FAQ Maps4Science

Version 1.2, 2 April 2012.

CONTENT

Note: The questions included in this FAQ were provided by the members of the two trial interviews (as captured by the M4S-team). The members of the two trial interviews are gratefully acknowledged for stating these critical questions: Peter Apers (UT, STW), Robert van der Drift (NWO), Herman Eijsackers (WUR, Min. ELI), Jacob Fokkema (TUD, NWO vz. ALW), Kees Kroese (o.a. ex-RvB Kadaster, Comm. van Wijzen), Peter Nijkamp (VU, ex-vz. NWO), Henk Ottens (vz. KNAG) and Wim van Vierssen (KWR Watercycle Research Institute). After the trial interviews, the questions were categorized into the above 10 groups and the answers were written down by the M4S-team. In addition, also the recommendations as stated by the members of the two trial interviews were documented. More questions and recommendations are welcome, and also suggestions for improvements to our answers are welcome.

1. WHY IS IT IMPORTANT?

FUTURE CHALLENGES

1.0 The angle is much more promising from the future challenges. (AB)
Future big challenges for modern society and science lie in Environment, Safety and Health (and reflected in our initial six scientific use cases). These challenges are becoming more and more complex in our networked society. As these challenges all have a spatial component, spatial information is one of the key ingredients to a better understanding and dealing with these challenges.
Spatial data is no longer just the result of a few experts measuring the world, but increasingly of huge masses of people using their mobile phones as well as an exponential growing amount of sensors from small, embedded sensors, to traffic cameras and satellites.
It is one the biggest challenges for spatial data sciences to deal with this new flood of data; to create order, to search and to combine it for the many purposes we have.
The Spatial CyberInfrastructure of the future needs search and analysis tools, methods, standards and resources to put these vast amounts of dynamic, multi-sourced spatial data to work. In many science disciplines, this data is needed, but adequate infrastructure needs to be developed. Seamless and painless data integration for this upcoming huge and complex dataset is urgently needed.
M4S is a distributed facility to empower the spatial information sciences with the infrastructure of the future. To allow spatial scientists to maintain their world-leading position through developing unprecedented science with this facility, and to enable other scientists in a large variety of disciplines to include spatial data and analyses in their research work.
M4S facilitates "spatial thinking" in other science fields. Spatial thinking is key for a competitive soc. See also National Research Council report (2006).
Pitch: *Future challenges are in the Environment, Safety and Health. These challenges are becoming more and more complex in our networked society. Spatial information is one of the aspects for better understanding and dealing with these challenges. We foresee a huge wave of spatial data coming from sensors, cell phones, and satellites that we want to use in new science. M4S updates and extends our current infrastructure to meet this challenge. These "social, dynamic geodata" is important to an expanding user community ranging from historians to medical epidemiologists.*

1.1 The proposal does not define clear challenges for assisting the described scientific domains in using GI. Could you elaborate on this? (AB)
The 6 scientific use cases have all been described by 'external' and in their field leading researchers. Within Maps4Science explicit project activities are planned to further explore the various challenges and needs from these domains for the Maps4Science facility. Moreover, it is a stated ambition of the facility to address the issue of ease-of-use, and accommodate scientists without the GIS technological background in the full and fair exploitation of the offered data resources. Such can only be achieved by combining properly the semantics of data and services with the understanding of those end-users' needs.

BOTTLENECKS

1.2 Why do you not reply (with achievements) on the three bottlenecks mentioned (in the presentation)? (Licence restrictions; Technical inability to handle data types (3D, massive data, …); Lack of awareness).(TvE)

The proposal takes care of solving all three mentioned bottlenecks. To summarize:

1. **licence restrictions**: Digital rights management is a separate science case ('*science case 2'*, Section 1.5, pages 14-16). Furthermore, open data policy is being applied for an increasing number of geodata sets (due to EU and NL policies). Finally, a national facility is a better counterpart for geodata owners to agree on w.r.t. licences.
2. **technical inability**: we identified 9 geo-information science research topics (together '*science case 1'*, see Section 1.4, pages 10-14), which should resolve the identified inabilities, and the results will be used to implement (materialize) the Spatial Cyberinfrastructure ('*science case 3'*, see Section 1.6, page 16 and Chapter 6, pages 24-28).
3. **lack of awareness**: we identified 6 scientific use cases: Health, Water resources, Crime, Agriculture, Cultural history and Navigation and positioning (together '*science case 4'*, Section 1.7, pages 16-17 and Annex B, pages 37-40) that will raise awareness in these fields. Together with the 'positive spread' effect and the fact that M4S is a national and recognizable facility (with communication budget, Section 5.1, page 23) this will resolve the last bottleneck.

Some more details to illustrate the first issue. Government is the single biggest collector of spatial data and therefore a very important supplier of data for the Maps4Science. Under the open data initiative governmental bodies will open up their datasets for the re-use without any restrictions. Key ministries like the Ministries of Infrastructure and the Environment and Economic Affairs, Agriculture and Innovation are amongst the staunchest supporters of this new policy. The first results of this change in policy are already in place. E.g. the important datasets Nationaal Wegen Bestand (National Roads Dataset) and Basisregistratie Topografie (Key Register Topography) have recently become available for re-use at no cost. Further, novel approaches to collecting geo-data are evolving; e.g. based on crowdsouring.

1.3 The proposal did not start in a strong way. It starts with complaining: "access is rather poor," "difficulties," "missed opportunities." So why would we bother with all your problems? First get your problems solved and then come to us for a facility. (AB)

Agreed, in the Netherlands we probably have the highest geodata density in the world and we have very good science, industry and government players (in top 5 according to various rankings). The proposal tries to draw benefits more from this good starting position, and even improve the Dutch position. Therefore, we analysed the potential for improvement, which may have sounded as 'complaints', but this was not intended as such. The main intention is to motivate and direct the development of a national geodata facility for research. The world is changing rapidly and data providers seek new ways to profit from their assets (i.e., the data). We need to deal with the interest of data providers (restrictive) as well as with the interest of science (progressive). Currently, spatial data is too often inaccessible due to license restrictions. Spatial data is also special in this case, as the geographic coordinate is a natural index to combine datasets and produce new, unforeseen combinations. M4S will address this in a dedicated, spatial Digital Rights Management construction, special for spatial data.

The new datasets are dynamic, formal (authoritative) and informal (social), multi-sourced, measured without purpose etc. M4S will build an infrastructure to deal with these new, exciting and huge amounts of data.

GI is an enabling technology. More and more scientists make breakthrough science based on spatial data and the spatial patterns detected. The majority, however, is unaware of the enormous potential of spatial data. M4S will provide easy access to these resources on-demand.

1.4 Is the field fragmented because of infrastructure barriers or are people working in fragmented modes, and how will the facility help us here? (AB)

The current fragmentation of geodata facilities is mainly due to the history developments. In the past the geodata was analogue (paper maps, aerial photographs) and had to be physically at the locations of the various Universities, often in the library or special branch thereof. Each University has its own emphasis and this is reflected in the available geodata sets, and every University used its

own limited resources to provide the most relevant data. In the digital age, these analogue facilities are replaced by digital counterparts and the physical aspect is less an issue with the growing availability of bandwidth. Each University (and research institute) looked for solutions and discovered that creating and maintaining these digital facilities is far from trivial. Given the successful example of a national geodata facility from the UK (EDINA), this provided a lot of willingness to cooperate. Also making all the data sets (of the various Universities) available for all researchers will be a big benefit. So, no fragmentation exists in the field, it is rather the opposite. There is an impressive track record of cooperation, probably not surpassed by any other country in the world. This is illustrated amongst others by a well-organized geo-business community in GeoBusiness Nederland (supporter, see their letter); an organized public sector through the existence of Geonovum (partner); four Universities (all partners), which jointly deliver the MSc GIMA; KNAW-NCG as our scientific platform and our academic libraries (nearly all partners, incl. DANS) having a joint platform 'UKB werkgroep kaarten en GIS'. The recently started programme ICT Innovation Platform Geo-information (IIPGeo, supporter) connects the well-established and organized Dutch geo-community with other (ICT) sectors.

1.5 It seems that everything today is in different databases and M4S just wants to link them. This will not impress NWO! (AB)

The geodata world needs a large scale facility to address the challenges of the future. Currently, universities are satisfying local needs with local budgets. M4S will be a distributed facility, meaning that it will connect the local repositories to become a large, connected and sustainable infrastructure. We must overcome: semantics issues in spatial data, governance of both formal and informal data, and user-producer gaps, in particular towards new, unexpected use. ICT is helping us a lot in this. ICT is not an issue, it is about how we apply it in the Geo-ICT for the benefit of the Netherlands' Science community. It is not only about linking, but also about making the geodata usable for scientific innovations in other disciples.

1.6 Can you mention the three largest problems that this project will solve? (TvE)

1) Science has become more and more specialised and a problem is the lack of an integrative, multidisciplinary view. M4S offers this integrative view — in a spatial way.
2) On a more concrete level, there is the problem that scientist do not have access to geodata, as for example Marianne Junger and Nick van de Giesen explained in their interview. M4S solves this problem once and for all.
3) From a technical point of view, the largest problem is handling of big geodatasets (including 3D space, time, scale and semantic aspects). This problem is even becoming bigger and bigger in the future because of sensors, satellite images, social media and the extra data that comes off of them.

1.7 Only 4% of all geo-information has ever been watched at, why is the geo-sector not able to improve this old statistic? (PvO)

First of all: The statistic was actually worse than the 4% you mention, at least globally. The statistic is related to the observation Al Gore 14 years ago (made in his speech on January 31st, 1998, California Science Center, LA.) on The Digital Earth, Understanding our planet in the 21st Century: "The Landsat satellite is capable of taking a complete photograph of the entire planet every two weeks, and it's been collecting data for more than 20 years. In spite of the great need for that information, the vast majority of those images have never fired a single neuron in a single human brain. Instead, they are stored in electronic silos of data. We used to have an agricultural policy where we stored grain in Midwestern silos and let it rot while millions of people starved to death. Now we have an insatiable hunger for knowledge. Yet a great deal of data remains unused."
We do not have actual statistics for the Netherlands and spatial data concerning the Netherlands. In general, we expect it to be more positive than the global percentage. And it needs to be mentioned that some spatial data sources find tremendous usage even today, while others find no usage whatsoever, yet applications for them are not hard to think up. The Netherlands has a better score

for a number of reasons: (1) we are a densely populated country, which means that commonly there are frictions over the multiple use of space, requiring spatial intelligence, (2) spatial information is fairly well embedded in a number of planning organizations, certainly compared to the global scene, (3) the quality of available spatial information in our country is very high.
The M4S consortium fundamentally believes that the quoted 4% (if true or nearby) is not so much a proof of the uselessness or lack of application of the data that is available, but more fundamentally an indication that the data exploitation is substantially hampered by a variety of obstacles. Those are closed data silo's, license and other protection schemes, old-fashioned data custodianship practices, sometimes a lack of sensible business models, and certainly also the lack of technical know-how with the potential users on how to work with the data. Spatial data has rather advanced formats, which require the data engineer good knowledge to exploit its fullest potential; that knowledge is not present everywhere. We thus aim at easing up the dissemination of the data, and also its various forms of usage.

WHY SPENT MONEY ON MAPS?

1.8 Why should we spend money on maps, health is much more important? (PvO)
Indeed health, safety, environment, and so on are the really important issues, which are all better enabled by M4S (e.g. the famous very first GIS analysis of Dr. John Snow: cholera-epidemic of 1854). M4S does not spend money on maps but offers possibilities for better understanding and working in fundamental issues like health. M4S provides the needed access to geodata and services. The actual maps (or geodata) are in most cases provided by partners and supporters (see the various support letters of the main government geodata producers: Kadaster, TNO, KNMI, RWS, CBS, etc.). We also need to emphasize that maps and geodata are not the same, much like eScience and ICT are not the same. Maps are a single, albeit important, type of visual product derived from geodata.

1.9 Why should the Netherlands spend money on maps? Why should the spatial information infrastructure prevail? (AB)
First part related or even equal to question 1.7 above. The spatial information is crucial for so many researchers, that a national infrastructure-based approach is the most effective, efficient, and sustainable solution (rather than having several individual or local University solutions). Also the various technological, organizational and legal challenges make it important to join forces to develop a national geo-data facility. The economy of scale will lead big players like the data owners to be more willing to connect and collaborate.

1.10 Is it worth spending this money and what are the valleys bridged by this. (TvE)
Spatial dynamics, i.e., physical/economic/social processes and their interactions, has become a key factor in our densely populated and highly developed society. All issues concerning the environment (emissions), climate change (sea level rise), mobility (economic and social impact), health (exposure), agriculture (food security), resource efficiency (sustainability), social cohesion, and many others, are issues with a dominant spatial angle. We have now the basic technology, the tools, the knowledge and the collaborative will to make M4S a facility to uncover the huge potential of spatial data for science disciplines, which until now only know spatial data from their TomToms and Google Maps. An investment in M4S is therefore also an investment in Health. It is also an investment in Climate. It is also investing in Mobility, etc. M4S will allow these disciplines to make a breakthrough in spatial thinking, and hence will make breakthroughs in those disciplines more likely to happen.
***Pitch**: Investing in Maps4Science is also investing in health, environment, mobility and many other sciences. It enables scientists in a broad range to include spatial thinking in their scientific tool kit.*

RELATION WITH TOP SECTORS

1.11 Do you think the top sectors need research and not just data?
What is in it for the minister(y) of Economic Affairs, Agriculture and Innovation? (AB)
In many top sectors (even the headquarters top sector) location is imminent. We have examples of scientists leading the field of spatially-informed research in their respective top sectors. Top sectors need data and the infrastructure to deliver it on demand. M4S is preparing for the Netherlands science community a unique facility for discovery, use and sharing of spatial data. See also the support letters of the Ministries of I&M and EL&I in this (and the supportive action of EL&I to establish Nationale Satellietdatabank, see support letter Netherlands Space Office).
Top sectors need research as well as data. Our project will create a facility that will improve access to data as well as facilitate the combination of data from various sources. Also, we will enable the reuse of methods and services and consequently enhance the necessary research facilities required for enabling breakthrough research. Such research is not born out of eureka moments, we believe, but is born out of bringing views and ideas together, letting them and mingle and meet. To allow for that you need to connect researchers, and one important way of doing this is bringing their data into a single facility.

SCIENTIFIC BREAKTHROUGHS

1.12 I like the scientific breakthroughs but you did not discriminate.
The facility should enable scientific breakthroughs, not just some operational problems. (PvO)
In the proposal we have 4 science cases: 1. Geo-Information (GI) Science, 2.Data rights management (and GII governance), 3. Spatial Cyberinfrastructure and 4. GI Use. In all these science cases, there are challenges. In GI Science, we identified 9 topics, and expect results, as smaller and bigger scientific breakthroughs. How and what precise, is difficult (or impossible) to describe in advance. But based on the track record of the GI Science community involved, it is likely that the breakthroughs will be delivered. Special emphasis is given to the challenges in (geo)data rights management. Also here, scientific (both theoretical and practical) breakthroughs are needed and together with results of the other GI Science research, the whole will materialize and be put together in the Spatial Cyberinfrastructure (SCI). As this is a very complex system, the realization itself marks a breakthrough. Then applying and using this infrastructure (useable geodata and related services) in many different scientific disciplines (6 scientific use cases identified), will result in breakthroughs in these fields of science also, we believe. But it is even harder to predict and describe in advance what these breakthroughs will exactly be. For one, M4S will equip the associated research teams with data that allows them to turn fieldwork-type of case studies into nationwide (wall-to-wall) studies, improving strongly the statistical basis for their findings.

1.13 Building a facility is not a goal in itself. Can you be more specific about the expected breakthroughs? (PvO)
Agreed, the breakthroughs are expected in the scientific use cases (when applying the facility) and before the facility can be created, also in GI Science and data rights management (see question 1.11). However, the facility is extremely complex. The facility is a very complex system with many aspects: data, hardware, software, network, standards, people, legal and organizational agreements, financial arrangements, etc. (see question 2.5). Putting all the needed pieces together will in itself be an unprecedented technological and scientific achievement.

1.14 You should make clear what are the scientific challenges. You need much more emphasis on the geographical aspects. (AB)
Scientists need good data. For many science disciplines, the field of location data is extremely challenging as it provides unprecedented insights in human behaviour and activity as well as in monitoring and organising our environment in the broadest sense. M4S is a facility that will empower scientists to make breakthroughs with and within the spatial data domain.

Inside the spatial data domain, breakthroughs can be expected in dealing with spatial data scales and heterogeneity, in dealing with geo-semantics, with the particular digital rights aspects typical for spatial data as well as discovery and analysis algorithms for spatial data mining in the exponential growth of location based information.
In other science domains, spatial data will become accessible and attainable allowing subfields to flourish such as spatial economics, spatial health analysis and spatial crime investigation.

DRAAGVLAK

1.15 Do the partners believe in the proposal? (AB)

The partners have a very strong believe in the proposal. This is based on many decades of using geodata in science, but with the transition from the analogue to the digital era the playing field is changing. The partners, or actors within the partner organizations, can roughly be categorized into three groups which each have their own motivation: 1. Libraries, 2. GI scientists, and 3. Other scientists using GI. The organizations responsible for the geodata provision (libraries) have experienced the difficulties (based on technological and organizational challenges needing better solutions). Therefore, there is an enormous drive to collaborate. Also, the Dutch GI scientists are highly motivated to find solutions for these challenges (and are in the right position to do so, based on their track record) and making these solutions all work together in a national research facility M4S. Finally, all the scientist from different disciplines state more or less the same: we are not able to properly conduct our research without a facility such as M4S (and they are therefore highly motivated). Cooperation in the GI sector was and is strong. E.g. joint master programme (GIMA), Geobusiness NL, NCG etc. There is also a high level of mutual trust. This provides a competitive advantage in this field for NL. Now, in the future (with M4S) and in the past centuries.

1.16 For the end-users in the water sector, this proposal is not in the innovation contract for the water sector. We (from water sector) will never step into such a project with such thin commitments. (PvO)

Being a national facility with national responsibility, M4S will be in a good position to create relationships and collaborations on solid commitments (or even as contract partners). *Also impact of the relationships will have national impact, so making it more attractive for others, than the partners, to collaborate.* The many support letters that we received give a clear indication that also outside the consortium there is a strong support for the proposal.

1.17 Which universities benefit from M4S? (AB)

M4S is an initiative of 10 partners led by TU Delft, each of the partners realizing that their current level of investment is not enough to reach scientific excellence in spatial thinking. We need an intervention to push it all to a higher level. We need a national initiative to boost this development. And M4S will do that. It will provide structure and propulsion to collaborations and harmonisation of local initiatives. In 7 years from now, it will be an unstoppable and undeniable facility, shared and distributed, that will create new science and great science. It will provide Dutch scientists in a variety of disciplines and universities an unprecedented data facility for unleashed results. In principle, the facility will be open to all scientists affiliated with a Dutch university.

2. WHAT IT IS

2.1 Consider the metaphor of the timetable for trains. A complex timetable for the variety of arriving and departing trains. Make it clear and simple. (PvO)

The metaphor of a train timetable / schedule / plan is slightly beside the point. We totally agree, M4S is a networked cyberinfrastructure facility, but rather than creating a central control structure, we aim at facilitating a networked environment. To make this operational, it is crucial that agreed upon protocols and standards are applied. This is slightly different from the suggested 'train timetable', as our major concern is to facilitate interoperability between various heterogeneous data sets and services, respecting various use rights, and even more so it concerns very rich and heterogeneous data (types) and services from many different origins to be used by users with a wide range of skills. The M4S consortium prefers the metaphor of a utility network: also with many producers and consumers involved, but always having a clear interface (connections) to access content and services (as with 220 V). Also, the facility is built for the collective community, but will be beneficial for the many individuals in it.

2.2 The example showed more the world of Google rather than the world of GI science. Is it, that this is all new to you? Is your distance to Google too large that you do not recognize that you should use their tools? Haven't you been overtaken by Google? (AB)

Google Earth and Google Maps belong to the class of applications (which includes the 'Digitale bosatlas') of which the goal is to produce visual map images. These applications do *not* deliver the scientific data needed for all kinds of analyses and simulations. Of course, M4S will use similar web-based cartographic techniques that Google, Esri's ArcGIS Online, and the required other players in the Open Source domain (such as OpenLayers and others) have developed and applied. These techniques will definitely be used in the (web-based) visualization part of M4S. But maps are only one kind of end product.

Both Google and Esri have provided their support letters: making their solutions available for M4S, but also showing interest in new solutions to be developed by M4S (to be potentially included in their future products). Unfortunately, Google is not an open model aiming for reuse (except for reuse by the company Google Inc. itself. We aim to benefit from linked open data, i.e., reusing data to the extent possible, but while respecting rights and obligations given various right holders' rights and (international) legal obligations, including the right on privacy and non-intrusion.

2.3 Is the GI research not going in the wrong direction: high quality, expensive data? (PvO)

GI research has many faces: from data acquisition (sensors), to interpretation, storage, analysis (processing), dissemination and use in applications. In many cases, the costs go down; e.g., cheap positioning possible (with GPS), more and more automated interpretation (instead of humans manually extracting features of aerial photographs), geoweb-based solutions providing very efficient low cost access to good data sources, etc. So, most of the time, GI research has resulted in lower costs while providing better data. In the GI field, data quality is often defined as the most suitable, cost-effective level of data fit for purpose (so, not per se the highest quality data). Of course, it is true that satellite missions are not cheap (and neither is GPS or Galileo), but so many users are benefiting that the cost per user is actually low. For some applications, highly detailed (accurate) data is needed and possible; e.g., to monitor deformations of buildings or the movement of tectonic plates. M4S will provide an infrastructure that will feed the GI and other research with all different types of data, so it will stimulate optimal use of the development of different types of data, both authoritative and volunteered.

2.4 What quality level are you aiming at for the facility and how will that compare to similar facilities internationally? (PvO)

The quality level should be sufficient to support a wide range of researchers with geodata and associated services. At the moment, there is a single similar facility available in the world: EDINA in the UK. EDINA (Digimap) has been in production for more than 10 years. Outside the scientific community there are quite a few (government and industry) developments within the Netherlands (and in the EU) to provide geodata to users. However, M4S includes more data sets (to support more researchers) and advanced functionality. To realize M4S functionally, it is foreseen that significant R&D has to be conducted. Once in operation, M4S would be the most advanced facility in the world and the NL would be in a good position to team up, or even lead further international facilities (an EU M4S). The high-level GI position of the Netherlands, both scientific and in the availability of the data (see also the text in the theme "Who"), is the base for this high ambition. This ambition fits very well with the current developments on open data policy, and the ambitions of the knowledge economy of the Dutch government.

2.5 Can you explain the facilities? (PvO)

It is about geodata, geodata, and geodata (or 'maps, maps and maps') and services to use them, for example the 3-dimensional spatial analysis of a phenomenon. A bit longer: geo data acquisition and storage, geo data processing services (analyses and simulations) and geo data dissemination and use. To realize this many components (of very different nature) have to be in the right place. The facility is a very complex system with many aspects: data, hardware, software, network, standards, people, legal and organizational agreements, financial arrangements, etc. Good examples of this are the EU INSPIRE directive (infrastructure for spatial information in Europe) and the US EarthCube. More details at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32007L0002:EN:NOT on INSPIRE with its 5 components: 1. metadata, 2. data specification, 3. network services, 4. access and rights of use, and 5. monitoring and reporting. More details on the US NSF call for Earth Cube proposals (at http://www.nsf.gov/geo/earthcube/) with its 8 components: 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.

2.6 We already have the 'Digital Bosatlas', why is this not sufficient? (PvO)

The 'Digitale Bosatlas' is a web application that enables high school students with a fairly large but fixed set of data layers to make maps for their educational projects. It should be understood that maps are static images in this application; ie jpg, png files, and not data sets. Finding correlations between different data sources is not possible. The focus is therefore on the cartographic end product. The user of the 'Digitale Bosatlas' has no direct access to the source data, and can only process the data source to map.

In M4S the focus is on freeing the data for scientific use, for example, but not only, in combining spatial data sets to find correlations. The map is just one (1) way to get those results then visualize a finished product. The real scientific work (data analysis, cross-correlation with non-spatial data only locally available to the user) is anterior.

Also we expect that the amount of spatial sensor data will dramatically increase over time. This data may be an important data set for social scientist, but only after their processing the data may lead to data interesting enough to be displayed as a map on top of other maps such as the digital Bos atlas.

2.7 Is this yet another GII? (AB)

To serve the six scientific use cases (Annex B of proposal) and other scientific use cases, Maps4Science will have to go well beyond the current state of the art w.r.t. GII capabilities. The following functionality will have to be included (and is today nowhere in the world available): geo semantic-web support, good and complete 3D/4D/5D support, explicit and usable links between resulting map data and source measurements/observations, be able to handle extreme large amounts of geodata, supporting geo-processing services within GII, delivering back geodata (and

parts of processing chains) through to the GII by the 'users', real-time streaming of data from live sensors, etc*. It is not another GII it the future of GII*

2.8 What is the bottleneck of the project (license agreement)? (TvE)
The overall complexity and the many different components of such a facility (organizational, legal and technical components all have to be at appropriate level and fit well together). How do all the components fit together, also given the very high ambition level: many different types of data and rich functionality (but easy to use). We recognize that data licenses are important issues, therefore we included a work package to address digital right management, allowing the users and providers of data sets to work compliant with (international) regulations, and enabling purpose limitation of data sets etc. Especially with the current trend of linked open data this aspect deserves sufficient attention, which is also true for software licenses (processing services).

2.9 At which target group is the project aiming at? (AB)
The target group is all researchers with a spatial component in the objects or people they study. To provide this very wide range of researchers with geodata and services. Starting point is Annex B of the proposal, which lists 6 scientific use cases in very different disciplines (Health, Water resources, Crime studies, Agriculture, Cultural history, Navigation and positioning), but many more disciplines could have been mentioned (a lot of them are still not yet aware of the possibilities – 'onbewust onbekwaam' in Dutch). See the results of the RGI project "GI in location", which resulted in the book "Geospatial Technology and the Role of Location in Science" by Henk Scholten, Rob van de Velde and Niels van Manen (Springer, 2009), with numerous examples for use of GI in Science (such as economics, Piet Rietveld). This book was an important inspiration for starting Maps4Science. A parallel was made between the role of GI and the role of a microscope in science. The EDINA statistics show that users are now really from all faculties and disciplines. Some reasons why researcher might be different (sometimes) compare to normal geodata users:

- Looking for unexpected relationships (more strange data combi's)
- More interested in source measurements (reinterpretations)
- Large areas/regions of interest, more data, discovering patterns
- Producing more own data (from measurements, computations)
- More requiring w.r.t. 3D and temporal (not only 'now') support
- Wanting to share geo-processing chains

Non geo-info professional (medicine, history, economics, etc.)

2.10 Can you make up a metaphor for the facility? (first train schedule, geodata web shop, bol.com for geodata) (PvO)
Several metaphors are possible (and more or less appropriate): but we favor the utility infrastructure metaphor enabling a geodata science community (see also question 2.1).

2.11 Do you have a slogan that captures the essence of the facility? (geodata on demand) (PvO)
Realizing the earth nerve system: geodata, enabling scientific breakthroughs based on real world information.

2.12 Will you also be using geodata from locations outside the Netherlands? (PvO)
Certainly, the research topics of various disciplines do not stop at our national border; e.g., comparing the economies of Rotterdam and Antwerp or studying crime differences between Enschede and Münster, or the spreading of a virus within Europe and beyond, or the river basins in Western Europe, etc. However, we have to be realistic and our initial geodata focus will be on Netherlands data (and even within the Netherlands select the most important types of geodata and provide these first). The priority ranking will initially be set by 1. the 6 scientific use cases (Annex B of our proposal) and 2. the currently available data sets (at the local University facilities and DANS).

The current map content of EduGIS can be seen as typification of future M4S geodata content: most data (themes) within the Netherlands (among others bases on our Dutch system of 'basisregistraties' and in collaboration with PDOK, our national government portal for geodata), then Europe (among others and where possible with content as defined by the 34 INSPIRE themes, based on agreements with the various countries; e.g., EDINA in de UK), and finally global datasets (in context of the GSDI development; see also their support letter).

2.13 There is a strange mix between high quality/expensive data and fast/cheap data. Please explain? (TvE)

M4S does not discriminate between expensive or cheap and various levels of quality and speed of collecting and disseminating. In principle all types and flavors should be supported. M4S will build up experience in working with those different types of data, where as the traditional SDI focus on traditional data. Note that high quality data (i.e., accurate and up-to-date) does not have to be expensive per se. Data costs are relevant, but outside the budget of Map4Science (and mainly covered by the data owners, providers as indicated in the support letters).

2.14 Also explain what is already available (AHN, Topographic Maps) and how accessible/ usable this is 'SDI 1.0'. Then explain 'SDI 2.0'. (PvO)

This is a really long story as there are really a lot of data sets (e.g., INSPIRE identifies 34 different themes relevant at EU level, but for sure there are many more types of geodata; e.g., have a look at the EduGIS, but also think about the many themes as collected by CBS, which very often also have a spatial component). Further, the availability of data and the cost for using the data are different issues. In general there is a trend that (geo) data becomes more open and freely available. In addition quite often data owners make their data available for free to researchers, but not always for commercial applications. The mentioned example: the topographic map 1:10:000 (produced by the Kadaster) has been 'available' via DANS to all researchers without cost for the last couple of years (not per se well known to all potential researchers and not per se easily accessible and usable, therefore VU and TUD developed a web based facility for this: http://maps.tudelft.nl/). This same data set has become available for the whole society for free per 1 January 2012 (i.e., paid by our national government, Ministry I&M with our tax money). For the other mentioned data sets AHN-2, the owners (RWS and all the 'Waterschappen') have the policy that the data is freely available for researchers (but with costs for commercial and other use). However, due to the incredible volume of this dataset, there is no efficient way of getting the data to the scientists. For example, within the TUD Kaartenkamer the state of the art 'approach' is: the scientist goes to Kaartenkamer with (set of) hard-disks and after discussing the needed content (area, flavor of AHN), the Kaartenkamer staff start working on the request and give the scientists a call when the hard disks are full with requested data. This may sound primitive in 2012, despite the fact that TUD has a good (high-speed) network and ICT Infra. Clearly this should be improved in the future. Maps4Science will play the role of coordinating and stimulation organization for this.

3. STRUCTURE AND PLANNING

3.1 Start with the structure of your 'sales' pitch. Do not start with 'all singing and all dancing' system at once. State "If we do not have, then the Netherlands will fall back." (AB)

We need Maps4Science to upgrade our local facilities and join forces to deal with the huge amounts of new spatio-temporal data. We need a national facility to manage this wealth of geo-data, we need a national facility to create cohesion and collaboration. We need to solve the issues to build such a facility together.

The Netherlands has a fine reputation on spatial data infrastructure. We have reached the top 5 of the world in GI Science. Maps4Science will provide us an impulse to stay on that position and even improve this further. Maps4Science will not only help the GI sciences, it will also enhance the use of GI data, tools and methods in many other sciences like epidemics, history, hydrology and criminology and many others.

3.2 More clearly explain what infrastructure is. Others see it differently. Databases alone are not the infrastructure. You must do research to get the facility running. Compare with Parelsnoer, GOFs and infrastructure proposals. Make it modular. (PD)

M4S will enable scientists to search and analyse, to store and retrieve, and to re-use spatiotemporal data from anywhere in the Netherlands. Although the core of the facility itself will be a bunch of connected data sets, M4S makes the effort to make it work and interoperate. Spatial data have specific characteristics needing specific attention. The integration between datasets is on the geographical coordinate and on time. We handle data describing complex systems, so we need to standardise on what different indicators and values mean. We work with representations at different spatial scales. We need to deal with privacy and other digital rights issues when combining datasets. We want to handle data coming from the explosive growth of sensors, mobile users (and content producers) and so forth. If we manage (which we do not doubt), we will be the first to achieve such a system at national scale, linking 10 excellent research organisations and facilitating the whole academic community in the Netherlands. We expect scientific breakthroughs when we enable scientists in domains that discover the power of spatiotemporal analysis, and are in great need of data, tools and methods. Epidemics, hydrologists, historians, criminologists and many others have great expectations in discovering the spatiotemporal aspects within their objects.

So what is the infrastructure in M4S?

1. It is connecting academic data repositories
2. It is integrating the spatiotemporal data within these repositories into one distributed database
3. It is upgrading the local facilities into a national one, giving it the power to deal with the expected wealth of geodata coming from sensors, mobile devices, cameras, satellites, and so on.
4. It is enabling scientists within GI to handle and deal with vast amounts of data from informal sources, to deal with the wholes and uncertainties, to deal with multi-scaled information.
5. It is enabling scientists in other domains to organise their data in a spatiotemporal way, and to combine it with other spatiotemporal data.

***Pitch**: M4S infrastructure needs to be developed to deal with the wealth of geodata and to allow scientists in all disciplines to make sense of it. It is hardware and software but most importantly, it is brainware: harmonisation, standardisation, organisation, governance, quality assessments and usability.*

3.3 The research components should be a 'sine qua non' that are required for the realization. (PvO)
Correct. Due to the functionally very high ambition level, it will not be possible to realize the facility without conducting research in a number of areas. We identified 9 of these topics in 'science case 1': GI Science. The initial research results (the rows 'R&D geo governance' and 'R&D GII' in the table on page 23) will then be used to create the initial version of the facility (row 'Spatial Cyberinfrastructure' in the table on page 23), which should be operational after two years. Meanwhile, the GI Science continues (row 'science case GII' in the table on page 23) to enable more of the required functionality for the later versions of the facility. Also feedback from initial use (row 'Demonstration science cases' in the table on page 23) will be taken into consideration, while conducting the research.

3.4 The NWO committee is concerned with infrastructure: it does not fund your research! (PD)
Maps4Science is a proposal for a large scale research facility. However, it is very ambitious and it cannot be realized with only state-of-the-art technology. Hence the need for accompanying research efforts. We believe that such research and development fundamental to the facility are fundable.

3.5 The set-up and operational phases are mixed. (PvO)
The phases are connected but not mixed. In the set-up phase, the (current) local facilities are still used to provide geodata. During these two years, the first version of the national facility will be developed, based on the GI research and also taking into account the solutions already available in the local facilities.

3.6 GI research should not be the core of the facility and user research should be done beforehand (PvO)
The GI research is important, but it is not the core of the proposal (that would be the Spatial CyberInfrastructure itself). However as explained under question 3.3, the proposed GI research is needed to be able to create the facility. The proposal has been written with the collective experience of the current local facilities (Libraries)as a basis, and with respect to the user research. However, this reflects current practice, and not the future ambition level of scientific use of geodata. Therefore the 6 scientific use cases have been selected, and they are expected to provide indicators also for future requirements. By error, these are now scheduled only in the operational phase (row 'Demonstration science cases' in the table on page 23) and they should actually also be active in the set-up phase and provide future requirements.

3.7 It seems an unhealthy mix of science and infrastructure and it is all about goodwill. (PvO)
The planned science within the M4S project is needed to realize the facility: 1. the scientific use case to initially provide requirements, and later on to assess the quality of the realized facility, 2. the GI science is needed in order to develop a SCI 2.0 (also see questions 3.3 and 3.6). Commitment of partners is more than goodwill (it is based on a project agreement with at least 25% own funding). Also the support letters are followed by actions in the past half year; e.g. more geo-data becoming open (of several of the supporter organizations).

3.8 Phases overlap and should have the GI science before the actual construction. Also more attention to the GEO-specific aspects. (PvO)
The proposal might be confusing (or be too compact) in this respect. There is indeed GI science (rows 'R&D geo governance' and 'R&D GII' in the table on page 23) planned in the set-up phase to develop the initial facility. These are very 'geo'-specific! However, not all functionality can be delivered in the initial facility, and therefore R&D continues also during the operational phases (row 'science case GII' in the table on page 23), to develop more advanced functionality (rows 'Maintenance' and 'Evaluation and Improvements' in the table on page 23).

3.9 How do you position GI-research in the project? (PvO)
The table on page 23 in the proposal indicating the planning of the GI Research may be confusing. Those activities will be carried out in the set-up phase (rows 'R&D geo governance' and 'R&D GII'). The GI research in this project provides the necessary methods, tools and standards to explore the distributed M4S facility. The 9 topics of GI research mentioned in the proposal are all needed to focus on the collaborative aspects of the facility. The GI science is needed to build the facility.

3.10 Are you allowed (within the NWO subsidy) to spend money for GI-research intended to build the facility? (PvO)
Without the GI-research, the efforts to build the facility are merely standard 'off-the-shelf' developments that are not of extraordinary scientific importance and will result in a standard infrastructure (SCI 1.0, as possible today). The GI research topics need to be carried out to make the facility worthwhile and they are needed to realize the SCI 2.0 of the future: Maps4Science.

3.11 The proposal contains an "unhealthy" mix of research and infrastructure development. Please explain. (PvO)
The research is needed to develop the software, achieve standardization, and the organizational parts of the infrastructure, and to make it work. As most universities have a local infrastructure in place, hardware investments are mainly for a central facility.

3.12 Name three concrete results that the project will achieve in 2 years, and in 7 years? (AB)
M4S will be the first 2nd generation SCI that will handle both formal and informal data of both static and dynamic nature. Data coming from numerous sensors, mobile devices, satellites etc. and measured by robots, professionals and citizens.
In 7 years time, M4S will have enabled 6 science domains to use spatiotemporal data and methods in their research efforts. This may have led to several major breakthroughs. Scientists in epidemics, crime science, hydrology, agriculture, cultural history, and many more will use spatiotemporal data as a new resource for data mining, data integration and advanced forms of analysis.

4. FOCUS

4.1 When I read the whole story, the headline was "I have a dream ..." But it is not concrete, not tangible. (PD)
The M4S is certainly an ambitious proposal, however it is phased in such a way that even after two years a very tangible result is foreseen. That result is measurable in terms of number of end-users, response times to requests, number of data sets and data themes supported, and the volume of data held, as well as those data sets that are virtually linked to. (also see 3.12)

4.2 There is no focus — if anything is possible, it is not realistic. (PD)
The M4S facility aims in the long run to serve all scientists in The Netherlands on whatever spatial data resource they may need for their work. This includes explicitly computational resources. But it is a long-term ambition that we hope to reach via a carefully laid out strategic pathway of development. That pathway picks the lowest hanging fruits first, and gradually improves on data coverage, functional coverage, and ease of use for the end-users.

4.3 What are exactly the questions from the scientific disciplines that you state to serve? (AB)
This question is best answered by referring back to the video quotes from the four scientific use cases at the beginning of our presentation today (also see the textual description of the six scientific use cases, in Annex B of the proposal). These are just a limited number of examples. In the book "Geospatial Technology and the Role of Location in Science" by Henk Scholten, Rob van de Velde and Niels van Manen (Springer, 2009) there are many more examples for use of GI in Science. But down to Earth: they somehow all need (a variety of) geodata and geo-processing services (selection, combination, aggregation, transformation, interpolation etc.) to conduct their research.

5. GI vs ICT

5.1 The case for eScience has been made before, and this is not exciting anymore. However, this proposal concerns a very specific type of eScience. (PvO)
This is a correct observation. M4S will deliver tools that enable eScience in a very special and important way: by leveraging the location information that is available in many datasets nowadays, it will allow non-GI researchers to construct rich characterizations of geographic context. This requires ICT of a special breed: Geo-ICT. But also for many researchers that do not consider themselves geo-scientists, geodata, for example the growing amount of data sets that contain sensor data, provide a rich source for scientific study.

5.2 Some members of the NWO committee are allergic for the statement that eScience is equal to ICT. (PD)
The two are not the same. M4S will deliver a facility with which to conduct innovative new forms of science, that is, eScience that is fully location-enabled. To create that facility requires baseline ICT, and on top of that geo-ICT, which consists of a rich family of spatial data types, semantic models for such data, and specialized operators to compute with those resources in combination. Generic ICT research does not address such issues as spatial uncertainty, spatial neighborhoods, spatial statistics, et cetera.

5.3 Quite a few challenges are ICT challenges, but not many are GEO challenges. (PvO)
This appears to be a slight misconception, probably caused by unclear communication on our side. We will focus on the specific Geo-challenges. We do not claim (it would not even be possible) that we will solve generic ICT problems that have shown to be hard to crack within ICT. Digital rights, security, interoperability are phenomenal problems by themselves. But the spatial data domain has historically had a rich but overseeable family of data types at its hands, and has been relatively successful in developing standards that allow data reconciliation. At the computational front a body of standardized functions is slowly emerging, with good practices being also developed. There is also the intrinsic uncertainty that always comes with spatial information that we are familiar with, and which requires special computational attention. The challenges ahead are now especially in reaching consensus on bridging differences in spatial semantics.

5.4 It should focus on the geo – that is what is new. (PvO)
A number of innovations is coming our way, and is giving new challenges. The paradigm will remain that location is the binding factor. New data types are arising: 3D data, variable-scale data, temporal data around spatial processes, sensor data to calibrate environmental models or to create new models from, volunteered spatial information with much larger uncertainties than we are accustomed to. Data sources that only implicitly carry location information, e.g., through text, requiring spatial information retrieval also are forthcoming.

5.5 You have a hard time convincing me you can solve the standardisation / interoperability by just building this facility. (TvE)
We cannot and will not solve the generic standardization/interoperability problems of this world. But we are `lucky' that the GI domain is not as wide as that, and that the S/I problem has been recognized for a long time already by many global players, such as ISO, OGC and INSPIRE. Various consortium partners have been involved in those bodies and continue to be so, in areas like spatial feature standards, web services, sensor web enablement, and we are not novices in the field. The open challenges right now are in the new data types, and also in thematic semantics and the geosemantic web. We are at the forefront of those developments. The M4S proposal has the

explicit support of the leading organizations in the standardization field; see the support letters of ISO/TC211 and OGC.

5.6 You also mention license problems. Quite a lot of publishers have digital rights problems. You seem to give the impression that you can work with it. Which unique approach will you have for your DRM issues? (TvE, PD)

The current license problems are in place because data producers have difficulty developing sensible cost models, and are treating all their customers identically. They are generally receptive to needs of scientists (and are willing to provide geodata under very favorable conditions), which is what M4S is catering for, but they have no canned solution for scientific requests for data. This is causing much of the current desk hassle that we mention in the proposal. Much of that can and will be alleviated through data convenants, meaning to say that part of the solution is in organizing it properly, not necessarily solving it in the technology. Some technological solutions, however, do exist or are within reach, such as spatial aggregation, and intentional data degradation.

A related issue is the governmental dynamics, which is pushing towards bringing more and more data into the open domain. Some data producers are waiting for those developments to materialize in full, before making a move into the open domain itself. These shifting tides give reasons to expect that over time the value is not so much going to be in the data itself, but rather in what you can do with it. Our DRM work will focus strongly on modeling those conditions of use.

5.7 What is the relationship with ICT research, specifically database technology research? (PvO)

On the connection with the current database research agenda, it depends what you are looking at. In the earlier collaborations; e.g., the Bsik RGI project GeoInfoNed extended the Dutch research database Monet with spatial data types. The PathFinder extension to Monet is currently being extended itself with spatial data types and operators to accommodate moving object applications, amongst others in context of the Commit project TimeTrails. Via collaborations with industry, new spatial data types have been prototyped in research and later on implemented by industry (e.g., the 3D solid data type in Orcale spatial). NoSQL databases and map/reduce techniques will certainly have their application in spatial data-intensive applications as addressed in GI Science, but GI researchers are exploring these developments. What remains specific to GI and large data sets are among others: (1) data joins are often not on exact value matches but on spatial similarity (ie., notions of neighborhood) and (2) spatial semantics need to be captured fully to run on inference mechanisms.

In general, Geo-ICT research picks up results from ICT research where this appears promising for spatial applications, and adds the complexity of spatial functionality to it.

5.8 How does this project relate to the eScience project? (PD)

The mission of eScience is to catalyse and enable the use of ICT in data-intensive, multidisciplinary research. The M4S consortium focuses on extending that same mission to the use of the growing and already rich spatial data environment in The Netherlands. That proposition means two things: (1) unlocking the huge potential of geodata assets already available and soon to be made available, and (2) providing the tools and know-how of computationally operating on that body of spatial data.

It is thus fair to state that M4S wants to catalyze and enable the science community also, but in the spatial domain, which asks for specific knowledge and experience which is available in our GI sector. For more on the possible collaboration with the NL eScience Center see question 7.6.

5.9 Are we able to work on both GI and ICT? (PvO)

Our work is GI and ICT, or if you want it is ICT of a special (and sweet) flavor.

5.10 Is Digital Rights Management the weak link of the project? (TvE)

Not necessarily, as we do not claim to solve the general DRM problem once and for all. We have a more focused user base, we have fewer products to manage, and there are a few tools available

already to handle important cases. Moreover, the tide is towards open data and the problems may to some extent disappear.

5.11 Can you operate the facility with standard (commercially available) software? (PvO)
In honesty, there is no yes/no answer to this question. The ambition level for the two-years deliverable can in all likelihood be reached by having a company like Esri, for instance, install a turn-key, large-scale SCI product, as they are installing internationally here and there. But the ambitions for later years go beyond what they offer at present, and the functionality for that level needs to be developed within our community on top of the basic GI infrastructure. It is rather painful, though not impossible, to do that on a closed, commercial platform.
The alternative is in working towards open-source solutions, which are also available and which have in recent times matured rather rapidly. A possible model for a kick-start is to have a commercial party install an open source solution to get us started. The levels of standardization in GI nowadays are such that data providers who might be operating proprietary solutions can still be included in the data provision network without too many problems.
These are just two examples of opposite ends of the solution spectrum. Hybrid solutions are possible and even likely the best alternatives.

5.12 More emphasis on geo-aspects; less on generic ICT. Many challenges appear to be purely ICT; which ones are GI-specific? (PvO)
The following aspects will be purely Geo-ICT:

- The online national atlas as a search facility for geospatial resources
- New 3D and spatiotemporal data types
- Multi- and vario-scale geo-information
- Spatial/geographic semantics for the web
- The geo-service toolbox
- Chaining geospatial services
- Metrics for geospatial data quality
- The M4S API
- The privacy issues due to binding through location

5.13 Explain very well that not a standard ICT-infrastructure is built, but a GI ICT-infrastructure. (PvO)
The M4S facility will be created as a stacked architecture of three layers, which are respectively 1. standard ICT + 2. standard GI + 3. new GI technology. We will keep an eye open for upgrading with new ICT possibilities whenever these arise and are appropriate.

5.14 What is your expertise in the area of sensor networks and the data management part of it? (PvO)
Sensor nets come in many different types and we cannot boast familiarity with all breeds of sensors and sensor nets. Nevertheless, there is an array of experience in the consortium on handling large amounts of sensor data. Just for UT/ITC, these are amongst others:

- Day to day operation on streaming data off of remote sensors on satellites, a.o. for water balance studies
- The Geonetcast facility
- The Human Sensor Web Google project
- Standardization and experimental work conducted on Sensor Web Enablement in collaboration with our German partners in 52North

5.15 How much closed / open source will be used? Is produced software within M4S open or closed source? (PvO)

The partners in the consortium are essentially neutral to the choice of closed versus open source software. This is made possible, amongst others, by the fair success of spatial data standards already reached under ISO and OGC aegis. Software components of closed source and open source nature can and do nowadays interoperate freely in many production settings. Thus, it will be possible to realize a closed source SDI product on top of which the academic partners can develop the advanced geo-ICT that will also be needed for M4S. It is to be expected that such development by academic partners will be done in open source style, to allow scrutiny, hence improvements, by peers. It is likely that software products coming out of M4S work packages will have much wider applicability, and thus be pushed to open software communities, as we have for instance with our long-time partners 52North.

At the same time, the M4S facility will not be one that will be cloned as a whole often to operate elsewhere. The community of M4S developers will thus be naturally small, and this means that the benefits of a code-sharing community will probably not be substantial. We also do expect a close collaboration with EDINA team, and their choices may affect ours. So, in principle, software developed within M4S will be in the open domain. One reason to use, especially for the foundation technology, closed sources Geo-ICT is to outsource that part of the development and allow the M4S partners to concentrate their efforts more on the scientifically challenging components.

5.16 How will performance and scalability be obtained (large data sets, many users)? Any strategy, w.r.t. to parallel processing, grid or cloud computing in system architecture? (PvO)

Though M4S is catering for the large-and-many aspect mentioned, it will not face that problem from day one, providing lead time to come up with good solutions. Another important factor is that M4S will first work towards the availability of existing data sets, which means that static performance parameters (storage, bandwidth) will be more important than dynamic performance parameters (like processing power). Scalability on the static performance parameters will therefore be given much more attention. Fortunately, with spatial data a number of straightforward schemes exists to distribute very large data sets over services, especially through compartmentalizing by spatial tiling and acquisition date. At the front of processing performance, we hope to study optimization techniques of spatial processing chains in the context of the M4S network of services and the computing facilities available to the requesting user. We intend to follow development around grid- and cloud-on-demand, as well as those around scientific databases (amongst others group of Martin Kersten, CWI). Also, the EduGIS implementation provides (Amazon) could experience within the consortium.

5.17 There exists the danger of drowning by data… computing centres in academic areas often underestimate the challenges. How is this addressed by Maps4Science? (PvO).

There are indeed serious challenges in handling the extremely large and dynamic volumes of spatial data. However, expertise is present within the consortium; e.g. the VU-team has created EduGIS facility running in Amazon cloud, and serving many different spatial data themes to a very large user community (primary and secondary education, but facility is completely open to everyone), the TUD-team is hosting the 3TU.datacenter (with a wide range of data sets; see http://datacentrum.3tu.nl/en), while the UT/ITC-team has installed SDI nodes in a number of government agencies around the world. Above all, the planned collaboration with the Netherlands eScience Center and SURF foundation is part of our strategy to provide generic and effective mass data storage, grid computing, and fast network connections.

5.18 Two parts of the proposal 'Linked Data' and 'Classic SDI' do not align very well. How will you solve this? (PvO)

Offering geo-data using the Linked Data principles must be seen as an extra layer on top of the Spatial Data Infrastructure (SDI = GII) that is the core of the Maps4Science proposal. There are a

number of initiatives to publish geographic information not only as complete datasets, but also as linkable information items; e.g. by the Ordnance Survey in the UK and the Linked GeoData project in Germany (http://www.ordnancesurvey.co.uk/oswebsite/opendata/linkeddata.html, http://linkedgeodata.org). These first attempts led to the discussion how the two can reinforce each other: the Linked Data Web needs good geo-data storage and rights management facilities, including a better way to handle complex geometries and relations between spatial objects. At the same time, SDIs can benefit by more variation in the way geo-data can be selected, navigated, explored and combined. Integration between data content and metadata for example can be simplified, and the possibility of exploring datasets and data mining by following links (and finding links) between individual items in the datasets, will be enhanced. Hence the use of semantic web and linked data technologies is meant to reinforce the SDI paradigm and set-up, and not as a replacement for current spatial data storage systems or content standards. It is in the field of making geo-information easy-to-find and easy-to-assess by 'outsiders', that semantic web and linked data innovations can contribute most to SDI's.

6. LIVING LAB

6.1 Living Lab? What is that? Also a cell is a living-lab. (AB)

The Living Lab notion is a rather new research phenomenon that introduces new ways of managing innovation processes. Our Living Lab for M4S is inspired by Brainport in Eindhoven. The underlying idea is that people's ideas, experiences, and knowledge, as well as their daily needs of support from products, services, or applications, should be the starting points in innovation.

In more detail, a Living Lab is a gathering of public–private partnerships in which businesses, researchers, authorities, and citizens work together to create, validate, and test new services, business ideas, markets, and technology in real-life contexts.

A Living Lab is not the objective, purpose or aim, but is the method to implement the process of innovation and validation.

The purpose of a Living Lab is to create a shared arena in which digital data services (M4S), processes, and new ways of working can be developed and tested with user representatives and researchers. Hence, a Living Lab is an environment in which people and technology are brought together, and in which the everyday context and user needs stimulate and challenge both research and development, since authorities and citizens take active part in the innovation process.

There are a few key principles:

- Continuity: This principle is important since good cross-border collaboration, which strengthens creativity and innovation, builds on trust, and this takes time to build up.
- Openness: The innovation process should be as open as possible, since the gathering of many perspectives and collaborative power to achieve rapid progress is important. The open process also makes it possible to support the process of user-driven innovation, including users wherever they are and whoever they are.
- Realism: To generate valid results for real markets, it is necessary to facilitate situations of use and behavior as realistically as possible. This principle of focusing on real users, in real-life situations is what distinguishes Living Labs from other kinds of open co-creation environments, such as existing GeoLabs.
- Empowerment of users: The engagement of users is fundamental to align innovation processes on a desired direction, based on the human needs and desires. Living Labs efficiency is based on the creative power of user communities; hence, it becomes important to motivate and empower the users to engage in these processes.
- Spontaneity: To succeed with new innovations, it is important to inspire usage, meet personal desires, and fit with and contribute to societal needs.

Some examples. Several companies (like Grontmij, Vicrea, Esri) need a high-quality testing environment for their software. The support letter of Esri by Jack Dangermond states: "we see many fruitful avenues for collaborative research and product outcomes that could address several of the geographic information research challenges to be tackled in your proposed large-scale research facility, including: -Usability and dissemination modes; -Management of very large data sets; -Services, searches and optimization; -Volunteered GI and citizen science; and -Satellites as a service with Esri's improving imagery deployment services." Also government agencies like RWS do no longer operate their own R&D department. With the development of PDOK and in the future – possibly – the Shared Service Organization Geo, there is the need for innovation on the geoinformation infrastructure and to work closely together with the government in this (see also the support letter of the Ministries of ELI and I&M).

To conclude. A Living Lab is a co-creation environment for innovation. It can be seen as the R&D outlet for business and government.

6.2 Remember it is an infrastructure! (AB)

Indeed M4S is an infrastructure, but an infrastructure is worthless when there is no use, or there are no guarantees for its future sustenance. To avoid building a too 'beautiful' infrastructure, we want to embed from the beginning the infrastructure in a strong user environment, with scientists, business and government (and education) in what is known as the "Golden Triangle". It is important that such development is pursued to optimize the chances of a financially sound set-up for M4S that will propel it deeply into the future.

6.3 The proposed Living Lab seems to be in its infancy; which steps are proposed to start addressing its big challenges? (AB)

The Living Lab concept is in use in other domains; see Brainport in Eindhoven (question 6.1). We also intend to learn from those domains on how they established a functional and active Living Lab. In the geo-information sector, the OGC has been conducting (international) testbeds to fast-track the development of standards (and their implementation) in a setting where industry, government and academia work closely together; see the support letter of OGC's Mark Reichardt: "Maintaining an operational linkage between OGC programs and the Maps4Science facility as part of a persistent testbed environment to understand and address the interoperability needs of the scientific community in the OGC process. I am delighted that several of the Maps4Science consortium members are also active OGC member organizations. TU DELFT as well as Geonovum, ITC / University of Twente, and the Centre for Geo-Information / Wageningen University have been OGC members for a number of years, and I invite the opportunity to help create and sustain an open line of communication and collaboration between the Map4Science project consortium and the OGC international standards process."

7. GOVERNANCE

7.1 Governance of the whole thing: governance has to be really extremely well thought out. Are the tasks, responsibilities clear? For an infrastructure proposal, it must be very clear and well thought out. It should not be democratic!! Do you have the responsibilities to stop things defined? To change things? (PD, AB)

Our governance structure is inspired by both the Clarin, and Parelsnoer initiatives and our own experience in the Bsik RGI programme for a geo-information infrastructure.

The Board is responsible for all strategic decisions, all 10 partners are involved, and there is one independent chair. The Board has the final responsibility of the project to NWO. There is a Daily Board (consisting of three facility directors, one of which is chair), leading the three subprogrammes (M4S blocks). They are responsible for realizing the facility. The National Advisory Panel (NAP) gives strategic advice; it has representatives from the Dutch geo-domain, both on the data producer and data consumer sides. Peer review takes place by the International Advisory Panel (IAP). Both NAP and IAP have a chair.

So, there are two levels of decision making: the Daily Board and the Board. The Daily Board acts within the year plan and year budget and their mandate. The Board is able to stop things or change things, for example after advice of the NAP or upon reviews by the IAP.

7.2 Is there an organisational shell around this proposal? (PvO)

See the answer above. Besides that, we are currently studying which legal entity would best fit the mission of M4S. This may lead to an agreement to establish a Foundation for M4S after granting of the proposal to guarantee good governance. See also the Letter of Intent (LoI) which is being prepared.

7.3 How can you be sure you will be able to bring the project to a favorable conclusion? (AB)

At the end, it comes down to the people involved. We have a strong believe in our team, who have their track records in for example the Bsik RGI-programme and worked very successfully together earlier. Moreover, much positive energy and early forms of commitment have been generated lately in the wider Dutch geocommunity, indicating real needs and possibilities.

7.4 Who is on the board? Who is responsible for what? (PD)

In the Board, all 10 founding institutes are involved via their high-level representatives. The future owners of the infrastructure for serving the geodata (DANS and TUD Library/ 3TU.Datacentre) play an active role in the Board. The Daily Board (facility directors) consist of a director of the Science block (Peter van Oosterom), a director of the Infrastructure block (TUDelft/DANS) and a director of the Living Lab (IIPGeo), all as part-time officers. They are responsible for realizing the facility and will among them divide the responsibilities for science, international cooperation, ICT, finance and daily management, network management and communication.

The National Advisory Panel is a representation of intended users from Science (profs), Government/Business directors (users/data suppliers) and top-level GI/ICT professionals, e.g., Kadaster, NCG, GBN, IPN**.** The International Advisory Panel is a representation of prominent, experienced researchers from international research organisations and businesses, e.g., Edina, JRC, GSDI, OGC, Oracle, Google; Members for both the NAP and IAP will be selected from the pool of support letters (and involvement in practice; e.g., because of PDOK or NMDC).

For responsibilities see the answer to question 7.1.

7.5 Setup well who has control over M4S. (PD, AB)

See the answers to the questions above (7.1 and 7.4).

7.6 What is the relationship with the eScience Center? (PvO)
We will be strengthening each other. The eScience Center is a supporter of Maps4Science. We want to make use of the eInfrastructure GOF if their proposal is granted. We would thus enable the eScience Center with geo-information facilities in the eInfrastructure.

7.7 Who will become the governors of M4S? (PvO)
The Daily Board, chaired by Peter van Oosterom, and finally the Board.

7.8 How can users influence the facility? (PvO)
There is a user group foreseen (formed by the six scientific use cases, and completed by scientific use oriented representatives from the member organizations present in our NAP and IAP), for a practical feed-back loop for the facility.
I would want to state that this is covered in the mandate of NAP. It is functional to put it there because user needs statements will be witnessed by both the M4S facilitators, and also the data providers.

7.9 How stable are the current 10 partners? Can current partners stop halfway the project (if so under what conditions)? Can new partners enter the consortium (if so, under what condition)? (AB)
The partners are very reliable in their commitment. They have proven in the past that they can cooperate (Bsik RGI, MSc Geomatics, etc.). We do not expect that they will stop. When the project is granted an internal cooperation agreement will signed that ensures the required contribution from the partners. The current 10 partners are very stable: the consortium is composed with care, all partners believe in Maps4Science and almost all of them worked together earlier in the Bsik RGI programme. Stopping halfway the project is not very likely and not possible, unless one of the partners doesn't meet his obligations. In that case they will have to leave the consortium under the condition to fulfill the financial commitments and deliverables as agreed and to compensate for damage.
About a new partner: We did not discuss this. New partners are, from the view of Maps4Science, welcome in a limited way as long as they bring significant added value and the consortium is able to handle it (in governance and finance). There is serious interest from TNO (see the support letter) and we would like to explore this further if also NWO agrees with this.

7.10 In the management structure and role, it was stated that IIP-Geo will provide the director of the Living Lab (one of the 3 main blocks of M4S). But IIP-Geo is not a partner, could you explain? (AB)
IIP-Geo is not a “standing organisation” It is a cooperation network. This network can provide the director for the Living Lab. It is the quality of the person that matters. At the moment IIP-Geo is a successful project which will end in September 2012. That was also the reason why IIP-Geo could not be a consortium partner. The Board of IIP-Geo (among them WUR, ITC, Geonovum, KNAW-NCG, SaGeo, Min I&M, GBN) has the ambition to continue the role of IIP-Geo in the future, playing the role as catalyst and facilitator for innovation in the golden triangle in the Living Lab. This will be a sustainable way to safeguard the investment of ICT Regie/ Partnership of NWO, STW and Agentschap NL in IIP-Geo.

7.11 You mention National and International Advisory Panels, could you please describe the composition of these panels more precisely? (AB).
The proposal does not explicitly mention the members, but at the moment the Advisory Panel and Board is being established. All consortium partners will be represented in the Board with an executive board with three members to ensure a decisive operation. The Advisory Panel will be mainly recruited from the supporting organizations both national organizations (currently 24 support

letters) and the international organizations (currently 12 support letters). Several organizations have already committed their collaboration in the Advisory Panel; for example see the support letter of the KNAW-NCG (Netherlands Geodetic Commission).

8. BUSINESS CASE AND PRIORITIES

8.1 Financial: we have 80 Million euros. Which part would you easily cut out or sacrifice. So which part do you cut out? Or are you only happy with the total? (AB)

We believe we have a well-balanced proposal. If we believed it would be possible to achieve the same result with a lower budget we would of course have proposed that. We want to emphasize that the participants already intend to contribute more than 25% of the funding, so we believe this is a quite reasonable proposal given the ambitions and requested subsidy.

8.2 Try to be more specific, e.g., priority nr. 1 is which are the 10 most important datasets and … (PvO)

A further prioritization is certainly scheduled, and has been hampered by the dynamics of the field, especially in the data that national and local governments are planning to move to the public domain. M4S does not intend to duplicate such efforts, but by M4S objectives, it wants to make such additional sources fully visible and thus discoverable by and exploitable to the intended scientific user community. For instance, future Rijkswaterstaat open data sources should be findable from the M4S platform, should be obtainable using transparent M4S methods, and should be as easy to use as those native to M4S.

A prioritization will identify those data sources relevant to conduct research in urban issues, logistics, water and health. The value proposition of M4s, however, is in part based on serendipity: making fortunate discoveries by accidental combinations of spatial data. This makes the proposition to define certain data sources as out-of-scope a tricky one. Similar to the prioritization of the offered geodata sets, also the offered geo-processing services will be prioritized.

8.3 BBMRI and Parelsnoer both had a clear story: local hospitals have databases that we must link. We need people, we need to standardize, we need more measurements. All man-power. (TvE)

There is some analogy with M4S, but there is also an important distinction. The distinction is that M4S caters for many scientists in many different disciplines, with the fundamental belief that spatial data carries an extremely powerful connecting factor in the registered locality. This allows one to compare distance to the nearest hamburger outlet with incidence of obesity at primary schools. The number of data combinations is just endless. But the location code is not always a simple coordinate pair, because spatial data comes in many shapes. Consequently, its interoperation is non-trivial and not quite for the non-initiated. Support and training in its usage for the large potential user population is therefore a sine qua non.

The analogy holds because The Netherlands has a substantial number of data producers and disseminators that have partially overlapping data holdings, giving intrinsic interoperation problems, and requiring standards to be in place. Our own libraries are one such source, but there are many more in the public and private domains as well. The setting up of standards, standardized services and novel modes of finding and binding the data sources requires substantial manpower, as well as new research in GI. That new research, a.o. is in the construction of semantically-informed processing services on spatial data sources.

8.4 The accents in the budget should be better emphasized. (PvO)

The budget has four main tastes, and we have used the time after initial submission (31 August 2011) to detail these out to a more detailed level.

A necessary first component is that of *hardware/software* for the facility, allowing us to host data sets at acceptable service levels, construct and maintain portal and services, and work towards a deeper standardization that will allow future free-form data cross-combinations. In this component, another important part is labor. The second component in the budget is devoted to *GI Science*

development in support of M4S, addressing the appropriate handling of new data types, new services, deeper (spatial) semantics and spatial processing appropriate for purpose, as well as models and services to support handling the rights on spatial data sources. By far the biggest part of this component is essentially labor, which is translated into production software necessary to turn M4S into a platform for the future. A third component is a party of *six scientific use cases* that will demonstrate the M4S potential, and will serve M4S to develop robust dissemination methods. The use cases have been chosen as truly diverse, underpinning M4S's ambition to serve all of science equally well. The scientific use cases are a mix of investment in equipment and labor. Whereas the use cases provide us with early proofs of concept, the fourth component of the *Living Lab* is meant to work towards long-term sustainability of M4S, addressing models of collaboration with public, private and citizen actors, aiming to find return on investment for M4S also.

8.5 The acquisition of data is not well described and covered. (PD)
This is a just criticism, and was the consequence of not explicitly describing the role and commitment of data producers in our project proposal. However, this has received sufficient attention since date of submission, and we are happy to be able to state that we expect to secure data agreements for M4S exploitation with a large number of parties already in the early stages of building the facility (also see the support letters of many data producers). M4S does in principle not plan to pay for data usage licenses, under the philosophy that novel modes of use will be discovered by the scientific end-user community, and which can later be leveraged potentially by the data owners in their commercial exploitation. The M4S consortium will seek to reach agreements with data providers in a style that allows M4S over time to tax for marginal data use by scientists, aiming to recover the costs of storage and provisioning in a non-commercial way.

8.6 Over 20 million is a large investment. What will be realized and what is the leverage? (PvO)
The budget for the project is certainly substantial, and by ratio so is the commitment of the parties involved. Our respective directors, deans and rectors believe this is a wise investment that will propel a number of scientific fields into novel and intensive forms of spatial data usage. The data-intensive and well-informed science is the science of the future, and with this comes an intrinsic need of understanding the contexts in which phenomena under study exist and take place. Spatial data and monitored spatial processes can fundamentally help paint in that contextual understanding. That is the expected leverage and fundamental value proposition of M4S. Next to that, the GI technology needs to grow with it and provide the exploratory and exploitation methods to make it all work. By the end of the subsidized M4S facility project, the Dutch science community has a platform that brings the extremely rich geodata ecosystem at the fingertips, and brings with it expert knowledge of how to practically operate on and with those data sources.

8.7 Budget is flattering. When going below a certain level (10 M€), then the proposal is not appropriate for the NWO call. (PvO)
The budget is needed to reach the levels of ambition that we have sketched. Perhaps we will need to downplay some ambitions because of a lower budget (see question 8.1), but more than a 50% cut will essentially turn M4S into a shallow undertaking that is not worth pursuing.

8.8 NWO only funds work with a strong connection with the top sectors; how strong is the M4S case in this respect? (AB)
This case is actually very strong, as spatial data is everywhere, and timely spatial data is becoming a much more common commodity these days. This true of the booming sensor field, the booming mobile field, and also the booming social media field: spatial is everywhere. In the overall ICT Roadmap of the top sectors, Geo-ICT is pointed out as an important development and there is a close connection with the top sectors Water (business case Digitale Delta), AgroFood, Energy (smart grids), High Tech Systems and Materials (use of satellite images, 'Convenant Informatieketen Water & Klimaat'), Creative Industry, and so on. Due to our well organised business partners in GBN, this is

not only wishful thinking. Via the Living Lab, co-creation for innovation is foreseen.

8.9 The duality is the fact that you're working to realize a facility and you also offer a service. (PD)
We want to realize a facility; that facility's function is to offer services. The M4S goal is to realize the facility. The proposal potentially could support at least 5 out of 9 top sectors, i.e., Logistics, Agriculture, Water, Food and Life Sciences. This proposal contains business cases that are examples of these top sectors. We intend to create a facility that will serve as a platform for future collaboration, providing access to existing data-sets and allowing for co-creation and sharing of new ones, also various services required for using and combining data from various sources will be provided. This will enable researchers to boost their research (re)using each other's results while respecting regulatory requirements. The proposed GEO-ontological work enabling the interpretation and consequent combination of geodata sets and the DRM of a distributed, i.e., networked solution is unique in the world and will enable us to produce ground breaking research results.

8.10 How is the money spent? Infrastructure, people?
8.14 How is the money spent? (R&D, standards, semantics, production facility)
8.15 Is it possible to split the money between the various components (hardware, software, development)? (PvO)
See above, but requires further detail. Section 5.1 of the proposal (page 23) provides a little more detail. A very rough, initial estimation would be: staff 70%, hardware 20% and software 10%, but see tha chart bin our presentation, for which we have (obviously) backup data. The staff effort can be further refined in various types of labor, of which the R&D efforts are dominant. No M4S money is spent on data (as this will be either from open data policy or provided by supporters and partners). Also, the cost for software licenses is relatively low due to favorable agreements with supporting software vendors.

8.11 If you have to give priority for the various components of the project, which part is absolutely necessary? (geen 3D, geen temporal) (PvO)
See answers to 8.1 and 8.2.

8.12 Can you leave out one block (facility, research, living labs) completely? How does this influence the desired effects? (PvO)
No, not without significantly disturbing the balance, because it is our philosophy to build a well-balanced demand-supply infrastructure, so we cannot leave out one block.
Let us play the game along a little bit here and sketch what results if we leave out one of the three.
If we leave out the facility itself, we might be able to construct a wholly virtual environment based on services provided by the current data holders only. In the strict sense, this will not allow us to host any of the new geo-data, and we will not construct any novel dissemination services either, because those will not be hosted easily.
Concerning the Science block (having 4 parts: GI Science, DRM, Cyberinfra, and Science with GI). If we leave out the GI research and Cyberinfra parts completely, we will be able to obtain an SDI 1.0 but not an SDI 2.0, and many of the exciting novel forms of use will be out of the question. There will also not be any support for the new geodata types. We might omit the Science with GI part 'scientific use cases' (and hope that this will continue somehow, but less controlled from M4S perspectives: requirements and fair assessment). Assuming that the open data policy will be dominant, we might even omit the DRM-part, but for the time being this would exclude a significant amount of data sets (currently not open).
If we leave out the Living Lab, then we will be able to construct an ambitious M4S facility as described, however, we will have made no effort towards a long-term continuity of the facility as a whole and there may be little lasting impact of M4S seen as a project. As with the 'scientific use cases' we might exclude the living lab from M4S, and hope that it will 'happen' anyways.

8.13 Can you make money with this facility? (PD)
While we do not intend to make money ourselves, the infrastructure that will remain after this project will certainly allow commercial organizations to develop business out of the available geodata sets and services. On the other hand, we will try to negotiate with the data holders the possibility to charge marginally, in cost recovery mode, for data services to scientific end-users. After all, these users nowadays make heavy investments in their research data acquisition, some of which we will ease up dramatically. It appears only reasonable that a small fee is paid for every 1Mb of data delivered, and for every compute service executed or delivered. Clearly, such an RTI scheme should leave intact any potential for commercial activity by the data holders. The latter are keen to see us develop a scientific playground that identifies for them those new delivery modes.

8.16 Are you willing to sacrifice parts of the project? What will remain if you receive only 25%, 50%, 75% of the support? (PvO)
In summary, 25% is probably doable but is asking for serious sacrifices, and will force us to make substantial focusing steps (see question 8.1, 8.2 and 8.12). A 50% budget cut will turn M4S into a proof-of-concept platform, not a production platform which we now aim at. A 75% cut is not thinkable, would take out the backbone of the project philosophy, and would disqualify us for the current NWO GOF round for which we are here today.

8.17 Does the project meet the NWO financing criteria? (PD)
Yes, we carefully analyzed the call text, we further discussed the proposal with NWO representatives and we have a consortium partner, DANS, having earlier experience with such a large-scale research facility.

8.18 Can EU funds be obtained for this proposal? (TvE, PD)
In its present setting, the proposal is not fit for meeting EU criteria directly, as it is strongly nationally oriented. (The reason being the unparalleled Dutch rich data ecosystem.) Successful application in EU frameworks is possible probably when a more international consortium is constructed, and one that is willing to work on the Dutch case. The selling points of M4S remain strong scientifically, but perhaps not EU-politically. The construction of such a consortium is quite well possible but has not been attempted, though we have received a significant amount of support letters from the international domain, including from partners that could form an European consortium. We do expect a successful M4S facility to create quite an impact internationally. It is our intention to further explore the EU possibilities in collaboration with other stakeholders, such as EDINA from the UK and the JRC SDI Unit (from the EC, responsible for the technical aspects of the INSPIRE directive). For example, the call as described in the 'Draft ICT PSP work programme 2012' offers good opportunities for an EU counterpart (of part) of M4S.

8.19 Any infrastructure has continuous depreciations and therefore requires continuous investments. How are these secured? (PD)
We are in the early stages of describing our plans to this. The fundamental parameters for the long-term sustainability of M4S are the following:

- There are facilities already in place to deliver geodata to scientists; the funding for that is expected not to disappear.
- Scientific data acquisition is a costly affair today, and so, scientists in the future will be willing to pay a marginal, cost recovery rate for spatial data usage, giving them financial benefit because they are spending far less time on it.
- The data dissemination cycles will become so short that data-intensive, iterative acquire-analyse-conclude cycles will become possible, leading to different and exciting new research modes.

- Current data holders will be willing to grant M4S the right to charge scientific end-users for cost-recovery tariffs only, in the hope of seeing new models of data usage arise, which they will have the right to absorb in their own business models.

9.20 Business case is not clear, could you elaborate? (PvO)

In the proposal, the top level of the business plan was presented. We use this question to elaborate on this aspect. Maps4Science will be mainly a virtual or networked facility. The principle of "Think Global – Support Local" applies, meaning that the facility will connect many sites through open standards and protocols, while the (current) support staff will be based locally at the different universities and research organisations. In addition, a centralized Maps4Science development and coordinating office is envisaged. The Business Case for Maps4Science consists of two elements.

1. Maps4Science is a tool for science: a national facility connecting and extending local facilities of spatial data provision (and leading international developments). Maps4Science makes local (single university) facilities more effective and efficient. As a whole, it will operate as a base facility for spatial data and processing services of interest to a large group of science domains. The value proposition to end-user scientists in these domains is the free and direct access to high-quality data, the ease of data discovery and the ease with which data will be embedded and made operational in their own systems.
2. The second element in the business case is the facility itself as topic for research: How can we manage these large sets of spatial data and how can we deal with governance and technical challenges. How can we create a facility of unprecedented coverage, quality and ease-of-use globally, that goes far beyond the current state-of-the art of a 'classic' Geo-Information Infrastructure (GII): geo semantic-web support, good and complete 3D/4D/5D support, ability to handle extremely large amounts of geo-data, supporting geo-processing services within GII (analysis, simulation, data mining), delivering back geo-data (and parts of processing chains) through to the GII by the 'users', Volunteered Geographic Information (VGI), real-time streaming of data from live sensors, etc.

In the Set-up Phase there are the 'construction and investment costs' for the spatial cyberinfrastructure, consisting of 'traditional' datasets of an estimated volume growth of 200 TB/year and 'non-traditional' datasets (such as sensor network data, VGI, LBS and other sources) of between 500 TB and 1000 TB/year resulting in a total storage (and processing) capability of 7 PB. Some important connections must be based on optical pathways to ensure fast and seamless access. The budget corresponds to current realistic costs.

The running costs for the 5-year Operational Phase follow the same distinction: (i) Are the expenses for maintenance, evaluation and improvements, personnel, general operations and demonstration science cases; and (ii) Are the expenses for the Living Lab and the science case GII. Personnel costs indicated are intended for a small coordinating staff and additional support. The partners contribute their own budgets now already allocated to staff and cyberinfrastructure at their premises. As part of the national facility these budgets will get better leverage. And new possibilities for science become available.

The planned collaboration with the Netherlands eScience Center and SURF foundation (see support letter) will deliver e-Science expertise and proper embedding in the Netherlands ICT and e-Science application landscape. These collaborations also bring solutions to common e-Science problems and will allow for the re-use of services and knowledge. In addition, the Netherlands boasts a good generic grid computing infrastructure (BiGGrid), fast network connections and substantial available data storage (see also "Building the e-Infrastructure" proposal by SURF for NWO Large-Scale Research Facilities). Dutch industry, academic research groups and libraries at the Dutch Universities are internationally at the forefront of the developments in the field of geo-data and have a great deal of expertise in handling, analysing and annotating this data. Furthermore, there is expertise in the distributed computing tasks on computer grids and clouds, and there is a good local infrastructure for data storage and data handling. Nevertheless, we will also require a dedicated national facility with fast network access to support this national infrastructure project, and the rapidly growing

needs for storage and processing of geo-data.
The developed knowledge will also be extremely useful to governing organizations of the GII, as can be read in the recently received support letters from the Ministry of Infrastructure and the Environment (I&M) and the Ministry of Economic Affairs, Agriculture and Innovation (EL&I). A quote putting Maps4Science is the right perspective: "*we [I&M, ELI] are working on the implementation of a National Infrastructure for Geo-information. This infrastructure includes a system of base-registrations, a national distribution and access point for geo-information (PDOK, Public services on Maps, National GeoRegister) and a Shared Service Organization Geo-information (SSO-Geo). … In the future, we will need a complementary infrastructure which is able of managing diffuse, high volume data and make it available for users in a practical manner. This is a next step in the evolution of geo-information infrastructures. Science has to make advancements to bring this to reality. We believe that Maps4Science can fulfil this need." and "We [I&M] also see possibilities for commercial spin-off which will contribute to a more knowledge-based economy. Application of geo-information within the 'topsectoren' spring to mind."*

9. FUTURE

9.1 It should be kept alive. So, if you can say that in addition to scientists, also industry and government can use the facility (as spin-off). (PD)
Primary users are the scientists. However, it is imaginable that there would also be users from outside that benefit from the facility (and would be willing to pay for this, depending on the data covenants we reach with data holders). Models for sustainable funding of the facility (after the project) will be investigated during the project (and will be one of the deliverables).

9.2 Make sure that after two years at least three companies have signed an extensive MoU. (PvO)
There are a large number of possible collaborations; see the current list of support letters. It is totally realistic to assume that quite a number of these will be translated into real contracts for collaboration (i.e., more than MoU's). The type of contract depends on the organization; e.g., could be collaboration with mutual benefits with NMDC, PDOK, or data agreements with data owners, or providing base facilities to commercial service developers, etc.

9.3 It says: decommissioning ... but running the facility will costs 2,5 or 3 million per year. Nobody wants to own the infrastructure – but all want services. How do you deal with that? (PD)
The use of this terminology may be unfortunate in our proposal. What we intended to state is that it is the decommissioning of the project, but not the facility itself. At the moment, the team of DANS and TUD Libarary/3TU.Datacenter have expressed their willingness to become the owners of the facility (after the project). Their motivation is that that they perceive the facility's great value to scientists in the future (as it has been in the past). Of course, such can only be achieved under reasonable conditions (and as this is 7 years ahead, there is sufficient time to prepare this well).

9.4 If you do not have agreements concerning the conditions for using the facility, then no one wants to own the facility. (PvO)
It is clear that a fair and healthy funding model of the facility (infrastructure) is needed. This is non-trivial for M4S, given our large and heterogeneous data sets, our range and intensity of use. These requirements direct it towards a flexible and scalable solution (which must somehow be reflected in organizational and financial aspects involved). Currently, a LoI is being prepared in which the future custodians of the facility (DANS and TUD Library/3TU.Datacenter) and the other project partners express their intention to continue with the facility (after the project) based on a to-be-developed model for sustainable operation (see also question 9.7). This model is also one of the deliverables of the project and obviously needs investigation. Also, part of the proof of the needs will be delivered by the (growing) use and needs during the project lifetime (and give better foundation to future decisions and risks for the future afterwards).

9.5 How is your international embedding: can we take the lead in Europe? How? (PvO)
The Netherlands geo-information sector has a very good reputation: our scientists are in the top-5 (publications) and top-3 (impact) of the world (see also question 10.1), our coordinating government body Geonovum is very influential in the European INSPIRE implementation (and NL is often seen as an example and provides many 'active builders'), our geo-information industry is equally strong and is well-organized (see 100 members of GBN). With the additional push of M4S we are in an even better position to further extend our leading position and generate benefits from this; in the first place for our scientists, but the benefits will be much wider (government, industry, citizens).

9.6 What is the meaning of the word decommissioning? Who will manage/operate the facility after the project? (PvO)
See question 9.3 above.

9.7 Can you ask end-users (universities) to pay for the facility after the project? (PD)
Yes, they have been paying in the past, they are paying now, they have committed 25% own contribution for the next 7 years, and it is very likely that with increased used by their scientists, they are willing to provide funding in the future (as the impact and importance of geodata is only growing), probably on a marginal, per-use basis. A national facility will offer at lower cost per user more functionality, that is, more for less.

9.8 Will you be able to embed the facility internationally? (PvO)
International discussions are on-going and will continue. Similar to the relationship between PDOK (base registration government) and Maps4Science, there is a healthy future for a European scale-up and have the parallel INPIRE (government base geo-data) – EU-Maps4Science. This could evolve bottom-up by collaborations with leading countries; e.g., NL-UK or top-down by EU decision (and developed in context of Horizon2020, FP8, FP9, ….). Many partners and umbrella organizations are already involved; see support letters: Eurogeographics, EuroSDR, AGILE.

9.9 To what extent do we cooperate with other countries? (PvO)
See question 9.8 above.

9.10 In LoI it is stated that DANS/3TUDatacentre intend to take responsibility for facility after project, how are tasks divided between DANS en 3DDatacentre (after project, but also during project). (PD)
DANS and 3TU.Datacenter are currently discussing a closer cooperation, the foreseen model for which is that DANS will perform back-office functions for the technical sciences, while 3TU.DC performs front-office functions for its users. The division of tasks for M4S will be part of the cooperation. During the project the tasks are defined by the work packages.

9.11 Same question but now between DANS/3TUDatacenter and other partners. (PD)
DANS and 3TU.DC will be primarily responsible for providing continued access to geodata on a basic level. The maintenance of the services on the data, for which specific applications will be developed during the M4S project, is a shared responsibility of all partners, for which an adequate business model will be selected (see other questions on this issue).

9.12 DANS is involved in multiple GOF proposals, e.g. also in Clariah. However, the NWO budget is limited and only one of these could be funded, which one would DANS select? and why? (PD)
It is DANS' explicit strategy to serve all disciplines where a demand exists in the area of data archiving and access provision. DANS prefers to play a limited supportive role in many research infrastrutures/facilities rather than being central in one. In the Roadmap 2008 humanities and social science were two separate fields, and we expect that this remains so in 2012. As a partner in M4S in the social sciences and in CLARIAH in the humanities, we hope the Roadmap budget will suffice to finance at least these two. DANS is also partner in ESS and LifeWatch, where ESS was already funded in 2008 and LifeWatch is in an other domain (Life sciences). Of course, both M4S and CLARIAH are very dear to me.

10. WHO

10.1 How would you describe the current position of the Dutch GI-research in Europe/World? (AB)
The Dutch GI research community holds a global top 5 position. WUR citation analyses in 2009 showed a strong increase in publications by that Dutch community over the course of the national RGI program. The Dutch GI community holds position 5 globally in publication numbers, and position 3 (after Germany and the USA) in citation impact (see for the whole report Eindrapport RGI-402 at http://www.rgi.nl/index.php?sid=10&nid=358: Effect van "Ruimte voor Geo-informatie" op GIScience in Nederland).
The position is due to the long history of the Netherlands in this field (for example cartography in times of the VOC) and because of the geographic situation of the Netherlands: we are a small country with a high density of population, partly below the sea-level, so we have to make smart use of space. The Netherlands is also the best measured country in the world, no one country has detailed maps like GBKN (BGT) and AHN-2.
Oracle recognized the TU Delft contributions to the geospatial community by awarding Peter van Oosterom the 2007 Oracle Spatial Excellence Award for Research and Development (see support letter Oracle). The TU Delft team had also won both Innovation awards in the Science category of the Bsik innovation program Space for Geo-Information (RGI)
The support letter of Esri (worldwide leader in GIS software) by Jack Dangermond (owner and president) states "We have been exposed to the excellent scientific production of your group and other groups in the proposal consortium, via our common membership in associations such as the Association of Geographic Information Laboratories in Europe (AGILE) and the Open Geospatial Consortium (OGC) … Funding the Maps4Science infrastructure will again demonstrate *Dutch leadership in the global field of geoinformation*. As leaders in the geographic information systems market during more than 40 years, and with many users in The Netherlands, Esri can confidently estimate the huge impact that such an infrastructure would have on scientific innovation, which is premised on access to information, of a wide range of scientific disciplines. To that end we stand ready to provide technical and capacity-building assistance to the project, when funded, to help make it successful."

10.2 We are not the Top5, we belong to the Top5 of the world. (PvO)
Indeed, this should be correctly formulated.

10.3 Committee members also have questions they do not ask: -Are YOU the top 5? (image) - Amazing technological things are promised (3D) and is this then a PPT that is convincing? - Are you talking for your own benefit or for the whole Netherlands? -Is this the consortium that is able to realize this facility and within 5 years belonging to the world top?
It's a vicious competition. (PvO)
The interview team has been modified to now include our top professors – the official interview takes place with the crème-de-la-créme of the geo-information science on this subject! In addition, four renowned professors in video interviews explain themselves 4 (out of the 6) scientific use cases. Other quality indicators are the very good scores of the involved groups involved University groups in the last research assessment ('onderzoeksvisitaties'; e.g., the TUD GIStechnology group scored, above an average of 4), has received various scientific awards (see also question 10.1), and the qualitative statements that can be found in various support letters; e.g., or Oracle and Esri (illustrating the high impact of the past GI research). The presentation with be improved, more advanced graphics and images fitting better to the high technological and scientific nature of M4S.

10.4 From the proposal it is not clear that there exists a proven willingness to cooperate, is this aspect lacking in the consortium? (PvO)

We are a bit surprised by this question as the geo-information sector in the Netherlands is well-recognized for its collaboration between private sector, government bodies, and scientific and research organizations. This collaboration is unprecedented and serves as an example for other sectors not only in the Netherlands, but also in other countries. This is illustrated amongst others by a well organized geo-business community in GeoBusiness Nederland (supporter, see their letter); an organized public sector through the existence of Geonovum (partner); four Universities (all partners) jointly deliver the MSc GIMA; and our academic libraries (nearly all partners, incl. DANS) have a joint platform 'UKB werkgroep kaarten en GIS'. The recently started programme ICT Innovation Platform Geo-information (IIPGeo, supporter) connects the well established and organized Dutch geo-community with other (ICT) sectors.

More proof of collaboration in the Dutch geo-community is the Bsik RGI (http://www.rgi.nl/index.php?l=eng), which gave an impulse to geo-innovation in the Netherlands with many results with great impact. This success is based on a established network involving over 300 (inter)national organizations. Examples are: (i) the EduGIS facility, (ii) the RGI 3D geo-information developments, resulting in the NL 3D pilot (conducted by a huge number of partners in networked setting and receiving last month an international award from the OGC); see http://www.geonovum.nl/dossiers/3d-pilot; and the report of the International Midterm Review committee RGI illustrates the Dutch willingness to collaborate by stating "A strong and vibrant network with a good critical mass has been built up. Care must be taken to ensure the viability of networking programmes over more years (e.g. 7-10 years)"; see: http://www.rgi.nl/downloads/files/MTR%20RGI%20FINAL%20VERSION.pdf

In summary, not only does a real team spirit exist among the Maps4Science partners, but also among the (inter)national supporters. Note that additional support letters continue to be received, showing the growing (inter)national support; see http://www.maps4science.nl/.

10.5 In the proposal you claim to have strong relations to important industrial organizations such as ESRI and Oracle. However, this is not indicated via support letters, so it's your claim valid? (PvO)

The relationship with leading industry is based on various long-lasting partnerships. The main applicant has a strong partnership (via the GDMC http://www.gdmc.nl/participation) with Oracle. The spatial DBMS innovations fuelled by the TUD research, such as topology structures and 3D functionality, are nowadays included in the flagship product of Oracle: their DBMS. The warm relationship between Esri Inc. and UT/ITC dates back to the early 1990s, and continues to be active today, for instance through joint execution of international consulting and implementation projects, especially in SDI and related capacity development. We have requested support letters form our industry partners and received them form Oracle and ESRI (while Google confirmed willingness to provide such a support letter); see website maps4science.nl