49VanOosteromPointClouds.html

- Architecture for disk-intensive and cpu-intensive queries;
- Combining large datasets. Many of the issues mentioned above have been (partially) solved for the non-spatial domain, but for spatial data many of the solutions need to be re-engineered.

1.4.4 Semantics of GI

This research theme should further develop the theory, methods and techniques that will allow the intelligent matching between user needs and resources made available through the facility. The multitude of GI resources, data and computational alike, combined with the wide range of users and uses, urges Maps4Science to look into the deeper semantics of the resources, as well as those of the respective user and use case. The application of semantic web technology has been tested in the GI domain already, but its exploitation to *filtering resources* on the one hand, and *matching resources* with use cases and perceived needs is still mostly unknown territory.

The innovation and challenges of this research theme are in the appropriate capture of the semantics of complex combinations of 2D, 3D and time dimensions, levels of uncertainty innate to the data, and the underlying physical processes the data represents. A correct capturing of the semantic relationships between datasets, for instance in lineage, and between data and computational resources is also an open issue.

An important (semantic) development is the so-called 'Linked Data' approach, which is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. The Linked data is a technology proposed by the W3C that allows for mechanisms to link data - as opposed to documents - on the web. The data storage format is based on triples (object-predicate-subject). Each part of a triple has a Uniform Resource Identifier (URI). The actual linking of the data is done by referencing from a URI to other URIs. The Semantic web is envisioned as the next generation of the Internet where content is provided not only using documents but rather with linked data, contained in concept models. The concept models are coded in resource description framework (RDF)/web ontology language (OWL). The instantiated (existing) concepts are stored as triples of linked data in the ontology concept models. From the W3C the Open Data Movement aims at making data freely available to everyone. There are already various interesting open data sets available on the Web. Examples include Wikipedia, Wikibooks, Geonames, MusicBrainz, WordNet, the DBLP bibliography and many more which are published under Creative Commons or Talis licenses.

For the GI-Science research community in the Netherlands the large-scale research facility is aiming at the same goal as the W3C open data movement, making data available for research and innovations. It means that the facility should be included in schema above. It also means exposing data using triples to interlink data based on common accepted ontologies or if appropriate own developed ontology to accommodate specific requirements of the GOF.

1.4.5 Services, searches and optimization

This research theme addresses specifically the non-data resources that the facility will also host: the computational resources. It will do this with the aim to grow a healthy and well-understood corpus of processes to operate on GI resources. Some corpora are already available for such use. What is missing is a proper framework to analyse these resources in terms of their I/O behaviour, their application to data resources and their intent. Eventually, one wants to allow the combination of processing services, known as service orchestration or service chaining. This can only be achieved with such an analytic framework in place.

The Maps4Science facility will need a number of standardized and basic services to support its objectives. Typically, these can be identified as the CRUD acronym: create, read, update and delete. The characteristics of GI force the facility to pay special attention to issues of data sanity for any of the data resources it takes in. *Curation services* for GI are rather special instruments, and care needs to be taken to organize them properly. Some of such services do not currently exist, and Maps4Science's ambitions of coverage are substantial. *Derivation* services are another class of processes that use more raw datasets as inputs and produce derived data sets as a result. These derivation services come in many flavours, and curation is in a way a very special case. Deeper semantic characterizations of services themselves may allow use to come forward with a theory of service chaining, specifically addressing the geo characteristic of such services. This is an open area of research. It should aim at tools that address the question which services are pair wise combinable. An extended theory should potentially also address the issues around service chain optimization: the redefinition of service chains that produce identical results, or perhaps similar-enough, or equally useful results, but with lower computational costs involved.

1.4.6 Standardized data models and data quality

Whereas RT 4 attempts to develop tools to allow matching user needs with facility resources, this research theme addresses the question of how to obtain standardized data models for facility resources to allow the facility to realize those resources. The commonality is in the data models of this theme, and the ontologies in RT 4. Harmonization is needed as the datasets are heterogeneous and from different sources and without harmonization before presenting the data to the user (researchers) it will be perceived as a chaotic mess. Harmonization can be done at the start (and then becomes the native representation) or at the end (before presenting to the users). Both approaches have their own advantages and disadvantages. In the Netherlands quite a number of data model standardization initiatives have been conducted: the basis is formed by the GI-standards of OGC and ISO TC211. The result is a base model geo-information (NEN3610) on-to which various (better harmonized) domain or sector models have been developed: TOP10, IMGEO, IMBOD, IMWA, IMKICH, etc. Though relatively stable, these models need maintenance due to chaining requirements or technology improvements (e.g. extend from a 2D to a 3D model). The second harmonization strategy (transform before presenting to user) may be quite attractive in a heterogeneous researcher/ user environment as the data may be transformed into the model (structure and terminology) which is most meaningful to the specific user.

An important aspect of working with geospatial datasets and in particular of working with reference datasets is the quality of the data and the means to assess the quality. High quality datasets are required when combining and aggregating data to, for instance, a larger scale. It is common to describe the quality of geographic data by several parameters, of which accuracy is most known. But other components, including precision (or resolution), consistency (are there contradictions in the data?) and completeness are equally important. Moreover, each of these components can be assessed in the spatial, temporal and thematic domain.

Although the above quality indicators can be measured in an objective manner for a given dataset their values must be compared to a quality standard, which is often dependent on the intended use of the data. For instance, land-monitoring applications (regional) might require lower resolution data than research applications on a local scale. Also timely availability depends on its application. Therefore, it is common to distinguish three types of quality assessment and reporting based on i) compliance and conformity control (comparing data against an a priori standard), ii) metadata standards (allowing fitness-for-use assessment based on the data documentation) and iii) user feedback (where users are given the possibility to report on the quality of the data).

A service offering spatial datasets should also provide the users with the means to assess the data quality. This could mean that users searching for data will have the possibility to filter the datasets based on one or more of the quality indicators, provided that these indicators are present in the metadata. On the other hand, users might be given the opportunity to rate the datasets based on their experience with using the data.

1.4.7 Volunteered GI and citizen science

This research theme addresses the exploitation of alternative data sources created by active, collaborating communities present on the web, often consisting of ordinary members of society.

The upsurge in online presence of people from all parts of society, fuelled by a ubiquity in mobile devices of which a growing percentage is equipped with sensors and positioning capability, is providing us with a growing body of volunteered information about many aspects of ordinary life. That information comes to us through social networks like facebook, foursquare, and in The Netherlands Hyves also, as well as through other web-based systems where citizens have organized themselves around a topic of their interest. Where such rapidly building up information resources centre around a theme, and the information resource is built up in some methodical way, we may speak of citizen science. Themes of interest come in many flavours, and in our western societies often involve leisure activities, tourism or peer-to-peer citizen trade. Societal reporting facilities, for instance on crime or pollution, are also on the rise, and have been instrumental in positive societal response in crisis situations like the Kenyan post-election riots and the Haiti earthquake aftermath, as the Ushahidi testimonial system has shown.

These resources created by the 'crowd' are considered to carry important value, even though they are normally considered to be non-authoritative. Their value lies in the sheer numbers, as demonstrated by the nature observation collecting site waarneming.nl, which in just over seven years time collected 14 million animal observations for the Dutch territory alone, turning our country into the most densely surveyed country of our times.

Where such an information resources carries locality information, whether explicitly encoded or implicit in the full text, we call it Volunteered GI, or VGI. One of the strongest cases made for VGI is in the collaborative Open

StreetMap (OSM) initiative, which has accumulated a road map of the world in just six years with 2.3 billion track points for 100,000 roads worldwide.

VGI has obvious potential but also comes with yet unresolved problems. Errors will always be included, and there is a question of reliability and trust around the accumulated data, even though the collaborating community typically includes data policing volunteers. VGI is here to stay, but we need to develop the methods to reap its benefits to the full, and to address its weaknesses at the same time. VGIs role can be manifold: it can be used as ground truthing and other verification purposes, it can also be used as raw data source for spatial phenomena for which no alternative survey methods currently exist.

1.4.8 GII assessment

Since the beginning of the 1990s, many local governments, countries, and regions have been building GIIs. The GII policies written for a specific jurisdiction (e.g., province, country or region) define specifically what the GII being designed and developed aims to achieve. For example, the specific goal of the INSPIRE directive is to create an GII that would assist policy makers in their activities affecting the environment (European Commission., 2007). Given this wide array of intended but rarely proven GII benefits and goals, it is natural for researchers, policy makers, government representatives and the public to be interested in the assessment studies measuring the benefits of GIIs and the level of realization of the goals.

The Dutch GII implementation provides an example of governmental demand for monitoring GII goals. The increasing demand for rational assessments requires generic assessment approaches that measure, for example, the extent to which GIIs programs meet their objectives. There is also a growing awareness across governments and communities of practitioners that much more attention needs to be paid to assessing the social, economic and scientific impacts of GIIs. Up until now, GII impact studies have mainly had an ex-ante character, i.e., focused on predicted GII impacts and benefits. Ex-post studies of GII benefits and impacts are still rare and only a few theoretical considerations and best practices in this matter can be mentioned. As GIIs can be viewed as Complex Adaptive Systems (CAS), the principles of assessing CAS can also be applied in assessing GIIs.

For further study, it is recommended to use existing expertise on current GII assessment views to measure the goals realization applied to the large-scale research facility being a GII. The design of the assessment view is generic, so it can be used to measure the extent of goals realization of any infrastructures with clearly defined and agreeable goals and where all of the stakeholders can be identified and approached. In addition to measuring the intended GII goals and benefits, the unintended impacts and benefits of GIIs cannot be neglected as they may be as equally or more interesting and important as the intended ones. Applying knowledge gained by assessing GIIs will improve use and usability if the realized large-scale research facility to benefit research and innovation.

1.4.9 Satellites as a service

Many researchers and operational services working in the fields of marine, weather and climate, land and astronomy applications recognize the importance of Earth observation satellite systems as a valuable source of geo-information. Application of Earth observation information requires scheduling of satellites and their sensors, acquisition and pre-processing of the sensor data and dissemination of the products. In addition, often these products need to be combined with the geo-information from multiple auxiliary data sources before they can be applied. The increased use and importance of satellite data has led to more time-driven user requirements. Instead of ordering off-the-shelf data products from a satellite operator and applying these products days to weeks after the actual acquisition, users have a need for real-time and on-demand programming, acquisition and delivery.

1.5 Science case 2: Data rights management and GII governance

Maps4Science will provide an infrastructure that enables researchers to access and use geo data from a large variety of data producers. Besides technical facilities, data policy issues play a key role in the actual access and use possibilities of data. A well-functioning and trustworthy infrastructure can only be achieved if a clear data policy is formulated and use restrictions (if any) respected. For enforcing certain use restriction technical as well as legal solutions can be used. It is the ambition of the Maps4Science programme to formulate, implement and evaluate a data policy framework that on the one hand respect ownership and privacy aspects of the data and at the other hand stimulate innovation and creative use.

With respect to the DRM this programme is innovative in two ways. Firstly it will address the issue of privacy. While GI accessible made accessible does not necessarily contain information that can be related to subjects and as such may cause privacy issues, the combination of GI may. Automated monitoring of rule infringements including privacy issues through will help to safeguard data owners from privacy infringements and consequent liability issues. Secondly the issue of automated monitoring of legal obligations in linked-data infrastructures is still rather underdeveloped. Neither the standards for expressing legal regulatory knowledge nor the mechanisms to reason with that knowledge effectively in an open-linked data environment have yet been developed and agreed upon. In this programme we want to push this aspect further. Although this may be beneficial to other sectors as well it is certainly relevant for a GI infrastructure.

1.5.1 Developments

The access and use policy can be arranged on individual data-set level or on a generic level for a large number of datasets within an organization or even a country. Besides a few exceptions, most geo-data access and use policies are historically data set and even application based. For every data set a special licences with use restrictions, costs and obligations were specified. It is now generally accepted that these strict regulations on data access and use hamper innovative applications, are time consuming and costly, and privilege certain 'power producers and users'. As a reaction on these strict regulations a more 'open data policy' evolves, with as goal to make the data available to all users at no cost and limiting conditions. A good example is the *Landsat* archive with remote sensing images of the world, which is now unconditionally at no cost available for all users worldwide. This policy has resulted in a strong increase in use and development of innovative applications. Some countries are also turning towards open data policies. A recent example is the UK, where in 2009 the government has decided to make their data 'open'. This policy decision has resulted in many innovative applications and sharp increase in data use. The Netherlands take an intermediate standpoint in the 'open data policy' arena. In generic terms the government is a supporter of open data, but at the same it maintains the specific regulations that apply for specific data sets. The provinces in the Netherlands are frontrunners in implementing an open (geo) data policy. For the implementation and success of the Maps4Science programme a supporting data policy that facilitate the access and use of geo-data is essential. The current overall data-policy of the Dutch government clearly supports the Maps4Science programme. What can be observed, however, is that the actual implementation of this 'open geo-data' policy is hampered by practical constraints and 'fear for change by data providers'. Within the Maps4Science programme experience will be gained to overcome these practical barriers, resulting in an overall data policy framework. This data policy framework contains legal, financial, organizational and technical components.

1.5.2 Towards a data policy framework

The development of a comprehensive data policy framework for geo-data is an important objective of this proposal. The full extent, structure and content are of course unknown at this moment, but amble attention will be given to the following items:

Practical implementation of generic data policy

The Dutch government has formulated a generic data policy of "open data". Open still means responsibilities of the data producer and conditions for the data user. The practical translation of a generic policy for individual data sets needs to be implemented and evaluated. This can result in different arrangements for different data-sets, depending on e.g. update frequency, privacy and ownership conditions.

Standardisation of geo-licences

Non-transparent and inconsistent licenses have often been identified as a major barrier to the sharing of geo data, and a clear need for harmonised geo-licences is increasingly being recognised. Only recently the development towards a standard for geo-information licences is starting to emerge. Examples of promising initiatives are found in the United States, Europe and Australia and at a global level. However, calls are growing for licensing models that have a broader reach than just on a national or sector level, possibly based on existing models such as creative commons (see e.g. GEOSS 2010).
Following these initiatives, but also in consultation with data owners, (generic) policy roles and rules will be defined that meet the needs of the owner as well as the user.

Securing data privacy

Quite some geo data are sensitive from a privacy point of view, with as a result that these data are hardly used for analysis. However, with special selection and aggregation facilities these data can be used, while safeguarding privacy. Within Maps4Science modules will be developed and tested that allow "secure" selection from privacy sensitive data.

Data funding and charging arrangements

The production and maintenance of geo data is costly. Available data does not automatically mean free data. At this moment a large variety of data funding and charging arrangements exist. Within Maps4Science an analysis of the different funding and charging models will be made and generic principles formulated.

Community feedback

"The proof of the pudding is in the eating". This does not only apply for pudding, but also for geo data. Currently, systematic feedback of users on the quality of geo data is hardly done. Within Maps4Science modules and procedures will be implemented for systematic user feedback on data.

Data protection issues

Maps4Science aims to develop tools that allow achieving compliance with legal and contract requirements. For this, Maps4Science will build on the findings of the EC-funded ACGT project in which a novel data protection framework has been developed which defines several legal and technical layers to demonstrate new (pragmatic) approaches to data protection while at the same time ensuring compliance with existing legislation.
It is advised to leave the responsibility of data policy implementation where it should, namely by the data owners. Within a distributed ICT environment (as Maps4Science intends to be) where datasets and data policies are spread out over the environment, the data policy and protection framework should be such that policies can be enforced at different parts of the system. It is recognised that the available standards for data protection and description (e.g. meta data standards) may not be mature enough to provide for all data protection issues. Within this work package these standards have to be assessed and updated if required.

1.6 Science case 3: Spatial Cyberinfrastructure

The science case for the spatial cyberinfrastructure (SCI) is linked to its development. Algorithms and scientific achievements to deal with processing and linking (semantics) are part of this science case. For a further elaboration of the science case on the spatial cyberinfrastucture we here refer to chapter 6: Technical Case.

1.7 Science case 4: GI Use

As can be seen abroad, a facility for GI provision serves many scientists in very different disciplines. A large range of scientific GI use cases clearly illustrates the need for geo-information and related geo-processing services. The book "Geospatial Technology and the Role of Location in Science"[4] also shows this in greater detail for selected cases. The list of scientific disciplines using geo-information include: architecture, landscape design, transportation science, movement studies, archaeology, history, demography, economics, spatial planning, spatial development, civil engineering, hydrology, water management, social and physical geography, geology, biology, ecology, health (epidemiology), crime studies, political science, disaster management, teaching, and many more.
In annex B we describe the following cases where location information bring scientific benefits to many types of science:

- Science case Health: Measuring and Forecasting the Spread of Epidemics
  Prof. dr. Peter Sloot, Dr. ir. Alfons Hoekstra, UvA and Drs. Carl Koppeschaar, Science in Action
- Science case Water resources: Better management through geo-information
  Prof.dr.ir. Nick van der Giesen, TUD and Dr. Rob van Swol, NLR
- Science case Crime: Geo-information and GI-Science as Crime reduction tools
  Prof. dr. Marianne Junger & Dr. Lorena Montoya-Morales, Dept. of Social Risks and Safety - UT
- Science case Agriculture: Avian Influenza – Don't spread the disease
  Ir. Henk Janssen, WUR

[4] Series: GeoJournal Library, Vol. 96, Scholten, vd Velde, van Manen, Eds., Springer 2009

- Science case Cultural history: The Integrating Heritage Program
  Prof.dr. Jan Kolen, Prof.dr. Henk Scholten, Dr. Niels van Manen and Maurice de Kleijn, MA, VU
- Science case GNSS performance: Support mission-critical applications by predicting GNSS performance
  Drs. Judith van Bruggen–van Putten, NLR

2 Talent Case

As has been set out in the previous chapter, this large-scale research facility will allow universities and other science groups in the Netherlands to maintain and improve their international scientific ranking. The improved efficiency in discovering and applying geo-data in all kinds of science domains will provide the Netherlands in a comfortable position to facilitate talent from all over the world. It also enables organisations to become leading partners in the European Framework programmes and in global initiatives. This also improves the attractiveness for foreign excellent young scientists to apply for scholarships and grants to work in the Netherlands.
In international science respect, we can be very proud of the international reputation of the universities of Amsterdam, Delft and Wageningen and the ITC – part of University Twente in the field of GI science. In addition, the Netherlands also hosts the ESTEC space research facility of ESA.
In addition to the international talent gain, Maps4Science is also improving the development of talents in education. The participation of EDUGIS allows us to directly interact with high schools and vocational education to create awareness and inspiration to new generations. Undoubtedly, this will improve the influx of motivated and talented students.
Maps4Science as a large-scale facility will be the second big impulse for the Dutch GI science community, after BSIK-RGI. The BSIK-RGI programme proved to be a very attractive research environment for foreign PhD students, Postdocs and international visitors.

3 Innovation Case

3.1 Introduction

In the past years the GI science has invested a lot to embed itself in a sustainable collaboration with government and industry. This has resulted in a collaborative strategic research agenda. Also, Dutch scientists participate in developments around the implementation of the European INSPIRE framework.
The NLR involves stakeholders in the development of their research agendas in a so-called Knowledge Arena. The latest edition of this arena was dedicated to "space for geo-services" where industry and government parties discussed the way monitoring and sensing instruments from space contribute to geo-information infrastructure.
The fact that geo-information is crucial in many disciplines (research subjects) is also apparent from an initial inventory of IIPGeo (ICT Innovation Platform Geo) in relation to the top sectors (http://www.top-sectoren.nl) as recently (spring 2011) defined by the Dutch government. In fact sheets for the top sectors AgroFood, Water, Energy, High Tech, Life Sciences and Creative Industry the role of geo-information is further explained, including the names of the geo-contact persons (see http://www.iipgeo.nl/IIP_GEO_kanaal/Topsectoren.html).
As can be derived from the recently published plans of the '*Topsectoren*', information technology in general, and in several cases geo-information in particular, is an enabling technology for expected innovations and breakthroughs. For most of these innovation sectors, the Geo Information Infrastructure will be a horizontal technology providing skills, tools and data for the benefit of all.
The Netherlands' GI sector has a large number of innovative spin-off companies originating from universities and research organisations. Typically these companies are quite small. Their dependence on a national science and innovation structure is therefore quite high. On the other hand, the uptake of new technologies and scientific results is quite high.
In several user domains large companies increasingly use GI in their business processes. Typically in domains like water (management), energy (oil companies, utilities etc.), logistics and agrofood GI has become a main ingredient for optimising businesses. These companies will be approached and involved in the development of this facility. Their role in particular in the continuation of the Maps4Science facility can be significant.
Institutional stakeholders in the Dutch geo-information infrastructure, like Kadaster, Rijkswaterstaat, Dienst Regelingen, Het Waterschapshuis, the provinces and municipalities are optimising their resources. With the initiative

of a Shared Service Organisation GEO, this process will only be enhanced. For these organisations a clear linkage with the Maps4Science facility is a means to connect to innovative solutions.

According to the branch organisation Geo-Business Netherlands, the total turnover of the Geo-Information sector in the Netherlands exceeds 1 billion Euros, and has an annual growth of 8% despite the current economic crisis. At the moment 15,000 geo-professionals are employed in the Netherlands, from which 3% are employed at research institutes, 65% at industry and 32% at government. It is therefore important that universities, government and industry stay in close relationship with each other, not only to increase innovation and application development, but also because universities supply the workforce needed in industry. To support this exchange, the Ministry of EL&I is considering hosting a National Satellite Database that allows for efficient dissemination of satellite data covering the Netherlands. The initial investment for infrastructure would be in the order of 1 million Euros, later followed by an investment in imagery. Being able to link to this database will connect the science community to government and industry and immediately reduce cost of purchase of data.

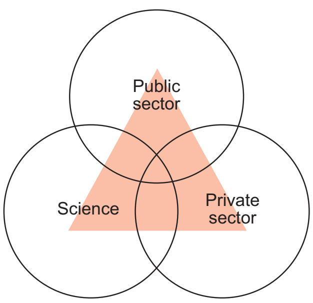
3.2 Living Lab

3.2.1 Introduction

In 2020 there will be more than 50 billion devices worldwide connected to the Internet. Together with the data already gathered by governments, scientists and companies this will create vast amounts of data, and 80% of it has a geographic component. At the same time structured data becomes available. EU implemented a European spatial data policy (INSPIRE) for the public sector and the Netherlands invested significantly in establishing a national geo-information infrastructure (NGII). This NGII primarily has been developed in order to improve public services to the general public and performance of public duties in the modern global Internet society (as the WRR calls it in their recent report '*iOverheid*'). An example of the use of this national infrastructure to facilitate innovation in public policy science is to develop an adaptation strategy for living in the Dutch delta, given the effects of climate change and honouring different stakeholders interests during the transition. As part of the e-government transition, an open data policy is tentatively being implemented to make government more transparent and accountable and, at the same time, to enable the private sector to innovate and stimulate the transition to a high-tech knowledge economy. This transition is essential in order to maintain national and European competitive edge by utilizing modern technology and compete in a global market.

In the private sector the Internet became an enabler for open source software and location-based date. Hence the Open Geographic Information System Consortium (OGC) improved interoperability of web-based geo information services. Access to global navigation data, Internet map data (Google Maps, Bing Maps etc) and the development of social media (Hyves, Twitter, Facebook) are game changers for doing business in a global market place. Development of mobile devices, location-based services, cloud computing and crowd sourcing are the next game changer for their business models. In Science the Internet technology enabled the formation of global science communities. Platforms like Wikepedia changes the way we find, distribute and even develop knowledge. Location Based Services enables the social scientist to observe human behaviour in space and time.

However all these organizational and technical developments have happened autonomously in science and the public and private sector. In order to maintain national and European competitive edge by utilizing modern location-based technology science, the public and private sector should closely work together at strategic and operational research. By co-creation of science, the public and private sector (the golden triangle) a strong growth in Location Based Services can be expected. The spatial grounding of public and private future challenges is at the heart of the Living Lab.



3.2.2 Innovation on Location Based Services in the Living Lab

In order to develop products and services that are 'game changers', the development must take place in an environment and context in which different systems and artefacts are used and operate similar to the real world (virtual) and in a later stage in the real world. A Living Lab is such a cross-over environment, involving all stakeholders and providing solutions that customer-organisations care about. A Living Lab is an experimentation environment in which integrated approaches are given shape in real life contexts and in which all stakeholders are considered 'co-producers'. All stakeholders (science, public and private sector) collaborate in creating solutions for our challenges. These challenges embody various realms of systems:

System Earth: System Earth consists of four spheres. These spheres are closely connected. For example, many birds (biosphere) fly through the air (atmosphere), while water (hydrosphere) often flows through the soil (lithosphere). In fact, the spheres are so closely connected that a change in one sphere often results in a change in one or more of the other spheres. Such changes that take place within an ecosystem are referred to as events. Events can occur naturally, such as an earthquake or a hurricane, or they can be caused by humans, such as an oil spill or air pollution. An event can cause changes to occur in one or more of the spheres, and/or an event can be the effect of changes in one or more of Earth's four spheres. This two-way cause and effect relationship between an event and a sphere are always location-based.

Global Society System: The Global Society System encompasses all-human society that exists on our planet from the beginnings of humanity in the form of many local societies and communities, through national states, to 24 hour economic system. In another words the Living Lab is an inclusive environment for different groups of people. The concept of global society becomes very fashionable in our times because with this concept most of today world social processes can be described and explained in very simple words. Behaviour or action in space and time of humans in conjunction with other humans, the System Earth and the Artefact System have to be understood for which the spatial frame is a natural but often underexplored base framework.

Artefact System (Man-made): Over time, humankind has developed tremendous artefacts located spatially. At the highest level the whole planet is the environment for all human technological activity. At a smaller scale, a large trading estate and the supply systems and support services that the companies use comprise an industrial environment. Transport infrastructures, be it public or private, cargo or passenger span the globe and have increasing influence on each other and on the earth and social system due to the ever-increasing competition for space. Cities grow and are not only produce by humankind but also influence society as established structures. The artefact system develops in space and time and choosing the spatial reference as a basis for interaction with the System Earth and the Global Society System is therefore a natural approach.

Natural, Engineering and Social sciences: In the Living Lab, scientists will be the hosts for public and private stakeholders, helping them to translate their real world problems into interactions with simulations, visualisation and computing-based on spatial environments.

Traditional natural science observed characterizes and mapped the static system Earth. Computer technology enabled natural science to model the system Earth processes in space and time. These dynamic models enabled science to study the cause and effect relationship. The models are also used to predict the impact of events and changes to the system over time.

In social science the scientists employ a range of methods in order to analyse a vast breadth of social phenomena; from survey data derived from millions of individuals, to the in-depth analysis of a single agents' social experiences; from monitoring what is happening on contemporary streets, to the investigation of ancient historical documents. The methods originally rooted in classical sociology and statistical mathematics has formed the basis for research in other disciplines, *such* as political science, media studies, and market research.

The engineering discipline produced the technical side of artefacts and is often the 'problem solver' for wishes in society, especially given the technology-savvy Western society. The use of models and simulations is imperative to the field, and often rely on the 'fundamental' natural sciences like math, physics and chemistry.

Location-based services enable natural science, engineering and social science, policy- and decision-makers to connect and model and simulate 'the real world' in order to predict dynamic changes in space and time of behaviour of humans in relation to the natural ecosystem. Research at the spatial-temporal interface of natural and social science is innovative as illustrated in Figure 2.

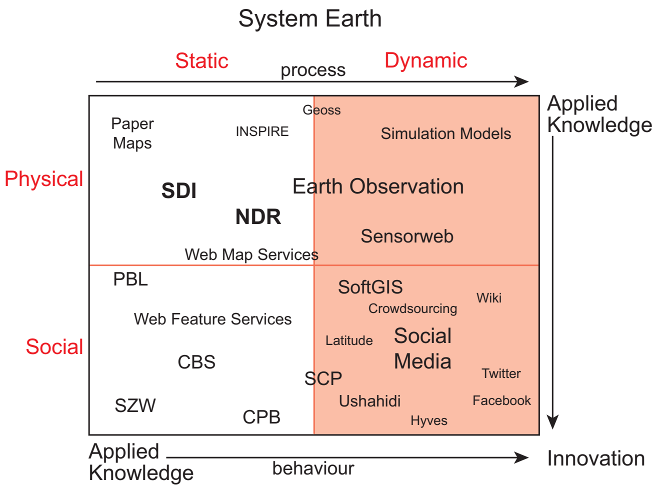


Figure 2: Innovation in location-based services.

In a living lab it is possible to cross the boundaries between the different disciplines of science, to use location-based and spatial analysis methods developed for natural sciences in a social science setting and tot location-based sensors as a basis for experiments in public policy making and business development. By combining location-based information from the various scientific disciplines with public policy and private sector innovations it is possible to achieve paradigm shifts in the way we do research and create public policy, implement infrastructures and do business in a global context.

3.2.3 Connecting sciences and systems for stakeholders

The Living Lab operates a proving ground for inventing, prototyping and embedding new collaborative arrangements of stakeholders. The Living Lab takes advantage of the available pools of creative talent, affluence of socio-cultural diversity and the unpredictability of inventiveness and imagination in the societal setting. Inhabitants and professional practitioners play significant roles in defining the scientific challenges and formulating the demands thereby shaping the research programme through processes of participatory design. To remain effective, a Living Lab must encourage and promote close interactions between science and practice, between public and private sectors. Serious gaming and simulation technology will be explored to test application of the processes of participatory design for location-based services.

By involving all stakeholders (science, public and private sectors) and customers from the start the living labs to increase the chance of successful product development and implementation and hence crossing the chasm (and avoiding the 'Valley of Death').

3.2.4 Contributions of the Living Lab

In the Living Lab the actual demand of public and private stakeholders will shape the actual activities. They will be organised along the following lines of contributions:

- *M4S testbed facility:* In this form a formal linkages with the OGC process will be implemented to assure that standards requirements identified by the Maps4Science program are addressed and to ensure that OGC standards are responsive as an interoperability "envelope" to facilitate the rapid transfer of research results into broad field application The M4S testbed creates a formal connection between the Maps4Science program and the OGC Interoperability Program. Significant future collaborative opportunity are envisioned whereby the M4S facility / program defines requirements in OGC international testbeds, pilots and experiments as a way to leverage OGC's international resources and knowledge base.
- *Live City lab based on real time sensor and crowdsourcing data and information:* Development of a platform that allows collecting, processing and distributing real-time data that originates in the city. Instead of focusing on any one specific application or limited number and type of data stream, this project explores up solutions that can cater for a large number of streams of very different kind of data, emphasizing the possibility of creatively combining multiple streams in the subsequent design of applications on top of the platform. As many users will be using this system simultaneously, key aspects will comprise building a scalable and distributed infrastructure to

smoothly distribute loads between platform and application providers without affecting the overall performance of the system.

- *Full Immersion Experiences using Layered Visualisation:* Based on successful approaches in the engineering sciences, stakeholders in a particular problem will be immersed in the spatial components of this problem. What are we talking about, what are the options, and what are the different viewpoints both physical and social? What would happen to the different systems (earth, social and artefact) in space (3D) and time (4D) when certain solutions are implemented? In a visual immersion theatre these options can be displayed in their full complexity and intertwined character. This leads to shared situational awareness, common grounding for decision making and provides opportunities for training.
- *Gaming Simulation:* The choices on the future state of the three systems often cannot be simply listed, but is the process of complex interaction between different stakeholders in a policy process. Gaming simulation can help to study, develop and train decision makers in the way they actually come to their decisions in full interaction with a virtual as-if-real environment. Traditionally, gaming approaches are powerful in immersing participants in the process of the policy making Maps4Science brings tremendous opportunities to bridge the traditional gap between (policy)(serious) gaming and computer simulation, by connecting this through a shared definition of the spatial domain. In the Living Lab gaming simulations will be developed that make the participants not only experience the process but also the detailed local and global consequences of their choices.
- *Multi-scale High Performance Computing:* The opportunities that Maps4Science provides for powerful computer simulation are nearly unlimited. The value is found in the combination of very detailed micro data and the opportunity to zoom out to the meso and macro level emergent outcomes. This connects to the rapidly growing field of multi-scale computer simulation that all of a sudden finds its natural application based on a spatial foundation. Multi-scale High Performance Computing will help stakeholders to formulate their questions and choices on their level of understanding and influence. This may even be at the level of the individual civilian. After running the models, the consequences of the local choices for the system-level outcome will be clear for policy makers. For the household level, the outcomes can be specified for consequences to their particular situation.

3.2.5 Living Lab in international context

This living lab will be part of the national, European and international community of living science labs and contribute by its unique value based on combining location-based information from the various scientific disciplines with public policy and private sector innovations. Also a number of international initiatives have been identified. In order to maintain a scientific competitive edge in this field this Maps 4Science program is evident.

- New York – city sourcing via the big apps programme – creates a more liveable city despite shrinking budgets.
- San Francisco – in addition to an open data program, this also includes initiatives such as SF Park - variable pricing based on use of parking spaces. Sensors measure how a parking is used: commonly used parking lots become more expensive, rarely used parking lots become cheaper. People develop a spatial and real-time location-based understanding.
- LIVE Singapore – a collaboration between MIT and the Singapore government implementing the principle "Making Decisions in sync with the environment" and applying open data, crowd-sourcing, sensors and dynamic pricing.

Note that in the Netherlands various cities (Amsterdam/Rotterdam/Eindhoven/etc.) each have similar open data programmes, with special attention for Assen sensor city (TNO experiment).

4 Partnership case

Maps4Science is a facility of and for the Dutch GI science community to improve the use of geo-information in research, and to develop innovative modes of spatial data use. It requires the cooperation of many parties inside and outside the national and international academic setting. The network includes parties who are not beneficiaries of this facility and have no direct role in creating the facility. These parties have provided a letter of support in which they explain the importance of the facility for their own activities.

At the national front, the consortium represents all the important academic players of the GI domain, as well as their associated library entities, and it is supported by the vast majority of the important national spatial data producers. We expect this network to synergize on data acquisition and production chains, and to be the important vehicle to develop new use cases, especially across the board of science. The partners have a positive history also, a.o. in the RGI programme. Expectations on the use cases of Maps4Science are high because of the unparalleled

Dutch data richness earlier mentioned. The strong support that our proposal has drawn from the Dutch private sector is ample proof of this. Positive intentions have mutually been expressed on collaboration with EduGIS, the geoportal for Dutch education entities. Executive bodies of several ministries are intending to collaborate in a shared service organization GEO. It is expected that they will host and maintain essential spatial datasets for public use; these will obviously be of interest to Maps4Science, and we are well-connected to this initiative. Currently, the Dutch Space Office NSO, by request of the government, is investigating the development of a National Satellite Database to make available from 2014 large amounts of space data from the European GMES programme. It perfectly relates to the intentions of Maps4Science and the two could strongly benefit from each other.
The consortium's embedding in the European GI landscape is strong: collaborations at project level exist with many academic partners inside and outside of GI science, and we are actively involved in AGILE, the association of European GI labs. Geonovum is coordinating the Netherlands public sector efforts in INSPIRE. We have strong ties with EuroGeographics, the combined European national mapping, land registry and cadastral agencies. The consortium will seek collaboration with EDINA, the UK national spatial data centre with a similar objective as that of Maps4Science. EDINA is in a way our competition, obviously not on content, but on functionality. Maps4Science seeks more innovative dissemination modes than what EDINA currently offers, and will be offering much more, and better tagged resources. ESA's GMES efforts will lead to substantial data products requiring Dutch ground truthing.
OUTSIDE EUROPE: The direct support offered by peer organizations in Canada (Geoide) and Australia (CRC-SI), both considered to be world leaders in the GI field, is proof of quality and timeliness of our proposal. Our group's members hold or have held steering positions in a large number of GI-organisations, or are otherwise active in these bodies: the globally operating GSDI and ISPRS, standardization committees like OGC, ISO/TC211 and CEN/TC287, as well as professional associations like FIG and ICA. Strong and active partnerships also exist with some of the leads of GI industry, esp. with ESRI, Oracle and Bentley. We maintain project connections with some national space agencies, specifically NASA and Brazil's INPE, and with some global sponsors for GI development projects abroad like Worldbank, Gates Foundation and USAID. In that context, members of the consortium have worked on national SDI implementation projects around the world. Finally, we are also globally active in GI education networks, with joint programs in a number of non-European countries, and hosting on behalf of the UN two UN University schools: Land Administration and Disaster Management.

5 Business case

5.1 Total Costs

	SET UP PHASE		OPERATIONAL PHASE					
	YEAR 1	**YEAR 2**	**YEAR 1**	**YEAR 2**	**YEAR 3**	**YEAR 4**	**YEAR 5**	**TOTAL**
Construction and investment costs								
Spatial Cyberinfrastructure	600	2200						2800
Customisation to educational needs	300	923						1223
R&D geo governance	400	615						1015
R&D GII	400	615						1015
Running costs								
Maintenance		150	542	555	569	583	598	2997
Evaluation and improvements			105	108	754	110	113	1190
Personnel costs	200	410	630	646	662	679	696	3923
Housing+office expenses	100	51	53	54	55	57	58	427
Demonstration science cases			420	431	442	453	113	1858
Living Lab construction	100	205	210	215	221	226	232	1409
science case GII			841	862	883	905	696	4186
Decommissioning								
transition						230	340	570
Other costs								
Communication	30	30	30	30	30	30	30	210
Total/yr	2130	5199	2830	2900	3616	3273	2876	22824

The costs consist of the following components:

- *Spatial Cyberinfrastructure*: The development of software and investment in ICT including hardware and network services to operate the facility;
- *Customisation to educational needs*: The extension of the facility to service high schools and vocational education with the tools and data developed. This is to support the talent case;
- *R&D geo governance*: In the setup phase research must be done in the way we can deal with spatial datasets in the heterogeneous, distributed environment;
- *R&D GII*: Also, research must by done on how the GII can be operated, in particular with regards to the high volume data of sensor networks and other;
- *Maintenance*: maintaining the facility and keep it running. Estimated at approx. 15% of the investment budget as annual expenses;
- *Evaluation and improvements*: Looking into the technical performance of the system and budget to adjust the system where needed. In year 3 a major revision of the system is foreseen;
- *Personnel costs*: Staff costs for the facility staff including the director. Envisaged is a total staff of approx. 6 – 8 FTE;
- *Housing and office expenses*: Operational expenses, including housing, printing paper etc., insurances etc.
- *Demonstration science cases*: Stimulating and showcasing some applications to draw attention of other scientists;
- *Living Lab construction*: facilitating funds to play its role as knowledge partner in the trinity with government and private sector;
- *Science case GII*: further elaborating the science related to GII;
- *Transition*: costs related to the transition of the facility from a temporary intervention to a more permanent and sustainable situation;
- *Communication*: budget for all kinds of communication, including website, printed material and workshop facilitation.

The operational costs in years 6-10 of the operational phase will remain on the level of year 5, including the costs for maintenance of the facility, evaluation and improvements, personnel costs, housing costs and office expences and communication budget. It is envisaged that the board will continue with the facility after the NWO funding stops. The costs will be distributed among the participating organisations.

5.2 Non-NWO contributions

	SET UP PHASE		*OPERATIONAL PHASE*					
	YEAR 1	***YEAR 2***	***YEAR 1***	***YEAR 2***	***YEAR 3***	***YEAR 4***	***YEAR 5***	***TOTAL***
TUD	*100*	*200*	*200*	*200*	*200*	*200*	*200*	*1300*
ALTERRA	*50*	*75*	*125*	*125*	*125*	*125*	*125*	*750*
DANS	*30*	*30*	*30*	*30*	*30*	*30*	*30*	*210*
GEONOVUM	*30*	*30*	*30*	*30*	*30*	*30*	*30*	*210*
NLR	*25*	*75*	*75*	*75*	*50*	*50*	*50*	*400*
UvA	*35*	*35*	*35*	*35*	*35*	*35*	*35*	*245*
UT/ITC	*100*	*150*	*150*	*150*	*150*	*150*	*150*	*1000*
UU	*35*	*35*	*35*	*35*	*35*	*35*	*35*	*245*
VU/EduGIS	*100*	*150*	*150*	*150*	*150*	*150*	*150*	*1000*
WU	*50*	*75*	*125*	*125*	*125*	*125*	*125*	*750*
TOTALS	***555***	***855***	***955***	***955***	***930***	***930***	***930***	***6110***

5.3 Requested NWO Financing

The budget calculated for realising the facility adds up to 22.8 M€. The sum of all maximum contributions of all partners adds up to 6.1 M€. The consortium agreed to provide 25% of the facility as contribution. Therefore the requested NWO financing is 17.1 M€.
The exact distribution of NWO financing over the different budget aspects needs to be done in a later stage, as the contributions vary, so the final distribution is depending on the final agreements of work distributions over partners.

6 Technical case

6.1 Spatial Cyberinfrastructure

The Spatial Cyberinfrastructure (SCI) component of Maps4Science realizes the facility to our user community as a federation of networked systems with the purpose to archive and serve out spatial data and computing resources. As such, it is scheduled to become part of the Dutch *National e-Infrastructure for Research*, as part of call for the Dutch research community when it comes to spatial data.
A cyberinfrastructure aims to support a user community with well-understood needs and fosters collaborations within that community. It integrates hardware and software solutions, as well as network facilities to sustainably archive data resources, as well as computational resources to operate on the data, offering the latest IT almost transparently to the end-users, and allowing them to focus on their domain specifics rather than on operating the machinery.
Modern cyberinfrastructures come in flavours: some are devoted to high-performance computing when computational capacity is scarce, while others have as prime purpose the archival and provisioning of data resources. Hybrid set-ups also exist. In addition, some serve out just a static, though possibly large, collection of resources, while recently more attention has been paid to the linked open data philosophy, and the idea of dynamic, unique namespaces for all resources. It almost goes without saying that such a modern system operates on the web, and uses web services as technology, typically in client-server mode.

For SCI, the user community consist of those scientists who have a need for geo-information from the Netherlands and who have difficulty identifying and gaining access to those data resources best suited for their own needs. SCI will at first be strongly data- and archive oriented, but is expected to develop more computational capacity over time, especially when the thematic and data coverage improves, and more complex data sources, involving time or scale variability, become available. We expect SCI to develop into a facility that can be accessed using transparently named resources that fully hide the distributed nature of the facility.
Below, we discuss SCI's embedding in the Dutch ICT infrastructural landscape for research, SCI's basic functionality as a cyberinfrastructure, and its extended functionality specific to being a *spatial* cyberinfrastructure. We also pay attention to the philosophy of resource-orientation, and the role that data models and open access play in it.

6.1.1 The Dutch ICT infrastructural landscape for research

In 2008, SURF proposed to NWO the creation of a National ICT Research Infrastructure, and the implementation of these plans started in 2009 under the governance of SURFnet. Two years later, they collaboratively launched the Netherlands e-Science Center. The latter aims to support multi-disciplinary data-driven research through the provision of various ICT resources. In addition, SURF is implementing a national high-performance computing and e-science support center (working name: SURF-SARA) based on the existing activities of SARA.
Maps4Science explicitly wants to connect with and fit in these developments, and is not aiming to create its own niche or independent existence. Our facility also targets data resources to be used in science, and this means that economies of scale can be reached when it comes to operating the basic technological substrate for networking, storage and the archiving functionality surrounding those two.
At this abstract level, the Maps4Science facility is not much different from other cyberinfrastructures that attempt to accommodate a scientific user community. Differences start to become apparent immediately when we look into the data content specifics: the *spatial* resources come with rather a long list of special characteristics. We come back to this further on.

6.1.2 SCI base functionality

A cyberinfrastructure facility needs to boast sufficient *capacity* as well as *capability* to serve its (future) user community. The first is a volume requirement that translates into available storage space, CPU cycles and network bandwidth, while the latter is a functional requirement translating into functions to support management of the resources, interaction with the users, and so on.
Typical activities under the heading of providing the base functionality are:

1. Acquisition and maintenance of the technology;
2. Training of and support for users;
3. Functionality development;
4. Monitoring and optimization of operation;
5. Interfacing with national and international environment, especially those efforts that provide similar cyberinfrastructures.

The key factor to success is in the understanding of the level of traffic and what it means for overall capacity, load balancing and scalability in case of increasing traffic parameters. The unit of data at this level is typically the file that holds a dataset, and the metadata record for it in the facility's catalog. The structural complexity of the catalog is highly variable, but is a not to be ignored factor in the level of intelligence in which the user can be supported automatically.
Functionality typically addresses authorization, write and read facilities for the data repository, explorations with the catalog, possibly the application of computational services with the data, as well as output reporting, specifically forms of scientific visualization.
An important special group of functions exist under the heading of curation. Curation can be seen as the processes put in place to ingest new datasets into the facility, and prepare and maintain them for sustained use. This involves a.o. formatting the data to be in line with the facility's data model. The Curation Life Cycle Model (CLCM) is a paradigm to structure and realize these processes. It recognizes full life cycle processes, sequential processes and occasional processes. These explicitly include detection and repair of data error, creation of metadata records, and countermeasures against datasets going stale for reasons of format or software upgrades. All in all, it prepares the dataset for optimal application and findability for prolonged periods of time. A second important group of functions

are those that operate the catalog. Essentially, they form the publish-find-bind cycle, allowing to make know to the catalog, to search the catalog and to serve up the data results of picked searches.
The approach that Maps4Science will take will be one of "following the leader", meaning that collaboration will be actively sought with e-Science specialists in the Netherlands, such as DANS and SURF. Modern ideas of such specialists indicate a move away from technology vendor specifics as well as data product specifics to *service-orientation*. In the latter, one will eventually not be able to distinguish a served up dataset from a served up computational result. This leads to notions of compositionality or service chaining, as well as to further abstractions in which, for instance, complete computational models are just a service also. Similarly, one aims at *interoperability* that allows different components to exchange data without problem, typically achieved by adhering rather strictly to a list of selected standards.
The base functionality of SCI sketched in this way can be realized with well-understood mature technology, which continues to be improved both in its capacity and in capability aspects. The costs involved are those that come with size of the facility, and quality staffing to make it a 24x7 undertaking.
Concerning data storage, data transport and data processing MAP4SCIENCE will cooperate with existing initiatives to implement an ICT Research Infrastructure.

6.1.3 SCI spatial functionality

A *spatial* cyberinfrastructure such as Maps4Science will be needs to address further issues that are rather specific for the geographic nature of the data content. Insiders state that "spatial is special", and there is some truth to this. First and foremost, spatial data has one common denominator in the ever-present location attribute: the data pins information to a location, and since this is done for all spatial datasets, for any location results a rather colourful barbecue pin with information across the board. If done well, this provides a wealth of data for all sorts of studies. When we add to the mix the location-specific history of such information, patterns of spatial processes emerge, the study of which will deepen our understanding once again.
Spatial data comes to us in many different forms, both conceptually and in the representation. Some geographic phenomena behave like continuous spatial fields (soil temperature, and other natural phenomena), while others are better conceptualized as discrete objects (administrative boundaries, and other often man-made phenomena). Some phenomena cover the study space, while others (for instance, in sensor networks) do not. Spatial data is often intrinsically uncertain, which calls for special techniques of interpretation and interpolation, in space as well as in time.
The GI research field has addressed the management of spatial data for a long time, and with governments building up their spatial data repositories, it has developed trustworthy methods of constructing Spatial Data Infrastructures (SDIs). Our SCI can be viewed from that perspective as an SDI for a rather specific user community. Much of the knowledge on SDI construction has been channelled through the globally operating Open Geospatial Consortium (OGC), and also ISO. Both have had a tremendous effect on the profession by putting in place abstract definitional and concrete realization standards on spatial data operations. This provides us with the definitions of the tools to realize SCI as a *spatial* facility. Various vendors and open source initiatives have come up with technological solutions that are abiding to those standards, and these solutions have matured in recent years, addressing at the same time the issues of interoperation.
Typical SDI facilities adopt service-orientation as a philosophy, and go beyond those offered in general cyberinfrastructures because of the commonality of their data holdings: location. This has caused a growth in spatial service types: one may request a map image for publication, or a map feature as a geometric element to continue computations with. Likewise, the catalog services have been spatialized, allowing for rather specific spatially-aware searches.
The benefits of the barbecue pin philosophy stretch further and have recently led to work in the geosemantic web. This is essentially an approach to apply semantic web techniques to capture more deeply the semantics of geographic phenomena and the characteristics of the spatial data that represents those phenomena in our systems. This is one interface point with Maps4Science's research program described in Section 1.5. Another connection must be made with the open linked data inclusion that we describe later in this section.
The barbecue pin also gives us a golden opportunity for exploration of datasets, because it allows us to create maps that combine previously uncombined data, or results of spatial computations. The map is used as the medium for exploration. From a SCI perspective this means client applications must be made available to allow for that

exploratory mode and for the (semi-) automated construction of professional maps ready for human consumption. Again, OGC especially has forwarded the standards, and commercial and open source solutions are available to create such client applications.

New spatial data formats are on the horizon, however, and MAP4SCIENCE has the ambition to accommodate these equally well. Sensor networks are appearing everywhere, for instance on the human body, and many bring us useful, spatially explicit data. Such data may come in the shape of live streams. Social networks are also a source of joy, and to various extents carry information that has some form of spatial anchor. Reporting sites, of crime, pollution, maintenance needs, hazard or accident events, are examples of these and may bring exploitable data sources. Another case must be made for the inclusion of open linked data, from official sources like government and other public bodies, but also again "from the crowd".

At the spatial functionality level, Maps4Science will adopt a special style of service-orientation known as resource-orientation. In brief, that style enforces strict rules of using communication protocols such as http, strict naming conventions for the resources offered, while allowing for fine-grain approaches to datasets, for instance by making selections amongst elements inside. This is justifiable because of the nature of spatial data and the locality principle often applied in operating on it.

The capture of spatial semantics and the resource-oriented style are two sides of the same coin: they collaborate to cast semantics of our datasets and computational services in a machine-exploitable representation that will allow us over time to develop highly generic forms of exploration and dissemination. Those semantics subsequently help us to make sense of the dataset, to improve its exploitation and to decrease the chances for erroneous or nonsensical use. Semantic technology involves ontological languages and reasoners to help us achieve such goals. An important development is that towards Open Linked Data, as proposed by Tim Berners-Lee in 2006, which can be viewed as a marriage of proper name spacing through URIs and ontological technology through RDF, to expose with a single metamodel the raw data behind the web's information sources, and to provide with materialized semantics. This makes them available for programmatic access rather than human access only. In all, this gives Maps4Science a rather challenging research and development agenda, aiming to capture and exploit deep semantics of spatial datasets in of computations with them. The objective of that new capability will be to support end-users with purposeful applications that fit their needs.

The core spatial functionality of SCI sketched above can be realized with now quite mature technology from the GIS and SDI communities. We can largely follow current SDI methods and standards. For the novel data formats, technology needs to be developed or improved, and this is in part addressed by the GI-Science proposals.

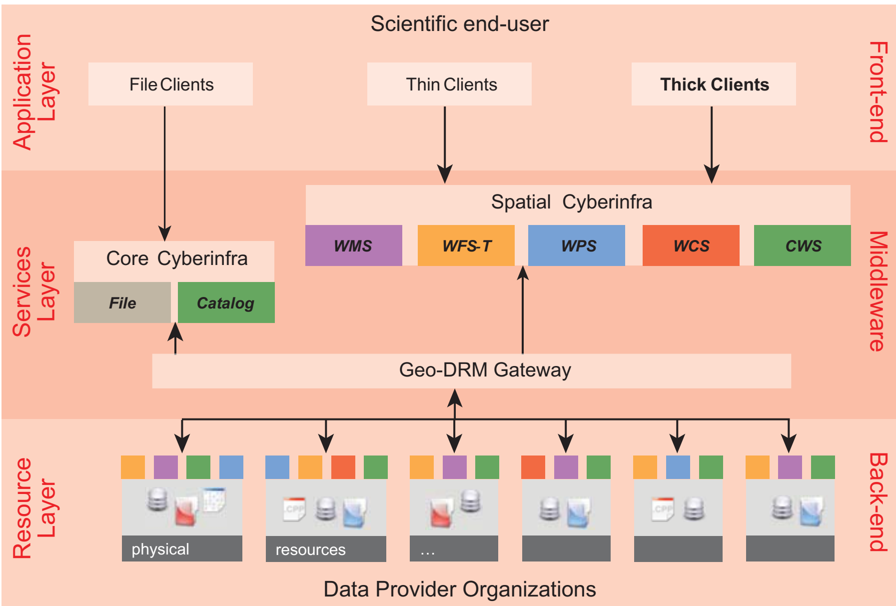


Figure 3: Initial functional architecture of SCI. File-based services (ftp,http) in the core cyberinfrastructure, and feature/map-based services (wms, wfs, wcs) in the spatial infrastructure. Catalog services for files and features/maps (CSW) are also included.

6.1.4 Maps4Science facility evolution

The construction of the necessary data infrastructure will develop in phases (see Figure 4):

1. In phase 1, the project will use existing infrastructure typically at partner nodes. The source data files will be stored on university servers. Students and staff of the Dutch universities will be able to access and download the data via Maps4Science. *SURFfederatie* handles the authentication. During this phase we will explore more fine-grain access facilities also;
2. In phase 2 of the project, the source data files will be stored on the servers of the Maps4Science consortium. Phase 1 and phase 2 will probably be used parallel to each other during most of the project. *SURFfederatie* handles the authentication. Much more full fine-grain access facilities will be put in place;
3. In phase 3 of the project, Maps4Science will access the data directly on the servers of the source data suppliers. Other consortia can in the long-term also use the infrastructure on these servers;

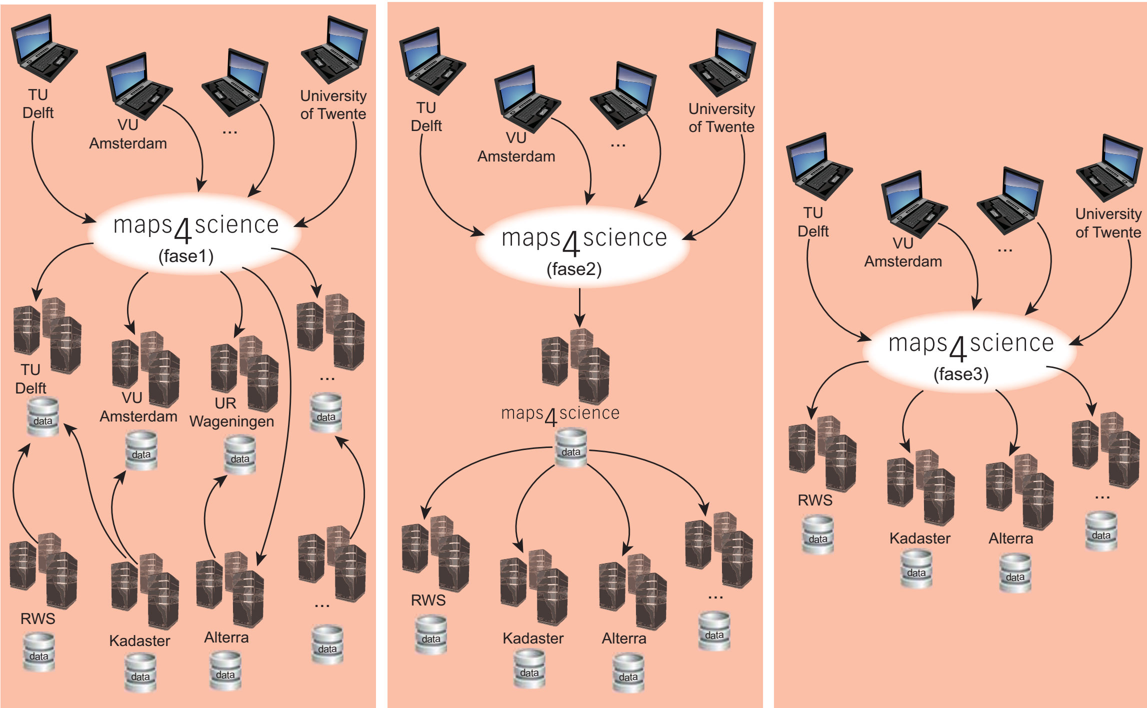


Figure 4: Development phases of the Maps4Science infrastructure.

6.2 RISK ANALYSIS

RISK	*IMPACT (1-3)*	*CHANCE (1-3)*	*RISK (Imp*Chc)*	*Measure*
Data charter and licence agreements not completed / insufficient	1	2	2	Early identification and commitment of crucial data providers.
Low end user acceptance due to technical barriers	2	3	6	Timely user involvement; showcasing and demos; courseware development. Also, monitor developments of large IT companies (MS, Google etc.) and other initatives.
Spatial DRM not implemented	1	2	2	Hardcoded rights management; Stricter (more rigid) licensing enforcement than needed.
Scope creep and loss of focus due to unrealisatic projections	2	2	4	Include providers, suppliers and users in advisory panels.
Drowning by data – greediness for large amounts of data with low added value	1	2	2	Include users in panels. Make adequate data plan.
Unrealistic user expectations	1	3	3	Early involvement of user

RISK	*IMPACT (1-3)*	*CHANCE (1-3)*	*RISK (Imp*Chc)*	*Measure*
				communities. Clear communication and awareness building.
Technical realisation results in low performance and low usability	3	2	6	Prototyping and testing in early stages. Involving experts. Using proven technology as much as possible.

Impact: 1 = light, 3 = threatening the facility
Chance: 1 = unlikely, 3 = very likely

Realising the facility is a major risk in terms of technical feasibility. However, managing the facility's environment in terms of users, providers, suppliers and other stakeholders carries equally large risks. It is clear that the facility should be demand driven as much as possible, so early involvement of (senior) users and (senior) suppliers is very important. A strong governance with clear goals is required to keep the developments on track.

An active risk management approach is a key component of our project delivery program. Besides the analysis made now, we include a monitor schedule for risks, results, and budget performance during the course of this project. This will ensure that risks and issues are addressed before they become major problems.

Risk management is a key element and is actively used in the process of planning and managing this project and lateron the facility. At the inception of the project, the risk list is developed and made available to the consortium. The risk list is used to capture and monitor the perceived risks to the success of the project. It identifies, in decreasing order of priority, events, which could negatively impact the project's success. Along with each risk is a plan for mitigating that risk. The mitigation plans allow the project to be structured to help reduce those risks or ensure that they are realized before significant investment is made. Throughout the facility lifetime, the risk list will be updated as new risks are identified. The list of risks is frequently monitored to ensure that each is managed effectively. Issues may be events that have occurred in spite of risk mitigation efforts, assumptions that have been proven incorrect, technical difficulties that have arisen, or needs for guidance that the consortium cannot resolve on its own. Most importantly, issues impede progress and need to be resolved as quickly as possible. The consortium will maintain an active list of issues that will include the nature of the issue, the individual responsible for its resolution, expected date of resolution. This tool will be used to manage the project schedule, ensure expeditious resolution of issues, and provide transparency to the development team and the (executive) board.

7 Possible focus for the Netherlands

Top 5 position in the world

The Netherlands excels as one of the top players in geo-information science worldwide. Bibliometric analysis[5] shows that the Netherlands is ranked fifth in the world in scientific output, as measured by the national contribution to the most relevant international journals. The Netherlands is responsible for a 6% share in the worldwide production of scientific articles in this field. It can be safely concluded that geo-information (GI) science in the Netherlands has an acknowledged position worldwide.

High data density

The Netherlands is historically known as one of worlds' best-measured countries. It is continuing this tradition today with unequalled new datasets, such as the nationwide large-scale topographic map, our unique digital height map (nationwide coverage; ten very accurate 3D points for every Dutch m^2) and a range of public and private collections of environmental and socio-economic geo-datasets. Despite the wealth of existing and well-maintained geo-information (GI), the access to GI for our academic community is rather poor. In particular, science domains discovering the enabling power of GI have difficulties in accessing and using it. Initial analysis shows that i) legal and commercial licence restrictions, ii) technical inability to deal with more advanced types of spatial data and iii) lack of awareness are the top three barriers preventing large uptake in science. This is a missed opportunity to achieve scientific breakthroughs in a wide range of scientific disciplines.

[5] Wageningen University, 2009. Bibliometric analysis of the geo-information science field.

8 Critical mass

The consortium includes 6 universities and 4 research institutes. In addition, there is support from a large range of interested and involved organisations. The long list of support letters underlines their involvement. The GI sector benefitted a lot from the 'Space for Geo-Information' BSIK programme and as a result, governmental organisations, private organisation, science and education have a solid base for cooperation and collaboration.
The leading scientists all have an excellent reputation in their respective expertises, as supported by visitation reports. This is a good starting point for improving chances for excellent research. External access is facilitated through SURF federation.
From the example of EDINA in the UK we learn that collaborative action to provide geo-information to the complete research community does indeed satisfy a huge need as can be seen from the usage statistics (see Figure 9).

9 Embedding

The organizational structure will have to take a number of central objectives of the program into account simultaneously:

- to bring together expertise on a range of sub-disciplines in order to address the core issues involved in setting up a research infrastructure for specific strategic research questions;
- to assure the coherence of the overall activities program in terms of the conceptual issues at the core of the program;
- to develop and apply new data sources;
- to assure high quality both of the overall program and of the individual projects;
- to keep to the schedule proposed in the planning;
- to stimulate communication among the researchers working in the programme

The Maps4Science governance is organised in a simple manner. A board, consisting of consortium partners, governs the facility. The TUD as lead partner will deliver the president to the board. Two Advisory Panels, a national and an international one will assist the board. The panels consist of stakeholders to the facility, including users, suppliers and stakeholders from public and private organisations.

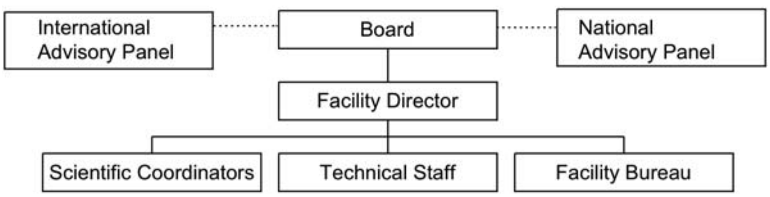


Figure 5: Organisation scheme of Maps4Science facility

The Maps4Science organizational structure consists of the following bodies:
Board, Consists of senior researchers and other experts with expertise in the field and in governance

- Max. 10 members; meeting twice a year;
- takes strategic decisions;
- is responsible to Funding Agencies;
- determines the Annual and Financial report; determines the Long Term Work Plan and the Yearly Work plan;
- decides on appointing members in the governance bodies;
- monitors progress of the project as a whole;
- acts as mediator in case of conflicts;
- assigns, if needed, new tasks to governance bodies;
- determines the procedure for avoiding conflict of interest;
- decides on admission of new participants;

Executive Board (EB)

- 4 members;
- Frequent meetings (once every month) and other contacts;
- The executive board reports to the Board;

- The EB is chaired by the programme director (Prof.dr.ir. P.J.M. van Oosterom), who has the full and final responsibility for the Maps4Science project reaching its goals and the effective functioning of its infrastructure. He will be in charge of the daily coordination of activities, implementation of evaluation procedures, finances, and personnel;
- The other members have specific subtasks, e.g. technical coordination, dissemination, education, awareness, IPR, etc., as their assignment;
- The EB prepares the long term work plan including success criteria for evaluating the project;
- The EB translates the long term programme into yearly work plans;
- The EB monitors the progress of assigned sub projects;
- The EB organizes evaluations of the project;
- The EB develops a procedure for avoiding conflicts of interest;

National Advisory Panel (NAP)

- 10 representatives from GI-Sciences;
- The representation of intended users of the infrastructure should be dominant;
- Typically meets four per year;
- Advises on a yearly work plan specifying steps to be taken with respect to the technical (prototype) infrastructure, the data infrastructure, the language technology service infrastructure, establishing user needs, creation and operation of centres of expertise; creation and operation of dissemination, education and awareness facilities;
- Evaluates and ranks subproject proposals from calls for project proposals especially taken into account the research environment in the Netherlands;
- Monitors progress of the yearly work plan.

International Advisory Panel (IAP)

- Appr. 8 members;
- Typically meets once a year;
- The IAP is a body of prominent and experienced researchers outside the Netherlands who are actively involved in the creation or maintenance of GI infrastructures, experts in technical infrastructures and experts in IPR matters;
- Advises the executive board and the board in international matters, e.g. cooperation, coordination and harmonisations with other international initiatives and programmes;
- Evaluates and ranks subproject proposals on request of the board resulting in an advice to the board.

The Maps4Science preparatory committee determines the initial composition of these bodies. All members of these bodies are personal members, i.e. they do not act as representatives of their organizations, cannot be replaced by colleagues, and will stay member if they would leave their organization. The executive board will develop a procedure to ensure that conflicts of interest are avoided. Maps4Science has a number of participating (research) organizations as its institutional members. Only participating organizations are eligible for funding of activities or subprojects. Maps4Science is open for new institutional participants. The Board decides on admission of new participants.

10 Proven willingness to collaborate

The consortium partners all had prior involvements in the BSIK programme Space for Geo Information (RGI). In this setting, cooperation of all kinds developed, like joint master course, joint research projects and partnering in national and international projects and consortia. As can be derived from the letters of intent, partners commit themselves to participate in the development of this facility.

A leading role is envisaged for the libraries and the so-called geo-desks of the different universities. It seems very logical to involve those who have similar local facilities running in the operational aspects of Maps4Science. In addition, they play a role at decommissioning and carrying on the facility after the end of this project.

11 Reflection of social trends

There is a transition of the geo-information role from the old paradigm card (describing a spatial situation) towards continuous identification, monitoring and control of processes at multiple spatial scales. This makes clear that GI is no longer a local or national concern. The GI facility and the use has now clearly an international context as regards the management and control of environmental systems at regional, continental and even global scales for both social and physical processes; this up to and including the System Earth.

Outside the research community also much geo-information is used. Gradually there is a transition from ad hoc geo-information provision towards a more infrastructure-based solution. So more information can be better found and used, and thus not need to be collected multiple times and is of guaranteed quality (as it originates directly from the source). This is based upon national and international regulations (Netherlands: base registers, Europe INSPIRE directive). In the Netherlands, the geo-registers and INSPIRE initiatives are implemented by the 'Publieke Dienstverlening op de Kaart' (PDOK), which is planned to be realized via a shared service organization geo (SSO-GEO). SSO-GEO is a collaboration between Kadadster, RWS and Min. EL&I. It further involves the following parties: LSV-GBKN, TNO Geological Survey, Geonovum, DR, Deltares, Logius, RVR, DLG, and Min. I&M. It is the intention that SSO-GEO will start in 2013 with the PDOK partners and will then increase until 2020. The SSO-GEO has two roles: 1. Executing operational tasks (provision for geo-information infrastructure) and 2. Stimulating innovation for the benefit of society (promoting new services). Maps4Science can be seen as the scientific counterpart (partner) of the SSO-GEO. It should be taken into account that researchers often need more information than other professionals: including historical series and going back to the source data / measurements.
An important development is the increasing number of freely available data sets, both by government organizations as well as private contributions, called Open Data initiative. The party that makes it data available for 'free' often has also (indirect) profit from this; e.g. via updates of data by third parties, better service from government to society, etc. In some cases, information is freely available for research but not for commercial applications. In such cases they need to be protected against illegal use by security measures through user authentication.
Geo-information (or geographically referenced information) has become one of our society's critical assets; indeed it is one of our foundations. Consumers, industry and government alike use geo-information in everyday situations. This has lead worldwide to the evolution of a geo-information industry, which has managed to develop products and services that help users answer all sorts of key questions that begin with: "Where...?". Geo-information is indispensable in the pursuit of policies in spatial planning, water management, environmental management, agriculture, energy supply, traffic and safety, to name but a few areas. Geo-information therefore has an essential role to play in the government's provision of public services and fuels one of the Netherlands' most innovative economic sectors.
The Netherlands, more so than most other West-European countries, is confronted with pressing needs regarding its spatial organization. These needs relate to the following:

- Being one of the most densely populated countries in the world, the Netherlands continuously needs to strike the right balance between urbanization and nature conservation, protecting areas with unique landscapes, flora and fauna.
- Having Europe's largest mainport within its territory, the Netherlands has to invest in infrastructure and logistic solutions to keep pace with the growth in mobility and transport and, at the same time, find solutions for the activity's environmental impact.
- Being a country with more than half its surface area lying below sea level, and one that is also situated at the confluence of major European rivers, the Netherlands, over the centuries, has felt the need for a sustainable water management.
- Being confronted with the imminent dangers of climate change, the Netherlands is now pressed to brace itself against a rise in sea level, changing climate and extreme river flows.

Typically, geo-information has a key role to play in resolving these issues. Adaptation to climate change is not possible without the knowledge of the places where the water defence structures are built on soft soil; water management is not possible without proper knowledge of local meteorological conditions; nature conservation is not possible without detailed information about landscapes, flora and fauna; and urban planning needs information on house stocks, underground infrastructure and traffic flows. In short: resolving today's issues regarding climate, water and spatial development calls for available, up-to-date geo-information.

Another important trend in society is the ubiquitous growth of mobile devices, many of which are location-aware, and allow their owners to collect data almost unnoticed. We expect that many of that automatically generated data will find its way into repositories, currently mostly associated with social networks on the web. This is an important new data source with strong spatial analytic value that we have yet to understand how to use optimally.
According to a recent survey (2008), the market for geo-information solutions is estimated to be 1.4 billion Euros in the Netherlands alone, and it is growing by 17% a year. Although these figures make clear that this market is

substantial and growth is impressive, average turnover of the related enterprises is 3.2 million Euros, which underlines the SME nature of most of these businesses: indeed, half of the companies surveyed had less than 10 employees. This clearly is an indication of an emerging, innovative technological sector. A sector that delivers services for water and environmental management, construction industry and built environment, oil and gas industry, agriculture, security and mobility. Over the past few years, Internet and Location Based Services (LBS) have extended this fertile ground of consumer applications. The success of virtual globes like Google Earth and Virtual Earth testifies to the reality of the great intuitive appeal of geo-information. Nowadays, driving your car, more and more implies being navigated by TomTom.

Other relevant information

The large-scale facility Maps4Science is positioned to be a sustainable component, consisting of geo-scientific information and services, within the Dutch research tools. The approach to achieve this is to jointly develop the four components. The embedding within academic structures next to the connections within the 'golden triangle' are anchor points for a successful and sustainable facility. The involvement of parties involved in the collection and provision of geo-information in the academic setting are crucial to the sustainable impact of Maps4Science.

Timetable

Duration of the project

Planned starting date 1 July 2012

Expected completion date 30 June 2019

The facility will be developed in a three-step approach (see chapter 6). After 1.5 years, the facility opens, and users can reach the current existing services but then in a seamless way.

	SETUP PHASE		OPERATIONAL PHASE				
	Year 1	Year 2	Year 1	Year 2	Year 3	Year 4	Year 5
Step 1: embrace current initatives		▼					
Step 2: integrating data sources			▼	▼			
Step 3: harmonising data sources					▼		▼
Decommissioning							▼

Declaration and signature

Have you requested funding for this research elsewhere?

☒ No

☐ Yes,

Please include details of any additional grants you have requested for this research project

Declaration

☒ By submitting this form through Iris, I declare that I have completed this form truthfully and completely.

Annex A: Other initiatives

Existing (supra institutional) initiatives in the Netherlands

University Libraries

Various Dutch universities have assigned the distribution of geo-information (traditionally maps) to an organization part of the University Library, referred to via names such as 'Geoloket', 'Geoplaza', 'Geodesk', 'Geodata Warehouse' or 'Kaartenkamer'. Following initiatives in other fields, umbrella organizations such as DANS and SURFnet provide the opportunity to a more collaborative approach for both data licence agreements and ICT solutions for distribution. The recent project 'Discover UKB maps', with the involved parties: TUD, VU, Geodan, DANS, Dutch Kadaster, SURFnet (and less the Royal Library as project name might suggest), resulted in an infrastructure that now operates in shared use by VU and TUD; see Figure 6.

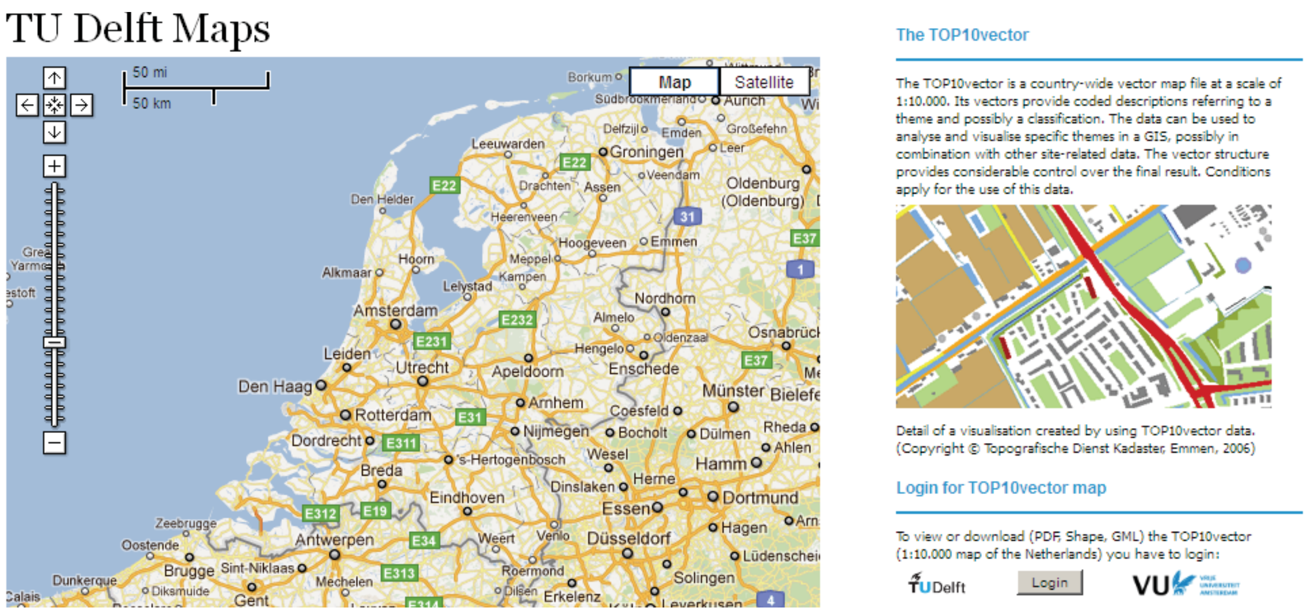


Figure 6: The operational result of the 'Discover UKB maps' project with options to obtain geo-information in various formats (shape, GML or pdf).

EduGIS: providing digital maps for education

EduGIS (www.edugis.nl) is an initative from the Space for Geo-Information programme and it provides on-line digital maps to a broad range of schools and courses (see Figure 7). The underlying geo-information is not directly available, except through the images. This is fine for visualisation but in general insufficient for research and analysis purposes. EduGIS is based on the Amazon cloud services to be scalable and efficiently serve the large numbers of users from the entire educational system. EduGIS also plays an important role as a sample portal in the European Comenius Network *Digital Earth*. Here, researchers and lecturers from various universities in Austria, Finland, United Kingdom, France, Belgium, Bulgaria, Greece and the Netherlands work together to make spatial media and geo-information accessible for education and educational research. The EduGIS use shows a steep growth over the last couple of years, not in the least because schoolbook publishers included EduGIS exercises. The last user survey conducted (2008) showed that in secondary education and the last years of primary education, 50% of the pupils used EduGIS several times (in exercises as a tool for analyzing social issues and exploring solutions).

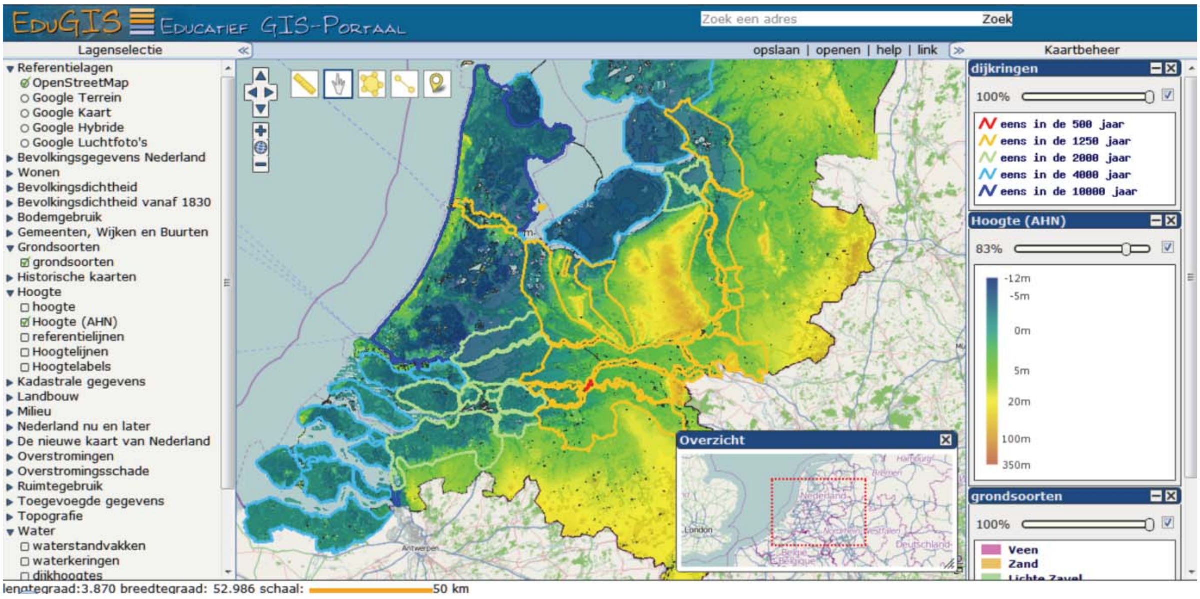


Figure 7: EduGIS with a large number of map layers (themes), in addition to the standard layers of Google Maps and OpenStreetMap, including satellite imagery (note most map layers are displayed here collapsed).

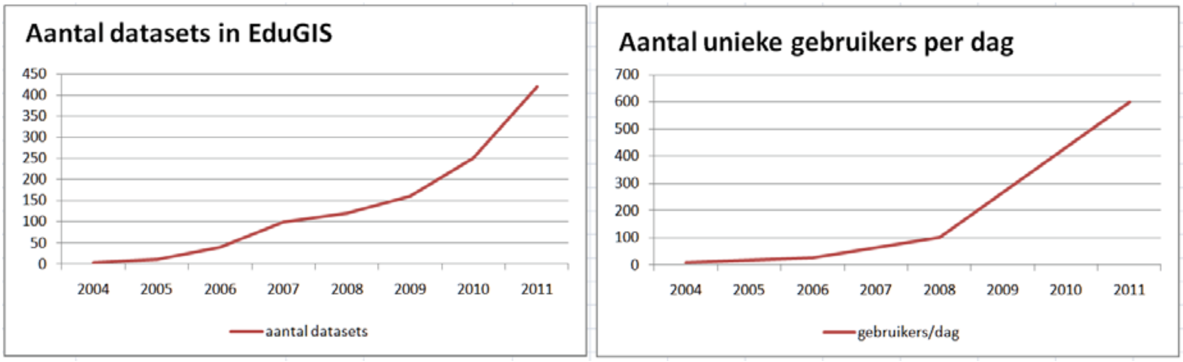


Figure 8: The EduGIS number of datasets over the years (left), and the unique number of users per day (right).

Related developments in other countries

Earth Cube

The National Science Foundation (NSF) in the USA is currently establishing '*Earth Cube*' as a national research facility for geo-information. The mentioned components of the facility are: 1. Governance; 2. Science Scenarios; 3. Cyber-architecture for Science; 4. Data Interactive Publications; 5. Semantics; 6. Processing, Models and Simulation; 7. Sensor Webs; 8. Curation and Archiving.

Senseable Cities

An example of using spatially oriented sensor data in Living Lab constructions, combining science, business and government is the MIT *Senseable Cities lab*; see http://senseable.mit.edu/. MIT researches the insights and possibilities of spatially oriented sensor data and design for studying and influencing human behaviour. A recent project is LIVE Singapore! (http://senseable.mit.edu/livesingapore), which provides people with access to a range of useful real-time information about their city by developing an open platform for the collection, elaboration and distribution of real-time data that reflect urban activity. Here, the scientific community together with government of Singapore are trying to learn ways to use large quantities of spatial information from maps, sensors, satellites and social networks to make Singapore a better city for its inhabitants.

EDINA

In the United Kingdom, EDINA (edina.ac.uk) is a scientific data facility, operational for quite some time already. It supports universities, colleges and research institutes in the UK by delivering access to a range of online data services through a UK academic infrastructure as well as supporting knowledge exchange and ICT capacity building, nationally and internationally. The latest EDINA user survey showed that undergraduates, postgraduates and staff, including information professionals from a broad range of disciplines, use their services. They found EDINA services easy to use, time saving and would recommend them to others. " (See Figure 9).

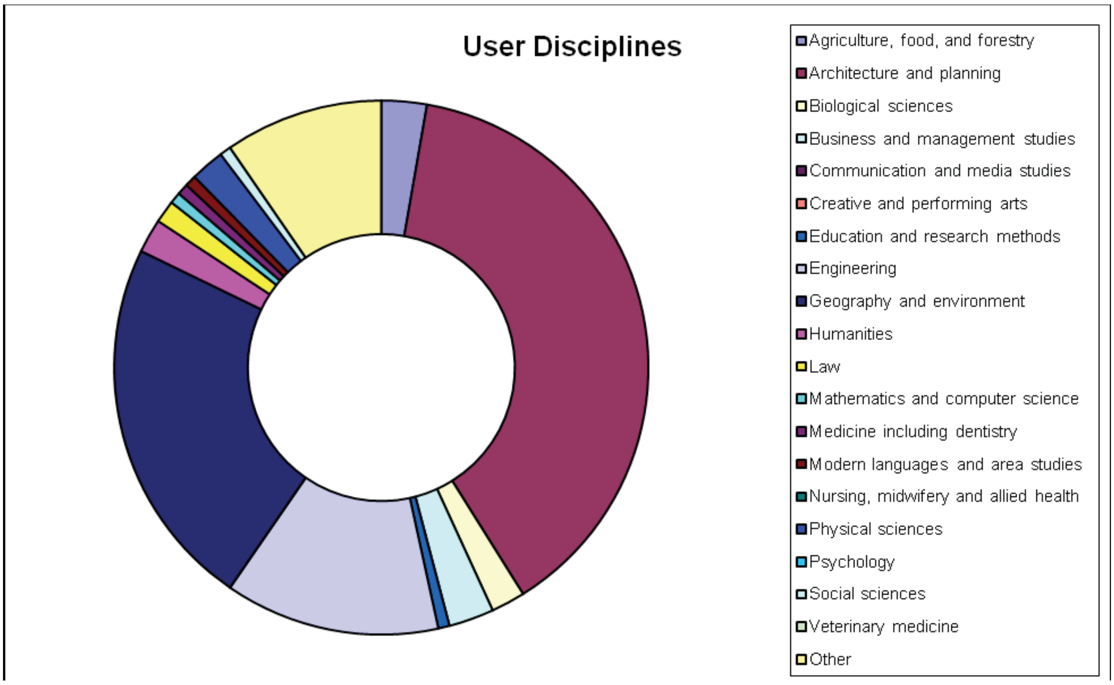


Figure 9: Example of the EDINA user survey analysis of the Geology Digimap illustrating the large variety of academic user disciplines, similar analysis is available for other geo-information products served by EDINA (source: http://edina.ac.uk/impact).

Annex B: Scientific use cases

Science case Health: Measuring and Forecasting the Spread of Epidemics

Prof. dr. Peter Sloot, Dr. ir. Alfons Hoekstra, UvA and Drs. Carl Koppeschaar, Science in Action

In recent years, a huge flow of quantitative social, demographic and behavioural data has become available, spurring the quest for innovative technologies to improve traditional disease surveillance systems, providing better localized detection capabilities and resulting in a broad practical impact. The availability of more and higher resolution spatiotemporal data, and a larger variety of geo-tagged datasets allow us to measure and forecast the spread of epidemics (of infectious diseases, and slow epidemics like obesitas). This should lead to 1) a Dutch National radar for infectious diseases; 2) data-driven models that allow forecasts of epidemics, and 3) forecasting different scenarios of containment/mitigation measures to support public health scientists and decision makers.

Epidemics are inevitably entangled with human behaviour, social contacts, and population mobility and mixing. Improved techniques and methods support the integration of datasets with geo-referenced information, economical and transportation databases. For the first time, epidemic processes can be studied in a comprehensive fashion that addresses the complexity inherent to the biological, social and behavioural aspects of health-related problems. Building upon our earlier work related to detecting, modelling and forecasting spreading of influenza and HIV, we expect breakthroughs through:

1. Include other infectious disease in the measurements, most notably (but not only) Lyme's disease and the noro virus;

2. Include other types of (slower) epidemics (e.g. obesitas) or age-related and work-related health problems;
3. Drastically intensify the collection of data by pushing to the limit all innovations as offered by the Maps4Science infrastructure, in combination with e.g. live geo-tagging of individuals;
4. Develop new data analytics by cross linking with as many social, logistic, demographic and behavioural geo-data available through the Maps4Science infrastructure;

To these ends, we need to both produce and consume several types of spatial data. Data on disease spread relies on volunteers and medical doctors who produce a dense and high-resolution spatiotemporal data set. Geo-information tools facilitate easy access and straightforward cross-linking of datasets to better understand the spread of epidemics. We also intend to exploit existing datasets on human mobility networks and on geo-tagging movements of large cohorts of volunteers, providing valuable data on human mobility and human proxy networks.
The Netherlands has by far the best datasets on spread of influenza (see www.degrotegriepmeting.nl), and other diseases. Further innovation and uptake of geo-information techniques and data will strengthen our position and should push further the development of a Dutch epidemic forecast measurement and computational infrastructure. In this way, we expect to maintain and improve the Dutch leading position in this field, but it would also strengthen European wide research and infrastructures for forecasting epidemics.

Science case Water resources: Better management through geo-information

Prof.dr.ir. Nick van der Giesen, TUD and Dr. Rob van Swol, NLR

Geo-information plays an important role in water resources management. On the input side, hydrologists use satellite images to estimate the distribution of water used by crops or the extent of flooding. Models are connected with spatial data through data assimilation to obtain the best possible state estimates to be used for management decisions. The final outputs are presented to decision makers in the form of interactive maps and ICT-based decision support systems. Presently, scientists involved in improving this information chain are not making use of an advanced cyber-infrastructure as is available to, say, geneticists or astronomists. Such an infrastructure is necessary to answer the many socio-economically relevant water issues in the world. Large providers of Remote Sensing data, such as the European Space Agency (ESA), are currently updating facilities to prepare for large data streams from new satellites and multi user access to this data. Being able to retrieve standardised data from these facilities is essential for the Dutch water sector to keep its leading international role.

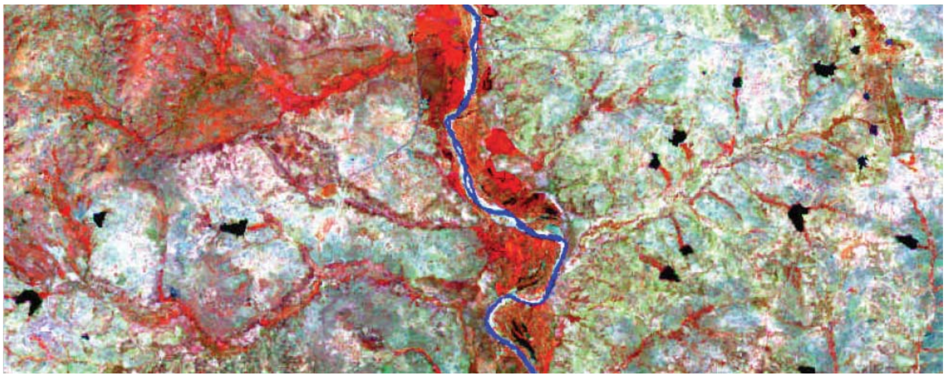


Figure 10: Mapping water resources from space, example Ghana, smallreservoirs.org.

In February 2011, 40 parties involved in the Dutch water sector signed the "Convenant Informatieketen voor Water en Klimaat" (Covenant Information Chain for Water and Climate, CIWK). CIWK addresses the need for spatial data management to improve water management by fostering cooperation between Remote Sensing, Geomatics and the water sector. In its advice to the government, the Topteam Water recommends to support the set of business cases developed by CIWK. These business cases cover a range of geo-informatics applications, from global evaporation mapping to dike monitoring. The infrastructure to be developed under the Maps4Science will enable universities and knowledge institutes to develop state-of-the-art hydrological and hydraulic tools.

Science case Crime: Geo-information and GI-Science as Crime reduction tools

Prof. dr. Marianne Junger & Dr. Lorena Montoya-Morales, Dept. of Social Risks and Safety - UT

Crime is related to place and the physical environment. However, research has been hindered to a large extent by the lack of suitable data. This is unfortunate because location is of utmost importance as a determinant of crime: a US study revealed that future crime is six times more predictable by the address of occurrence than by the identity of the offender. Crime science needs to address our understanding of location and the impact it should have on policy. Three example, yet untested hypotheses are:

1. *land use* and *accessibility* affect crime. Modern urbanism strives for intensive and mixed land uses and high accessibility and predicts that this increases safety. However, studies on crime suggest that this can lead to more crime, not less. Urbanism policies in the Netherlands apply the principles of mixed land use and accessibility.
2. *property price* affects crime. More expensive houses seem to be more at risk of burglary. However, the relation between crime and property price has not been studied at the building level because of a lack of available data.
3. *amount of light at night* affects crime. It facilitates supervision and it could be related to crime in different ways for different types of crimes. However, hardly any research has studied this relationship because information on a sufficiently large-scale to produce reliable data is not readily available.

Few publicly available datasets allow for a comprehensive study of these associations. We believe that many aspects of place and the physical environment are important determinants of crime but we know very little about their actual impact. This is unfortunate from the scientific point of view. It is also disturbing from the policy point of view as many of the issues described above are under control of policymakers and therefore have the potential of leading to important reductions in crime.

Science case Agriculture: Avian Influenza – Don't spread the disease

Ir. Henk Janssen, WUR

On a farm a case of avian influenza is detected. As in every outbreak of a contagious animal disease in the Netherlands, the Food and Consumer Product Safety Authority (VWA, Min. EL&I) puts procedures into action in an attempt to prevent the disease from spreading further. These are described in scenarios and for avian influenza in 2007. The procedures entail certain measures for quarantine areas surrounding the outbreak or outbreaks. In an area of ten kilometres surrounding an outbreak, additional checks are put into place. In an area of five kilometres surrounding an outbreak, protection measures are taken. Livestock within a three kilometre radius of an outbreak is vaccinated and livestock within a one kilometre radius of an outbreak is culled. The boundaries of the five and ten kilometre areas consist of topographical features such as roads, railroads and watercourses and should follow a ten or five kilometre buffer around the outbreak as close as possible. The three and one kilometre areas consist of buffers around the outbreak.

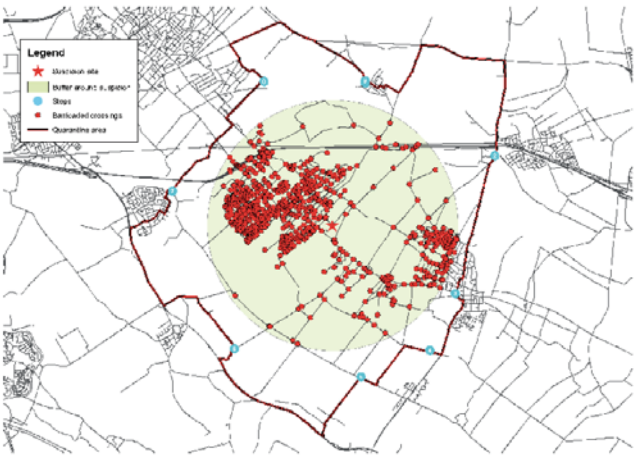


Figure 11: Quarantine areas around an infected farm

Workflows/business processes are typically based on spatial data and dependencies. Agriculture is a domain with many spatial components and the Ministry of EL&I has already been implementing spatial databases for years. Many agricultural business processes are supported by IT systems and these systems will benefit from the inclusion of GI technology. The GI can be used to minimize epidemiological risk and economic damages. Integration of GI technology in main stream IT systems has been addressed by research and software vendors, among others in the field of business intelligence. Analysts could then extract queries on these data, where location is the only relation between the features. Previously parts of such a system were implemented in research studies as prototypes. This allowed assessing whether the realization of such a system is feasible. Also the research shows that the amount of spatial data which has to be administered increases. A more important issue is the lack of up to date datasets. This compromises the reliability of the outcomes. Therefore additional research is needed in which in which Large-Scale Research Facilities will play a crucial role.

Science case Cultural history: The Integrating Heritage Program

Prof.dr. Jan Kolen, Prof.dr. Henk Scholten, Dr. Niels van Manen and Maurice de Kleijn, MA, VU

Parallel to the Maps4Science proposal, the CLUE institute (VU, WUR, UL, RUG) will submit a proposal for an Investment subsidy NWO Large. It aims to implement an NSDI for research of the heritage and history of the Dutch landscape, which was designed in a pilot funded by NWO (file: 380-57-001). The proposal aims to integrate research data related to the heritage and history of the Dutch Landscape and is therefore called the "Integrating Heritage Program" (the IH Program). The resulting NSDI would not only enable scholars to address queries specific to their disciplinary frameworks, but also facilitate interdisciplinary collaboration between experts in the field of landscape research and the humanities scholars – thus assisting the development of more holistic understandings of the historical landscape.

The Maps4Science facility will be a perfect basis for stimulating the interdisciplinary aims of the IH program and the building of its NSDI. By integrating the landscape NSDI into Maps4Science, landscape archaeologists and historians would gain access to additional information (helping them to contextualise their research) and data related to the historical landscape would become available to non-humanities scholars (reinforcing the interdisciplinary character of landscape studies).

Maps4Science will also benefit from such integration. It could draw from the strategies tried and tested in the NWO pilot for developing conceptual tools to facilitate interdisciplinary debate and for initiating and nurturing collaboration between scientists and non-academic partners. The two programmes could share expertise in data standards, user control and network connections between data sources. Therefore, although each of the two programmes could be a valuable addition to humanities research in The Netherlands in their own respect, their joint and combined implementation would serve the academic community best.

Science case GNSS performance: Support mission-critical applications by predicting GNSS performance

Drs. Judith van Bruggen–van Putten, NLR

The world of Global Navigation Satellite Systems (GNSS) is evolving. In 2010 EGNOS became operational and in 2011 the first official Galileo satellites will be launched. Numerous location-based services are developed, each with different demands on availability, integrity, accuracy and accountability. For example, when Safety-of-Life or mission critical location-based services are involved the trustworthiness of the signals is important and the error of the calculated position has to be within centimetres.

To gain insight into GNSS performance characteristics (availability, integrity, accuracy and accountability) permanent monitoring activities are required. Among others, the European Space Agency (ESA) gathers GNSS performance information spread out over Europe.

GNSS performance depends on the composition of the atmosphere. This research focuses on the design and update of models related to the troposphere. Relationship between characteristic of the troposphere, meteorological forecast information and GNSS performance characteristics will be identified and monitored. The expertise gained is valuable for the prediction of location-based and time-based GNSS performance. The Maps4Science infrastructure may support this science by providing the means for integrating geo-information with GNSS performance measurements.