SDI Convergence

Research, Emerging Trends, and Critical Assessment

SDI Convergence

Research, Emerging Trends, and Critical Assessment

B. van Loenen, J.W.J. Besemer, J.A. Zevenbergen (Editors)

Table of Contents

Foreword

This book is the result of a collaborative initiative of the Global Spatial Data Infrastructure Association (GSDI), the European Commission's Joint Research Centre, the European umbrella organisation for geographic information (EUROGI), the Dutch innovation program Space for Geo-Information (RGI) and Geonovum, the Netherlands SDI executive committee. In addition to the traditional call for papers for the joint GSDI-11/ INSPIRE/ Geonovum RGI conference "Spatial Data Infrastructure Convergence: Building SDI Bridges to Address Global Challenges" contributions of full articles were solicited for publication in this peer reviewed volume.

In several instances, the articles submitted addressed the theme of the conference. In others they stuck to the more traditional fields of SDI. The reviewing process resulted in twenty articles that together can be summarised as Spatial Data Infrastructure Convergence: Research, Emerging Trends, and Critical Assessment. These topics are represented well in this volume. We thank the authors of the articles and the members of the Peer Review Board.

We are grateful to the Netherlands Geodetic Commission (NCG) for their willingness to publish this work under a Creative Commons Attribution 3.0 license. It allows all to use the experiences and research presented in this book to their advantage. We especially thank Frans Schröder and Leni Verhoog of the Netherlands Geodetic Commission (NCG) for their assistance during the entire publishing process.

Rob van de Velde, Chairman
Local Organising Committee for GSDI-11
Director, Geonovum, the Netherlands

Bas C. Kok, President
GSDI Association (2008-2009)

Alessandro Annoni,
European Commission, Joint Research Centre

Mauro Salvemini, President; João Geirinhas, Secretary General
EUROGI

Jacqueline Meerkerk, Director
Space for Geo-Information (RGI), the Netherlands

Peer Review Board for SDI Convergence: Research, Emerging Trends, and Critical Assessment

Spatial Data Infrastructure Convergence

Bastiaan van Loenen, Jaap Besemer and Jaap Zevenbergen

Delft University of Technology, the Netherlands
{b.vanloenen, j.w.j.besemer, j.a.zevenbergen}@tudelft.nl

Spatial data infrastructures (SDIs), and their underlying geographic information technologies*,* play a critical role in allowing governments, local communities, non-government organisations, the commercial sector, the academic community and common people to make progress in addressing many of the worlds most pressing problems. The approaches in building spatial data infrastructures within and among nations are in many respects *converging*. Not only SDIs are converging, also many developments in society stimulate convergence of SDIs.

Geographic information (GI) is special because it refers in one way or another to a location relative to the earth. Other information can be linked to GI, for example, health care information, telecommunications, financial information, and traffic information. This specialty has not changed. However, the landscape of geo-information processing has. Where previously only public parties collected, provided and used geo-information, today this is done by almost everybody, varying from professional commercial data and service providers to traditional providers in the public sector, and (volunteered) user generated content providers. The last are often regular citizens that upload pictures on Google Earth, update the route in their navigation software or the public transportation route planner. This has resulted in geographic information rich societies in many places around the globe where accessibility, coverage, accuracy, and timeliness has never been as high.

The sector has seen a development from private editors of printed atlases meeting the needs of the 1950s customer to the 2009 atlases and APIs which can be accessed anywhere at any time. Similar to the world outside GI (see the replacement of *Encyclopedia Britannica* by Wikipedia forcing Brittanica to move towards Britannica 2.0), the major publishers of printed atlases are replaced by companies originating from outside the traditional mapping sector (see Microsoft, Nokia, Google) satisfying the GI needs of customers, but this time digital and online, and sometimes even real-time. Traditional users of GI such as engineering companies are still there, but the end-user role is now dominated by ICT-companies.

Not only GI-software and data providers are converging (e.g., TomTom and TeleAtlas) also providers of services outside the geo-domain are now converging with geo-partners (e.g., Nokia and NavTeq). We have moved from traditional geographic data acquisition by surveyors, aerial imagery and professional bureaus using GPS, to location information collection through non-geo techniques such as mobile telephone networks, wireless telecommunication networks, and radio frequency identification networks. This resulted in convergence of geographic information with other domains, all contributing to the geographic information component of the information infrastructure.

Also globalisation makes that SDIs do not stop at the national border. Increasingly national SDIs adhere to or consider the international arena. New domains are now fuelled with geographic information. Well-known examples are geographic information use in organisations like the United Nations and the Worldbank, but also the Intergovernmental Panel on Climate Change (IPCC), Greenpeace, WWF, International Federation of Red Cross and Red Crescent Societies, and parties operating in the national security

domain are or may become important players in SDI communities. This may result in initiatives that combine global challenges such as examining the relation between climate change and national security (see CAN, 2007).

"When the walls are down, the bridges can be build"
(Jesse Jackson, April 11, 2009, Nova College Tour).

It is probably fair to conclude that the internal orientation of the SDI community is moving towards a more external one. The walls of the SDIs are coming down which paves the way for building bridges. In this respect, the conference theme *Spatial Data Infrastructure Convergence: Building SDI Bridges to Address Global Challenges* is very timely. In this book several of these bridges are explored, described and examined, although the focus in many articles is still on making SDI more ready to go external. The articles presented in this book have gone through a full peer review process where each article was reviewed by three members of the peer review board. It appears that this volume covers three foci areas in the SDI research community: (1) Issues evolving around service-oriented SDIs, including catalogues, funding, legal aspects, modelling and metadata, (2) Critical assessment of SDIs, and (3) Community bridging.

TOWARDS SERVICE-ORIENTED SDIS

The focus of SDIs have moved from a data orientation in the 1990s to a process orientation in the late 1990s-2005 towards service-oriented SDIs exemplified by the INSPIRE directive in Europe and the Spatially enabling government initiative in Australia. Several of the challenges that service-oriented SDIs are facing are discussed in this volume.

A prime service needed to fully exploit the notion of SDI deals with the comprehensive access to the underlying data. A nice service to do so is presented in *The Potential of a National Atlas as Integral Part of the Spatial Data infrastructure Exemplified by the New Dutch National Atlas* by Menno-Jan Kraak, Ferjan Ormeling, Barend Köbben and Trias Aditya. National atlases contain comprehensive combinations of spatial datasets represented by maps that each completely covers a country, with an added narrative function. National atlases would benefit from an up-to-date data flow, and the SDI would benefit from integrated visual summaries of available spatial data and geo-services in well-designed comparable maps using the narrative characteristics of the atlas. As such the national atlas provides alternative interactive and dynamic access to the SDI. It may very well be that the Atlas interface and functionalities are the standard for the next generation Geoportals.

Not only access to data is important, but also access to services that allow the data to really contribute to the user's needs. In *Development and Deployment of a Services Catalog in Compliance with the INSPIRE Metadata Implementing Rules,* Javier Nogueras-Iso, Jesús Barrera, Antonio Frederico Rodríguez, Rocío Recio, Christian Laborda and Francisco Javier Zarazaga-Soria argue that services catalogues will constitute a key element for SDI development. They are essential to facilitate the reusability of services. Therefore, it is necessary to move towards more interoperable descriptions of services. The authors present the development and deployment of a first prototype of a services catalogue in the SDI of Spain. This new prototype allows for providing online services such as metadata creation, dynamic indexing and metadata search, as well as the connection with generic clients, which allow verifying the correct operation of services.

SDIs are not for free, and thus need to be funded. Frederika Welle Donker discusses in *Public Sector Geo Web Services: Which Business Model Will Pay for a Free Lunch?* what business models and financial models may provide a solid financial base for public sector geographic web services. Web services are an effective way to make public sector geo-information available. They allow information to be accessed directly at the source and to be combined from different sources. Such a situation is envisaged by INSPIRE. However, the costs of web services are high and revenues do not always cover the costs. Her assessment concludes that the current business models and financial models, especially the cost recovery models, have to be reconsidered to provide sustainable funding for quality services.

An issue closely related to the funding model is in enforcing and managing use rights for these services. In *Standard Licences for Geographic Information: the Development and Implementation in Local Government in Italy* Luigi Garretti, Silvana Griffa, Roberta Lucà and MariaTeresa Lopreiato introduce a model that is promising to satisfy some of the concerns of service providers. Web services and digital media make it easy to access information, including geographic information. Geographic information moving across digital networks may limit the power of content providers to control the use of their 'intellectual property'. Piedmont Region, Italy, is developing a new business model for licensing geographic information. It aims to realise a complete policy guideline to regulate the use and dissemination of Piedmont's geographic information through a set of standard electronic licences. Based on digital rights management principles, a new electronic licensing model was developed fully adhering to the INSPIRE principles. The model promotes sharing and re-use of geographic information through services.

In *Legal Simcity; Legislative Maps and Semantic Web Supporting Conflict Resolution,* Rob Peters, Rinke Hoekstra, Tom van Engers and Erik Hupkes present an atlas service, this time focused on legislation applying to a location. Their Legal Atlas should stimulate participatory decision-making by providing answers to questions like "where will I be able to develop wasteland?" Four levels of legislation (European, national, regional and local) are assessed by Legal Atlas. Through a Simcity approach users will be able to go through the map and see per location the possibilities. The application has been successfully tested by end-users in the Dutch province of Flevoland. Legal Atlas also overcomes the difficulty of visualising temporal factors in environmental legislation, like INSPIRE, such as breeding seasons.

New technological opportunities becoming reality through new services may also challenge legislative frameworks established in periods in which these services were non-existent. Location based services, for example, now raise these new questions to privacy legislation. In *Power and Privacy: the Use of LBS in Dutch Public Administration*, Charlotte van Ooijen and Sjaak Nouwt assess how the use of location-based information about citizens in public administration, such as road pricing systems may affect the meaning of citizenship by shifting the information and the power relationship between government and citizens. The authors explored three cases of LBS in Dutch public administration and argue that LBS may affect the balance between the roles citizens can have in their three relationships with government: (1) as subject of the state, (2) as client of the state and (3) as citoyen (partner of the state). LBSs increase governments' knowledge of whom is where at what moment, and the underlying technologies may also influence the level of control of whom goes where at what moment. The authors conclude that the right to privacy exists in public places where citizens can be monitored and information about them collected. When interacting with citizens, government should be aware of the conflicting values of the subject, citoyen and client role to avoid extremes such as Big Brother scenarios.

Services taking advantage of multiple data sets depend on the interoperability of the information models underlying these data sets. Jantien Stoter, Wilko Quak and Arjen Hofman explore in *Harmonising and Integrating Two Domain Models Topography* the feasibility of harmonising and integrating two independently established information models topography, expressed in UML (Unified Modelling Language) class diagrams. They used two datasets representing topography at different scales for different purposes in the Netherlands. For both datasets information models have been established that describe the content and meaning of the data. Since the information models were developed for different application domains they often do not align. The authors propose an information model topography that integrates the two information models. In this way they attempted to bridge the differences of two information models and arrive at fluent data integration of similar data types available at different scales.

In *An Analysis of Technology Choices for Data Grids in a Spatial Data Infrastructure,* Serena Coetzee and Judith Bishop address grid-enabling SDIs. A grid is a system integrating, virtualising, and managing services and resources in a distributed, heterogeneous environment. It supports virtual organisations across administrative and organisational domains. A data grid is a special kind of grid in which data resources are shared and coordinated. The authors present a scenario that describes how data grids can be applied to enable the sharing of address data in an SDI. They developed *Compartimos*, a reference model for an address data grid, and identified the essential components for sharing address data on a data grid in an SDI environment. So far, data grids have been applied to traditional data (text, image, sound). The authors now created a bridge in the sense of applying relevant developments in non-spatial data to spatial data. It is a promising way to manage the service-oriented SDI technically.

The importance of metadata documentation as a prerequisite for SDI is raised in *SDI and Metadata Entry and Updating Tools* by Abbas Rajabifard, Mohsen Kalantari and Andrew Binns. In this article, the authors assess metadata entry tools (METs) that should allow for integrated management and updates of spatial data and its accompanying metadata. This would prevent separated data collection processes, and two independent data sets that must be managed and updated - spatial data and metadata. The article highlights the significance of spatial data and metadata integration through developing a set of criteria for metadata application development. These criteria are used to assess a selection of METs.

One of these tools might have been the metadata tool developed and described in *A Prototype Metadata Tool for Land Use Change and Impact Models – a Case Study in Regional Victoria, Australia* by Stephen Williams, Christopher Pettit, David Hunter and Don Cherry. They present a prototype tool for storing and managing model metadata. It extends the utility of the more traditional model register allowing storage of details associated with each instance of a model run. This Model Information Knowledge Environment (MIKE) metadata tool shows promise in assisting the use of Natural Resource Management models within Victoria, Australia and in providing details of modelling activities throughout Victoria.

Practical issues that need to be overcome to successfully document metadata are discussed in *Implementation of Recent Metadata Directives and Guidelines in Public Administration: the Experience of Sardinia Region* by Luisa Manigas, Michele Beneventi, Luca Corvetto, Rita Vinelli and Marco Melis. They provide insight in the process of implementing metadata documentation. Metadata should take into account different metadata requirements of applications. A general description in a national spatial data register requires a limited number of metadata for each single datum, while at the re-

gional level specific and complete metadata need to be described. This article explores a most appropriate database structure and organisation for metadata documentation using the author's experiences in Sardinia, Italy. The article presents a tool, the metadata manager, that should help to manage metadata at the appropriate levels of an SDI.

CRITICAL ASSESSMENT OF SDIS

SDI scholars increasing pay attention to the evaluation and assessment of SDIs. One of the latest publications on SDI assessment is Crompvoets et al. (2008) in which a rainbow of assessment approaches is provided.

A new approach on evaluating SDIs is provided in Ara Toomanian and Ali Mansourian's *An Integrated Framework for the Implementation and Continuous Improvement of Spatial Data Infrastructures.* Their article adds techniques and methodologies from business management literature to the SDI assessment spectrum. The applicability to SDI evaluation of the methods Six Sigma, ABC (Activity Based Costing), BSC (Balanced Scorecard) and TQM (Total Quality Management) is assessed. It appears that these new methods can very well be used for SDI assessment purposes.

Another new SDI assessment approach, this time from an economic perspective, is provided by Elisabetta Genovese, Stéphane Roche and Claude Caron. In *The Value Chain Approach to Evaluate the Economic Impact of Geographic Information: Towards a New Visual Tool* they apply the value chain approach to assess the value that is created step-by-step along the chain. Identifying stakeholders and understand their roles in defining value of geographic information is very important in assessing the value of GI. The prototype *Socioscope* currently provides cartography of the links existing between various public and private contributors. These are the professional users. Other users will be considered and their contribution in the value chain added in a further phase of the research. Only then it will be possible to follow the generation of added value on a specific network of GI flows from original producer to the end-user.

Silke Boos and Hartmut Mueller assess SDIs in twenty-six Mega Cities around the globe in their *Evaluation of Spatial Information Technology Applications for Mega City Management.* Through an internet literature study, the authors sought evidence that developments in the national SDI are correlated with the SDI of the Mega City of that country. The research shows that in many Mega Cities a correlation between NSDI development and urban SDI development exists. They found that often the national SDI development is ahead of the SDI development of its largest urban areas.

Kevin McDougall, Abbas Rajabifard and Ian Williamson explore in their article *Local Government and SDI – Understanding their Capacity to Share Data* local government SDI within Australia to assess its capacity to contribute to higher level SDI initiatives. They undertook a comprehensive survey of over a hundred local government authorities to assess their SDI capacity and collaborative initiatives. Contrary to the general belief in many countries that it is wise to unburden local government from SDI implementation (see, for example, INSPIRE Directive), this research has found that local governments have mature spatial data holdings and the ICT infrastructure to facilitate SDI development through the wider sharing of data. The authors convincingly argue that local government should be viewed as an equal partner in SDI development.

On the crossroad of evaluation of SDI developments and future challenges, Ian Masser examines in his article *Changing Notions of a Spatial Data Infrastructure* some of the

changes that have taken place in the notion of a spatial data infrastructure (SDI) over the last 15 years. He addresses several challenges facing SDI implementation. He questions the 'one size fits all' approach, which is utilised in many SDI initiatives. Where governance activities are composed of specialised task-specific jurisdictions, SDI implementation is likely to stick to the perception of one relevant jurisdiction at each level of the administrative hierarchy. The required flexibility to respond to new needs and circumstances is lacking and Masser advocates that this is one of the future challenges SDI implementation has to overcome.

Cooperation is very important for a successful SDI implementation. This is stressed by Olof Olsson in *Cooperation – a Key Factor for Sustainable Spatial Data Infrastructure*. He, however, also notices that little research explicitly focusing on cooperation in the SDI domain has been accomplished. As a consequence, cooperation as a concept is generally not fully understood within the SDI context. He also notes that the SDI community does not necessarily have the best experts on organisational issues. In this respect he argues that the SDI community should welcome those that are specialised in cooperation to improve our understanding of cooperation as a critical component of an SDI. He introduces the Talk, Decision and Action theory of Brunsson as a means that may contribute to advancing our knowledge on cooperation in SDI.

COMMUNITY BRIDGING

In this part, three articles bridge the SDI community with other (SDI-) communities. Sheelan Vaez, Abbas Rajabifard and Ian Williamson raise in *Seamless SDI Model - Bridging the Gap between Land and Marine Environments* the issue of multiple SDIs covering adjacent and sometimes overlapping areas. Although interaction between these SDIs may be beneficial, currently they develop in isolation. The authors argue that there is a need to create a seamless SDI model that bridges the gap between the terrestrial and marine environments, creating a spatially enabled land-sea interface to more effectively meet sustainable development objectives. The authors identify issues and challenges that need to be overcome to converge terrestrial and marine SDIs.

Rohan Bennett and Abbas Rajabifard address in *The RRR Toolbox: a Conceptual Model for Improving Spatial Data Management in SDIs* large-scale data such as built environment information. The article assesses the applicability of the rights, restrictions and responsibilities (RRR) Toolbox, a holistic framework for understanding, creating and managing land interests to SDI. If a jurisdiction wishes to manage coherently all its RRRs, then each of the eight components of the Toolbox needs to be addressed and acted upon. Preliminary analysis suggests that seven of the eight RRR Toolbox principles hold for SDI. The authors argue that SDI researchers and practitioners should examine the requirements of Land Administration in terms of these seven principles, if only to improve their knowledge and understanding of the RRR based datasets within SDI. Their research suggests that the principles revealed in the RRR Toolbox are generic enough to apply to many SDI initiatives.

And finally, in reverse, SDI theory is applied to catchment management in *Building SDI Bridges for Catchment Management* by Dev Raj Paudyal, Kevin McDougall and Armando Apan. A catchment can be defined as a natural collection area where all rainfall and run-off water eventually flows to a creek, river, lake, ocean or into the groundwater system. A reliable SDI is needed for appropriate decision making and conflict resolution in catchment management. However, the integration of spatial data in such environments has been problematic as the available spatial data often have different scale, content and formats. One major cause is the community centricness of catchment

management issues. They do not follow the rules of administrative hierarchies, but cut across political-administrative boundaries. Since SDI development approaches mostly chose political-administrative boundaries as a starting point, the authors stress the need to re-examine SDI development approaches to accommodate the needs of catchment governance and management.

The contributions in this volume should contribute to SDI development around the globe. It is likely to support the ultimate dream of many in the SDI community: to reach a status of a true infrastructure: it becomes only visible upon break-down; "we are most aware of it when it fails to work- when the server is down. The electrical power grid fails, or the highway bridge collapse" (Star and Ruhleder, 1996). Although it will probably never happen that one will say that "the SDI is down", the many developments in society, all impacting on the status of SDI, are such that in the future the SDI community will truly converge into the information infrastructure community which will also cover geographic information.

REFERENCES

CNA (The CNA Corporation) (2007). National Security and the Threat of Climate Change, at: http://securityandclimate.cna.org/.

Crompvoets, J., A. Rajabifard, B. van Loenen, T. Delgado (2008). A Multi-View Framework to Assess SDI. Space for Geo-Information (RGI), Wageningen University and Centre for SDIs and Land Administration, Department of Geomatics, The University of Melbourne, Melbourne, Australia.

Star, S.L. and K. Ruhleder (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces, *Information Systems Research*, 7 (1): 111-134.

The Potential of a National Atlas as Integral Part of the Spatial Data Infrastructure Exemplified by the New Dutch National Atlas

Menno-Jan Kraak[1], Ferjan Ormeling[2], Barend Köbben[1] and Trias Aditya[3]

[1]International Institute for Geo-Information Science and Earth Observation (ITC), the Netherlands. [2]Utrecht University, the Netherlands. [3]Gadjah Mada University, Indonesia
kraak@itc.nl

Abstract

The recent developments around national spatial data infrastructures have stimulated interest to review and renew the national atlas concept. In a recent project a feasibility study and prototype implementation of an automatic visualisation of spatial data available through the spatial data infrastructure (SDI) have been executed in a systematic and cartographically accountable way to guarantee an up-to-date national atlas. The objective was to investigate how the national atlas could be organised as an integral part of the spatial data infrastructure. The atlas would benefit from an up-to-date data flow, and the SDI would benefit from integrated visual summaries of available spatial data and geoservices in well-designed comparable maps using the narrative characteristics of the atlas. As such the national atlas provides alternative interactive and dynamic access to the SDI.

Keywords: national atlas, spatial data infrastructure (SDI), web mapping.

1. INTRODUCTION

National atlases present a synthesis of the spatial knowledge that characterises a country. They contain comprehensive combinations of spatial datasets represented by maps that each completely covers a country, with an added narrative function. All information in national atlases refers to the same area, the national territory. Generally these maps are based on datasets that are the most detailed available on the national scale. An important aspect of atlases in general, but particularly of national atlases, is that all information is rendered at the same scale and resolution by applying the same level of generalisation. As far as possible, all information in national atlases is collected for comparable reference periods. In addition, as much as feasible or relevant, similar classification methods are used. To allow for fruitful temporal comparisons, national atlases aim to use similar class boundaries and legend colours. Together, this allows for comparing maps in the national atlas and for deducing information. Today national atlases are found both as books and as digital publication on DVD or the Internet (Sieber and Huber, 2007).

Before the maps can be compared, the structure of the cartographic image of the mapped phenomenon in each individual map has to be studied. What is its extent? What is its distribution? What is its pattern? What is its spatial association? What is its spatial interaction? (Board, 1984). In an atlas environment one can elaborate on these questions and process those into atlas use scenario's (see Figure 1).

When knowledge about the overall pattern of a phenomenon (scenario 3 in Figure 1) is required, for example, the overall extent of the phenomenon can be ascertained as well as its sub-regionalisation. Highest and lowest concentrations/values of the phenomenon can be described as well as the spatial anomalies. Studying spatial areal proc-

esses (scenario 5b in Figure 1) requires definition of the initial and final extent of the phenomenon, as well as determining the changes between them and the growth patterns they result from.

Figure 1: Atlas use scenarios.

Scenario 1:   What exist at a specific (point) location?
Scenario 1a: Where is a specific geographical object located?

Scenario 2:   What is (the value) of the phenomenon for a specific area?
Scenario 2a: Which objects are located within a specific area?

Scenario 3:   What is the overall pattern of that phenomenon?

Scenario 4:   What are the relationships between these patterns? [*]
Scenario 4a: What are the temporal trends in that phenomenon?

Scenario 5a: Which linear spatial processes are taking place in that area?
Scenario 5b: What areal spatial processes are taking place in that area?

Scenario 6:   What are the overall geographical characteristics of that area?

[*] Comparison of atlas maps can take different forms: there can be *temporal comparisons* when we compare the situation for a specific theme at different moments in time. Then there are *geographical comparisons*, and for this we produce maps of the same theme, with the same specifications, but for different areas. The point of such comparisons would be for instance the discovery of specific types of settlements, or distribution patterns over an area. Finally there is *topical* or *thematic comparison*, when different map themes are compared for the same area.

Based on the above scenario's the required functionality needed in the atlas can be established. Examples are the ability to put two maps alongside in order to allow for visual comparison, and a direct link between those maps to reflect action in one map immediately in the other maps such scale changes. Zoom and scroll functions are basic, but one could also consider functions like querying the database underlying the data used for the map, or the ability to query all map objects.

Why use national atlases at all when searching for spatial data? All users of geospatial information had atlases when they were first confronted with this kind of data. At school they were taught how to deal with them through the concepts that the school atlases were building on: areal and thematic subdivisions, map comparison, geo-referencing, datum's, among others. A national atlas will not only provide access to the atlas information, but also access to the underlying datasets and - when functioning as a geo-portal - also to all other related datasets made available by the national geospatial data providers. Thus, it offers these data providers also a 'presentation outlet'. The main benefits of having the national atlas as the portal towards the nation's geospatial information are: ease of use because of familiar concepts and ease of access because of the topical atlas structure.

2. THE DUTCH SITUATION

Although the first actions to produce a national atlas started already in 1929, it was only after the second world war, that the first edition of the Netherlands national atlas materialised (1963-1978). This was due to the global economic crisis in the 1930s and the decision to first produce a national atlas of Indonesia To that end a Foundation for the Scientific Atlas of the Netherlands was set up in 1958, backed by the Ministry of Educa-

tion, the Royal Netherlands Geographical Society, the Topographic Survey and universities. An Atlas Bureau was set up at the national physical planning agency, paid for by the Ministry of Education. This atlas was an inventory of all geospatial data, perhaps more targeted at the area than at its inhabitants. A most detailed soil map 1:250 000 formed its backbone, but it is difficult to perceive it as a narrative of the country, more as an incidental combination of contributions from various fields of science.

In the second edition, published 1989-1995, this was mended and this edition is clearly centred on the inhabitants of the country and only deals with aspects of the sciences (climate, geology, soil, etc) when this was deemed relevant for explaining the way Dutchmen provided for themselves. So the atlas contained no geological maps per se, but maps of economic geology or of the strata from which natural gas could be mined. The Atlas Bureau, transferred to the national mapping agency, could continue its ministry-backed work for the production of this second edition, but when it was completed the Ministry's outlook had changed, claiming that from now on such endeavours as a national atlas should be self-supporting, and resulting in the atlas bureau's closing. However, the foundation was to keep the national atlas concept alive, and to this end it was decided to make all maps from the previous two editions available on the web (through http://www.nationaleatlas.nl). With 1,5M hits since mid 2000 this website seems to answer the needs of high school students that have to do projects as well as the geospatial information needs of a larger informed audience. But, as the information contained in these scanned maps is getting out of date, new initiatives were needed.

The developments around the SDI in the Netherlands are seen as a great opportunity to revive the National Atlas. It might stimulate and improve the accessibility of the SDI by the narrative of the atlas. The atlas itself makes use of the spatial data and geo-services available through the SDI to create interactive and well designed atlas maps. The various data sets provided would be made comparable, and be visualised according to specific templates. Apart from serving as any map in a national atlas would, these maps may also function as an alternative entry to the SDI. They should be clickable and provide the underlying statistics used, but also allow to zoom in on a specific area and taxonomy level in the atlas so that all the data sources for that query combination are displayed as footprint on the atlas maps and accessed via hyperlinks. The highest levels of high school are likely to show interest in the ready-made maps, and professional users in the geoportal function.

The philosophy behind the realisation of the Dutch national atlas is to enable the creation of the 3rd edition of the national atlas from datasets available from national data providers (like Central Bureau of Statistics, or Geological Survey) by using a style template that would make the resulting maps comparable. The shell in which these maps could be produced requires functionality that allows for, for example, map comparisons and data queries, based on the atlas use scenario's mentioned in the map use section. This functionality is covered in section 5 of this article, but first the atlas will be positioned in the national spatial data infrastructure.

3. THE DUTCH SPATIAL DATA INFRASTRUCTURE

The stimulus for spatial data infrastructures originates from the motto "collect once, use many times". They have been defined by Groot and McLaughlin (2000) as "A set of institutional, technical and economical arrangements, to enhance the availability (access and use) for correct, up-to-date, fit-for-purpose and integrated geo-information, timely and at an affordable price, with the goals to support decision making processes related to countries' sustainable development". In the European Union, the INSPIRE Directive

implements the SDI concept. INSPIRE emphasises several basic principles which are also applicable to the national atlas concept as presented here (see Figure 2):

1. Data stewardship and data security. Data is collected once and maintained at the level where this can be done most effectively. For the atlas it means that it does not need a database of its own.
2. Conditions of data accessibility, and data interoperability. It is possible to combine seamlessly spatial data from different sources across Europe and share it between many users and applications. For the atlas this translates to a national level where similar problems occur.
3. Data reusability and data synchronisation. It is possible for spatial data collected at one level to be shared with other levels, e.g. detailed for those performing exhaustive investigations, but more general for strategic purposes. For the national atlas this might be less relevant since data is only presented at a single national level.
4. Data availability. Spatial data is abundant and widely available under conditions that do not restrain its extensive use. The atlas concept as such follows this principle.
5. Data discoverability, data validity and data rights. It is easy to discover which spatial data is available that fits the needs for a particular use and under what conditions it is available. An atlas has several relevant facilities for data discovery, such as an index for geographical names, a topical index, and index maps. The atlas can act as an integrated visual summary of available spatial data and geoservices and as such act as alternative interactive and dynamic access to the SDI.
6. Data usability. Spatial data must become easy to understand and interpret because it can be visualised within the appropriate context and can be selected in a user-friendly way. This is the atlas' natural habitat. It has a narrative to tell the story of the atlas objective is well design maps.

Figure 2: The contributions of a national atlas embedded in the spatial data infrastructure.
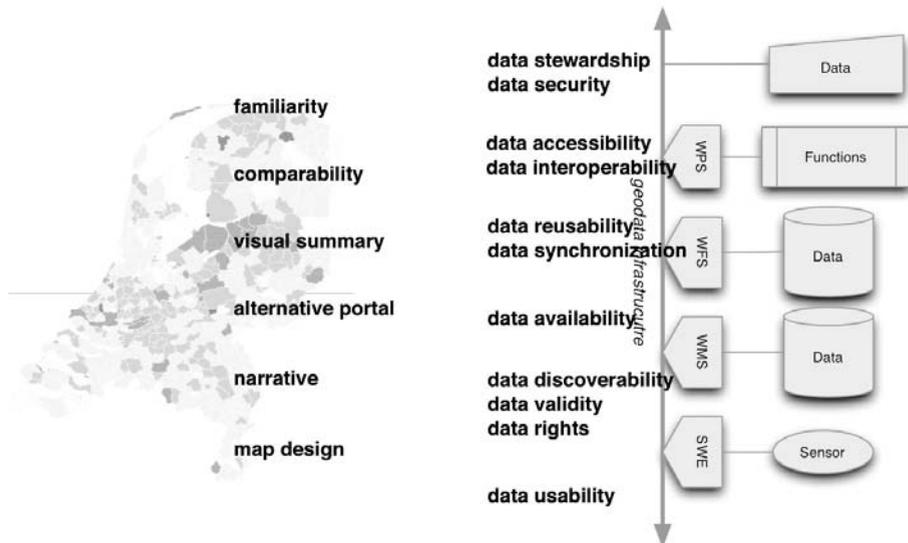
In the Netherlands the INSPIRE implementation is translated into practice guided by the GIDEON report (Ministry of Housing, Spatial Planning and Environment, 2008). The

12

challenge for the National Atlas Foundation is to have the new national atlas concept embedded in the national spatial data infrastructure, based on the recognition that a central place is needed where all different spatial datasets are made comparable, and that can moreover function as a most useful geoportal.

4. THE NEW NATIONAL ATLAS CONCEPT

In accordance with the scenario's and required functionality described above a proto-type of the national atlas (3rd edition) was designed and implemented (see http://www.nationaleatlas.nl/) Figure 3 summarises the main components of the atlas. The sections 5 and 7 address the atlas in more detail. Displaying maps is the main objective of the atlas. If a user selects a topic, for instance the number of inhabitants per municipality (A-I), the request goes to a geo-service, which returns the necessary data that allows for the creation of an interactive map (A-II). Alternatively it is possible to search for a topic or for a geographical name (B-I). In this last case the geo-service will return all names with the text string entered and map topic on the base map (B-II).

All maps are interactive and allow for the display of the data behind the symbols. It is also possible to search the SDI for alternative datasets (C-I). Through the atlas maps, which will display the footprint of the available datasets, the metadata of those datasets can be evaluated (C-II).

One of the characteristics of an atlas is that one can compare different themes, for instance the distribution of the young or of the elderly (D-I). Such request results in two maps that allow for the comparison of spatial patterns. For each topic the atlas provides a narrative, the story behind the map in a wider context (E-I). This will also result in access through web-links to other related information accessible via the SDI (E-I). Finally, it is possible to export (F-I) atlas maps to a Google Earth environment where users might combine the particular maps with their own data (F-II). However, it will also be possible to import (G-I) user data to be combine with the atlas maps (G-II).

The necessary technology to make this concept work has been elaborated by a PhD study (Aditya, 2007). For the realisation of the organisational framework it is necessary to convince the national spatial data providers to make their data accessible and comparable through this newly developed GUI (Geographical User Interface), and make them realise that the atlas would provide for added value apart from providing an extra presentation outlet.

5. THE ATLAS FUNCTIONALITY

As indicated in section 1 (see Figure 1), a number of scenarios has been developed based on the functionality needed (Simon van Leeuwen, 1996) to enable those specific types of map use. Elementary atlas functions include the ability to put maps next to each other to allow comparison. Database functions allow for querying the database 'behind the map' such as the statistical dataset the map is based on, but also the topographical elements regarding their object type, name, class or category they belong to or exact value for the phenomenon in question. Basic cartographic functions are the ability to zoom or scroll. In some cases, for larger countries than the Netherlands, this should also include the ability to change the projection of the map. Educational functions are those that monitor achievements of the students. Navigation functions refer to the possibility to follow specific pre-set paths through the atlas, in order to present the various maps in a specific self-explaining sequence that makes sense (the narrative). Generic functions such as import, export or print the data or the resulting maps are in-

cluded. Map functions refer to pop-up legends, the highlighting of specific legend classes, or the use of hotspots, but also offer the possibility to annotate maps, measure distances or surface areas, or use buffers and overlays.

Figure 3: The national atlas concept. The specific atlas functions have been integrated with geo-web services available via the spatial data infrastructure.
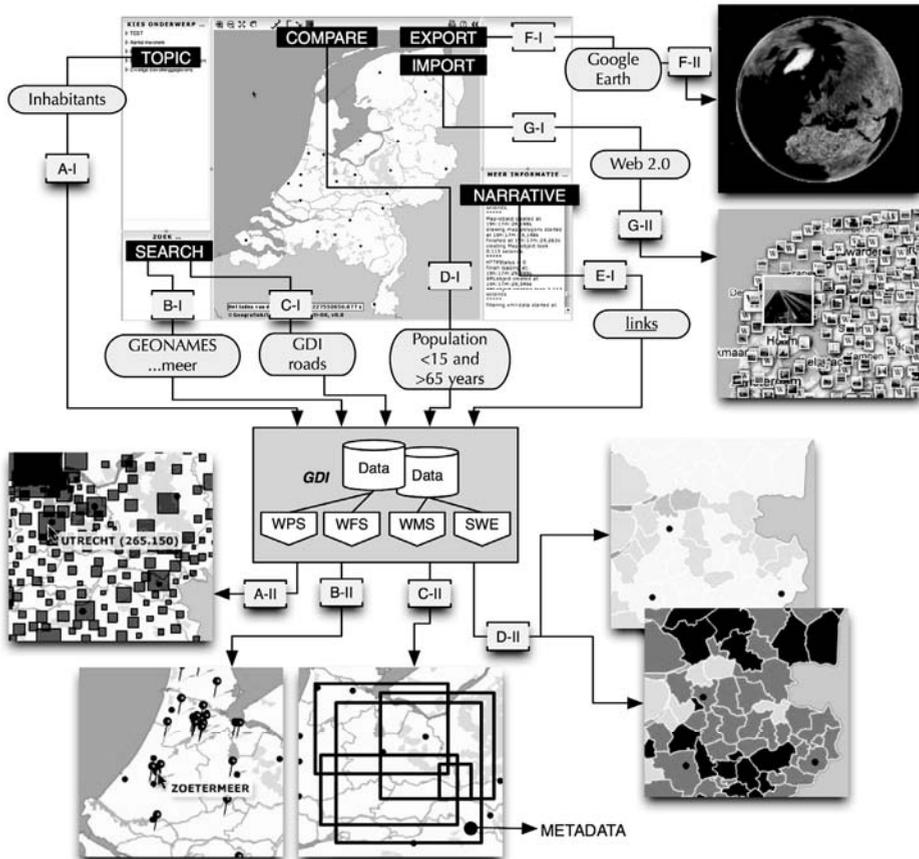
These functionalities provide for a first level of map use, to be extended to include the possibility to describe and summarise spatial data, to generalise concerning (related) complex spatial patterns, to use samples of spatial data to infer characteristics for a larger set of spatial data (population), to determine if the magnitude or frequency of some phenomenon differs from one location to another, or to learn whether an actual spatial pattern matches some expected pattern or the pattern of another phenomenon (correlation coefficients). Not all of these can be realised in a national atlas, so some form of prioritising has to be applied.

The development of the national atlas interface focuses on three principles. First the Atlas will provide for a uniform interface to the Dutch SDI where specific attention is paid to well-designed maps. Providing an overview is more important than in-depth analysis, which limits the scale of the maps. Secondly, the Atlas will have a modular design and therefore be able to serve different groups of users. Finally, the user should

experience 'instant satisfaction' using the Atlas. Speed while loading and manipulating the maps and a clear and easy to use interface are essential to achieve this goal. The interface of the national atlas will eventually include two components, an editor's interface and a user interface. The editor's interface, a tool to manipulate maps and data, will be developed later on.

The user interface basically has three windows, each divided in two panes. The window on the left side contains a list of topics and a search module. The window in the middle contains a toolbar and map area, and the window on the right side contains the key and the storyteller, where the user will find additional information to the map. The panes of the interface (except the toolbar) can be resized by dragging bars (see Figure 4).

On entering the Atlas, a default map is automatically shown. The topic of this map may pay attention to current events and should change regularly. To browse through the Atlas one can either search by keyword or location, or click one of the predefined topics from a list. For advanced search operations the search pane has a link to an advanced search pane. Here one can create queries using a combination of keywords and location, exclude certain keywords or locations plus specify a timeframe for the topic one is interested in. In the pane below the search pane, a taxonomic list is also giving entrance to the Atlas. For selecting more than one topic at the time, one can link to an expert pane where this is possible. The visualisation of multiple subjects chosen from the list of topics has three options, considering the nature of the chosen topics. The topics can be displayed on top of each other in one map, displayed in two maps next to each other, or after each other as an animation or slideshow. On selecting a topic, a link to the producer of the data underlying the specific topic is shown in the storyteller.

Figure 4: The user interface of the national atlas (http://www.nationaleatlas.nl/).
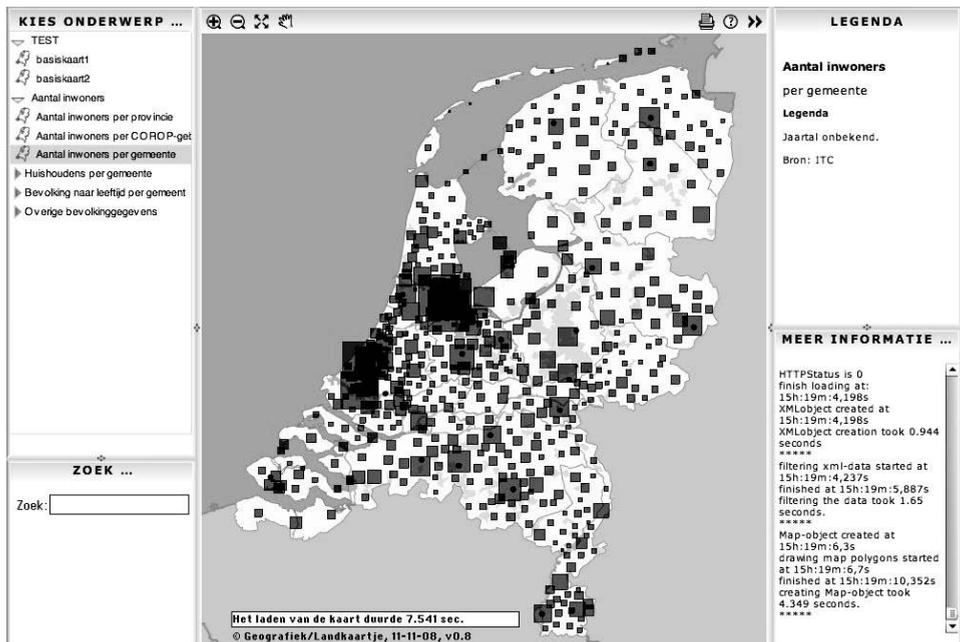
The toolbar at the top of the map area contains several easy recognisable pictograms giving access to zooming, panning and printing operations. Here one can also click a help-button for more information on how to use the atlas, while an additional button offers access to more tools like measuring and exporting the map or underlying data in a desired format. A tool-tip explains the functionality of a buttons when the cursor moves over. Clicking the 'more tools' button not only gives access to more tools, but also affects the way in which the user can manipulate the map and thus entering an 'expert mode'. In a movable pane one can, for instance, turn layers off and on or enter thresholds. The map area can contain raster or vector images, or a combination of both. The maps in vector-format will offer interactivity like clicking on an area shows additional information of that area in the tool-tip or in the storyteller pane. However, the design of the map cannot be manipulated. The visualisation of the maps will be predefined in style-sheets to maintain consistency in design throughout the Atlas. For overlaying two topics, however, the alpha of the upper layer and the saturation of the underlying layer can be manipulated.

The key to the map will explain its content, the internal identification and show elements of external identification such as the title of the map, the scale, source and copyright. Depending on the type of map the key will be more or less complicated. The storyteller is a window where all kind of additional information can be found. Here the modular design of the Atlas becomes very apparent and functional. In the standard modus one can take an atlas tour, which will guide the user through the Atlas in a sequence comparable to a paper atlas. By selecting topics from the list one can move from one topic to another whereby the accompanying storyteller will have a link for reading further on the matter, and thus entering the expert modus of the atlas. The atlas maps are designed according to cartographic guidelines found in text books and common practice.

6. NATIONAL ATLAS AND SDI

For a new Dutch National Atlas to be part of the national spatial data infrastructure, it should fit into the framework of interoperable geo-web services that make up such a SDI. There are many geo-web services available (e.g., Google Maps, Yahoo Maps, MSN Virtual Earth or MultiMap) that can be used by anybody, as their interfaces are publicly available. However, they are still proprietary since they are defined, developed and owned by commercial companies.

There is also a set of well-defined open standards for geo-web services: the Open Web Services (OWS) of the Open GeoSpatial Consortium (http://www.opengeospatial. org/standards/). There are OWS specifications for most parts of the spatial data storage, analysis and delivery process: for describing and finding spatial data there is a set of metadata specifications in the Catalog Service Web (CSW); for geographic vector data encoding there is the Geographic Markup Language (GML); for spatial data delivery the Web Coverage Service (WCS) and Web Feature Service (WFS), for querying and retrieving raster and vector data respectively; for processing of spatial data there is the Web Processing Service (WPS). And for data visualisation in the form of maps we have the Web Map Service (WMS), by far the most mature and widest adopted OWS specification. There are numerous open source as well as commercial solutions offering WMS functionality. Related to WMS are the Styled Layer Descriptor (SLD) specification, for map styling, and the Web Map Context Documents (WMCD) specification, for map setup and layout.

The envisaged architecture of the national atlas in the national spatial data infrastructure will be employing the OWS specifications in a multi-tier setup. At the server side

data are found offered by the data providers as well as the atlas itself. The data layer contains internal data that supports the atlas visualisations, for instance base map files, descriptive text, images and charts. The external data is not stored at the atlas server but retrieved from data providers. This can also be non-spatial data services, for instance statistical data. Metadata summaries describe all these datasets.

In many existing SDIs the maps are delivered to the end-user from WMSs. Mapping clients can combine the output of several such services in a compound map. However, this is not a desirable solution, as the portrayals of these layers are not matched to each other and to the larger overall atlas goals described earlier. Therefore the national atlas should primarily use data services instead of portrayal services: It will consume data from WFS sources, and combine and portray that data in a cartographically sound client-side application. This application is supported by the data integration and mapping components on the atlas server. The former is important in the harmonisation of data for our specific visualisation purposes. As an example, it is not uncommon for various data sources for socio-economic mapping to employ different spatial datasets of the same mapping units (e.g. municipalities), with different accuracies, and generalisation levels, for example. The data integration component will always use the one best suited for the scale and map types used in the atlas, and for other data providers only extract the attribute data and re-map these onto the preferred spatial data. The mapping component is used to determine which cartography (map type, classification) is suitable for specific data layers and their combinations, and which design templates are required to achieve this.

Once the mapping output and required template are defined the data is displayed on the client side. To achieve high quality and interactive visualisations, a powerful vector graphics technology should be used, such as Scalable Vector Graphics or Flash. Alternative mapping platforms such as Google Earth will also be supported, and the mapping component will also provide output formats suited for that, such as a combination of KML (now also an OGC standard) and Google Earth Network Links.

The technical implementation of the current prototype can be seen in Figure 5. Being a prototype, it does not implement all of the elements mentioned above. This is partly because the Dutch National SDI does not yet include the data services needed. We therefore host some of these, such as the WFSs for socio-economic mapping (from the provincial to the municipal level), internally on the atlas server itself. This also applies to the place name gazetteer that powers the location search. However, the actual data from the original data providers are used. As soon as these data providers implement their data services, only a simple URL change in the OWS requests is required. Furthermore, at present only the Flash version of the mapping client is implemented, since it is based on existing efforts. Open standards, such as SVG, instead of the proprietary Flash technology would in the longer term be desirable. The technical implementation of the Google Earth component has been tested (see Graham, 2008; and the test site at http://geoserver.itc.nl/natatlas/GE/), but has not been integrated in the prototype yet.

7. UPDATES, EXTENSIONS AND FUTURE

The integration in the spatial data infrastructure should guarantee that the most recent data available can be used in the atlas. If for instance the Census Bureau releases new annual population statistics through the geo-services that are already in use by the national atlas, an automatic update would theoretically be possible. However, since the atlas works with its 'own' basic administrative boundary data it might happen that administrative boundaries have changed and the atlas boundary set have not. Also the

Figure 5: Technical implementation of the national atlas prototype in the national SDI.
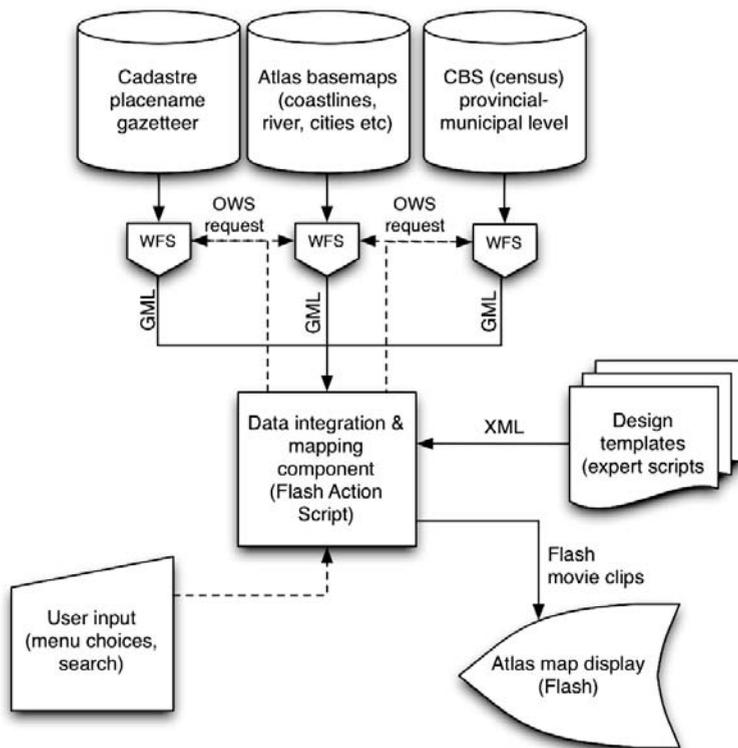
statistics themselves may require a different classification due to strong increase or decrease resulting in new class boundaries. This may result in a changing legend and possibly a changing map design. Therefore the national atlas cannot do without an atlas bureau with staff. They are responsible for the quality of the atlas. This atlas bureau, apart from editing maps and regularly extending the number of ready-made maps available through the site, should keep tabs on both the new spatial information made available by national data providers as well as considering the changing needs and interests of users such as the general public, schools and professionals.

The national atlas users are also familiar with other mapping tools such as Google Earth and Google Maps. The atlas has an export function to Google Earth where users can combine the atlas data with their own data. It will be a challenge to see how the cartographic design (the 'atlas template') fits and functions within a Google environment, which still mainly consists of satellite and aerial imagery. Atlas data availability through Google might also attract more visitors to the national atlas and the SDI.

The trend of users combining Google Earth spatial data with data collected by themselves is a recent development also know as neo-geography (Turner, 2006) or Volunteered Geographic Information (Goodchild, 2007). Users could combine their data with the atlas maps as well if the atlas would have an import function. However, form the perspective of some topics it is an interesting question if it would be possible to use these communities to update/extend the map content. Not all content benefits from this approach. Topics not directly observed or measured (like geology) or those which are already very well measured (like weather) would not qualify. However, topics like the

spread of flora and fauna, especially in the light of changing climate, could benefit. It raises all kinds of questions. How to evaluate the observations of these communities? In the traditional national atlas map workflow the scientist would provide the data, the narrative and draft maps, and one might expect a certain quality. If one intends to include Web 2.0 communities it is likely the workflow has to be changed, but how? A national atlas bureau would not be able to check all observations, irrespective from which community these would come. Goodchild talks of citizen scientist when these informal communities have their own serious protocols, such as for instance bird watchers. Their approach is probably self-cleaning enough, but how to deal with the input from others? Similar questions are relevant for the formal SDI world.

For professionals and students working with spatial data the ease of accessing and combining spatial datasets through the national atlas interface can be harvested in geo-collaboration. Here multiple users at different locations can address and amend the same image on line in planning sessions or emergency situations, deciding interactively with the shared cartographic image as medium, on the course to follow. Way beyond its static and almost per definition outdated information provision image, the national atlas that is incorporated in the spatial data infrastructure is getting a new lease of life, with many opportunities to play a useful role in middle of the dynamic world of geoinformation where everyone is a cartographer.

ACKNOWLEDGEMENTS

REFERENCES

Aditya, T. (2007). *The national atlas as a metaphor for improved use of a national geo-spatial data infrastructure*. PhD thesis, Enschede: ITC.

Board, C. (1984). High order map-using tasks: Geographical lessons in danger to be forgotten, *Cartographica*, 21(1): 85-97.

Goodchild, M. (2007). Citizens as sensors: The world of volunteered geography. *Geojournal*, 69: 211-221.

Graham, M. (2008). *Integrating large volume data services with Google Earth mapping*. IFA report, Enschede: ITC.

Groot, D., and McLaughlin, J. (2000). *Geospatial data infrastructure - concepts, cases, and good practice*. Oxford: Oxford University Press.

Ministry of Housing, Spatial Planning and Environment (2008). *GIDEON - key geo-information facility for the Netherlands - approach and implementation strategy [2008-2011]*. Den Haag: VROM.

Sieber, R. and Huber, S. (2007). Atlas of Switzerland 2 - A highly interactive thematic national atlas. In W. Cartwright, M. P. Peterson and G. Gartner (Eds.), *Multimedia cartography*. Berlin: Springer, pp. 161-182.

Simon van Leeuwen, W. (1996). *Het evalueren van geografische software*. MSc thesis, Utrecht: Utrecht University.

Turner, A. J. (2006). Introduction to Neogeography. O'Reilly Shortcuts.

Development and Deployment of a Services Catalog in Compliance with the INSPIRE Metadata Implementing Rules

J. Nogueras-Iso[1], J. Barrera[1], A.F. Rodríguez[2], R. Recio[1], C. Laborda[3] and F.J. Zarazaga-Soria[1]

[1]Computer Science and Systems Engineering Department, University of Zaragoza Spain, {jnog,jesusb,rociorm,javy}@unizar.es
[2]National Geographic Institute of Spain, Madrid, Spain, afrodriguez@fomento.es
[3]GeoSpatiumLab, Zaragoza, Spain, claborda@geoslab.com

Abstract

In order to facilitate the availability of and access to spatial data, Spatial Data Infrastructures must set up a series of services to be reused by their community of users in the construction of different applications and value-added services. One of the key elements to exploit the resources provided by these infrastructures is to facilitate a catalog of services describing the features of the services offered to their users. This article presents the development and deployment of a services catalog within the Spatial Data Infrastructure of Spain, a catalog in compliance with the INSPIRE implementing rules.

Keywords: Metadata, Services, Spatial Data Infrastructures, Services Catalog.

1. INTRODUCTION

According to the Global Spatial Data Infrastructure Association Cookbook (Nebert, 2004) "the term Spatial Data Infrastructure (SDI) is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data". SDIs provide the framework for the optimisation of the creation, maintenance and distribution of geographic information at different organisation levels (e.g., regional, national, or global level) and involving both public and private institutions.

From a technical point of view, in order to facilitate the access to and exploitation of spatial data, SDIs must set up a series of services to be reused by their community of users in the construction of different applications and value-added services. In addition, these services must be accessible on the Internet through standardized transport protocols and interfaces established by organisations and standardisation bodies such us OGC (Open Geospatial Consortium), ISO/TC211 (International Organization for Standardization, Technical for Committee Geographic Information/Geomatics) or W3C (World Wide Web Consortium).

Therefore, one of the key elements to exploit the reusability of the resources provided by an SDI is to set up a services catalog that provides metadata describing the features of the services offered to the community of users. The objective of a services catalog is to help users in the task of finding the most appropriate service according to their needs and requirements.

Up to now and as far as implementation of SDI initiatives is concerned, most of the efforts have been devoted to the development of spatial data catalogs. However, catalogs for describing services have received little attention. The objective of this article is to present the development of a services catalog, which has been deployed within the

Spatial Data Infrastructure of Spain (Infraestructura de Datos Espaciales de España - IDEE). The Web client application of this services catalog to be integrated within the IDEE geoportal provides an easy and intuitive interface to browse and evaluate the features of services accessible through this SDI. Additionally, it must be noted that this services catalog complies with the most relevant standards and specifications defined at international level, making a special emphasis on the compliance with the metadata implementing rules of the INSPIRE Directive (EC, 2007), a Directive for establishing an Infrastructure for Spatial Information in the European Community.

The rest of this article is structured as follows. Section 2 describes the state of the art in metadata schemas for service description. Section 3 presents the design of the services catalog, including a description of its architecture, the adoption of INSPIRE metadata implementing rules for metadata modelling, and the automatic method proposed to derive metadata from the capabilities information provided by OGC services. Section 4 describes the deployment of the services catalog in the Spain SDI. Finally, this article ends with some conclusions and proposals for future work.

2. STATE OF THE ART IN APPROACHES TO SERVICE DESCRIPTION

Within the context of SDIs, the first approach for service description arose with the definition of the first specifications for OGC services more than ten years ago. For descriptive purposes, all OGC service specifications support a *getCapabilities* operation to obtain service-level metadata (also named as capabilities) describing the content and acceptable request parameters of an OGC service. The XML responses to *getCapabilities* requests include information about service identification (general metadata for discovery such as information as title, abstract, or keywords), service provider, and available operations.

In 2005, ISO/TC211 defined an extension to the ISO 19115 (ISO, 2003a) geographic metadata standard for the description of the specific features of services. This extension has been defined within the context of ISO 19119 standard (ISO, 2005). This standard aims at establishing the foundations of geographic information services. It provides a taxonomy of geographic information services, and it extends the ISO 19115 comprehensive model with specific elements and data types for service descriptors. This proposal has been also adopted by OGC in the definition of one of the profiles of the CSW (Catalogue Services for the Web) protocol binding for catalog services specifications. This profile, called "ISO Metadata Application Profile" (Voges and Senkler, 2007), proposes a combination of the models in ISO 19115 and ISO 19119 as the information model followed by the metadata records to be managed through catalogs.

Recently, the European Commission started a procedure for the adoption of measures to implement the INSPIRE directive as regards metadata (EC, 2008a). These implementing measures define at an abstract level those descriptors that are essential for the discovery of data and services. Besides, these measures are accompanied with non-binding guidelines to establish the mapping between this set of abstract descriptors and the most important metadata standards such as ISO 19115 or Dublin Core (also adopted by ISO as ISO15836 (ISO, 2003b)). In particular, the guideline containing the mapping to ISO 19115 (EC, 2008b) also includes the mapping to the specific service descriptors defined in ISO 19119.

However, the above proposals seem insufficient to satisfy the current needs for building applications following a Service Oriented Architectures (SOA) (Lieberman, 2003), where the syntactic and semantic description of services is highly relevant. Nowadays,

the development of the services offered by spatial data infrastructures, and in general the development of services in any type of networked infrastructures, is usually guided by the Web Services Architecture proposed by W3C (Booth et al., 2004). This architecture aims at providing a standard means of interoperating between different software applications (the Web services), running on a variety of platforms and/or frameworks. According to the W3C Web Services Glossary (Haas and Brown, 2004), a Web Service is a software system designed to support interoperable machine-to-machine interaction over a network. Web Services are self-contained, self-describing, modular applications that can be published, located, and invoked across the Web (Doyle and Reed, 2001). The great impact of Web Services has increased the importance of metadata that describe the processing capabilities of services. The details of a Web Service can be published in a catalog, so that a client's (or another service's) request for such a service can lead to the client invoking that service.

The leading and most accepted standard for service syntactic metadata is WSDL (Web Services Description Language). WSDL (http://www.w3.org/TR/wsdl) is a means of describing a service connection (operation signature or binding) for software to connect to it. Service Directory specifications like UDDI (Universal Discovery, Description and Integration of Web Services protocol, http://www.uddi.org/) can use WSDL to express the machine-readable connect to a service. The main disadvantage of WSDL is that it does not have the ability to characterize the semantic capabilities of services. OWL (Web Ontology Language, http://www.w3.org/TR/owl-features/) is one of the options to formalize the semantic description of services, and in particular OWL-S (http://www.w3. org/Submission/OWL-S/), which is an ontology based on OWL to model the features of Web services. OWL-S is used to annotate services with semantic descriptions. Other options for semantic annotation are WSMO (Web Services Modeling Ontology) or WSDL-S (Web Service Semantics). On the one hand, WSMO (http://www.wsmo.org/) provides a conceptual model to describe the semantic aspects of services in order to automate their discovery, invocation and composition. The WSMO infrastructure has defined WSML (Web Services Modeling Language) as representation language. On the other hand, WSDL-S (http://www.w3.org/Submission/WSDL-S/) defines a mechanism to associate semantic annotations to Web services already described with WSDL.

As a conclusion of this state-of-the-art discussion, it can be stated that the current initiatives in the SDI context still tend to use metadata models (e.g., ISO 19119 extension of ISO 19115) that are more focused on the description of services for mere discovery purposes than in providing the means for the automatic invocation and composition of services. However, it can be noticed that there is an increasing consciousness about the need to move towards more interoperable descriptions of services. For instance, OGC has defined an application profile of the CSW protocol (Martell, 2007) that is compatible with the ebXML (Electronic Business using eXtensible Markup Language, http://www.ebxml.org/) infrastructure, which consists in a modular suite of specifications to exchange business messages, conduct trading relationships, communicate data in common terms and define and register business processes. It must be noted that one of the requirements for the services catalog proposed in this article is to be compliant with the INSPIRE implementing rules. Therefore, the metadata model proposed will follow the guidelines for mapping the metadata implementing rules to ISO 19115/19119 (see section 3.2).

3. DESIGN OF THE SERVICES CATALOG

This section describes the architecture of the services catalog. Section 3.1 presents the architecture. Then section 3.2 presents the metadata model followed by the metadata

records managed through the services catalog. Since one of the main requirements of the services catalog presented in this article is the compliance with the INSPIRE implementing rules, this section describes the implications in the modelling and encoding of service metadata. Finally, section 3.3 describes the process followed to derive service metadata from the capabilities information obtained through OGC services.

3.1 Architecture

The services catalog has been structured following a multi-layer architecture model, where the different components have been grouped in different levels according to their functionality with respect to data access, processing, or interaction with the final user. In particular, three architectural layers have been distinguished (see Figure 1): (1) the *Data Sources* layer includes the different storage repositories used by the services catalog, (2) the *Services* layer integrates the components in charge of the access to data (*Access Services*) and their processing (*Application Services),* and, (3) the *Web Applications* layer consists of the components that interact with the end user, either receiving its requests, or providing the results generated in the lower layer.

As it can be observed in Figure 1, the main component in the top level of the architecture (*Web Application* layer) is the *Services Catalog* application, which is the application that a final user can access through a Web browser. This application provides its functionality thanks to the use of two components called *Metadata Manager Client* and *Search Client*. Despite the fact that they are embedded in a single application, they could be the base of independent applications. Whereas the aim of the first component is the creation, elimination and modification of services metadata; the second component is in charge of querying the catalog and showing the results to the user. This *Services Catalog* application also enables online connections with the services returned by the catalog thanks to the integration of generic clients compliant with the most common OGC specifications. In order to facilitate the integration of components in this layer, they have been developed using the *Google Web Toolkit (GWT)* technology. This technology, sponsored by Google, provides a set of free software tools to build web applications with AJAX using Java as programming language.

With respect to the *Services* level, we can distinguish two categories of components: an *Application Services* category including the components that carry out tasks of data processing; and an *Access Services* category integrating the components that deal with data and information retrieval. Within the *Application Services* category we can find four components using the Java servlets technology: *Edition Server*, *Selection Server*, *Search Server*, and *OGC CSW*. The first two servlets are designed to give support to the *Metadata Manager Client* application. *Edition Server* provides a machine-readable definition (in XML format) for the Graphic User Interface (GUI) of the service metadata edition forms to be displayed by *Metadata Manager Client*. *Selection Server* provides management operations (i.e. insert, update and delete operations) to update the contents of the metadata repository. The *Search Server* servlet, invoked by the *Search Client* application, provides query and present operations. It processes the restrictions found in client queries and returns a list of results satisfying these restrictions. The list of results may be optionally grouped and sorted according to different criteria. And finally, the *OGC CSW* component offers a standardized interface to the services catalog according to the OGC specifications (Nebert et al., 2007). In particular, this component implements the CSW (Catalog Services for the Web) protocol binding to allow the communication between catalog clients and servers over HTTP.

Figure 1: Services Catalog general architecture.
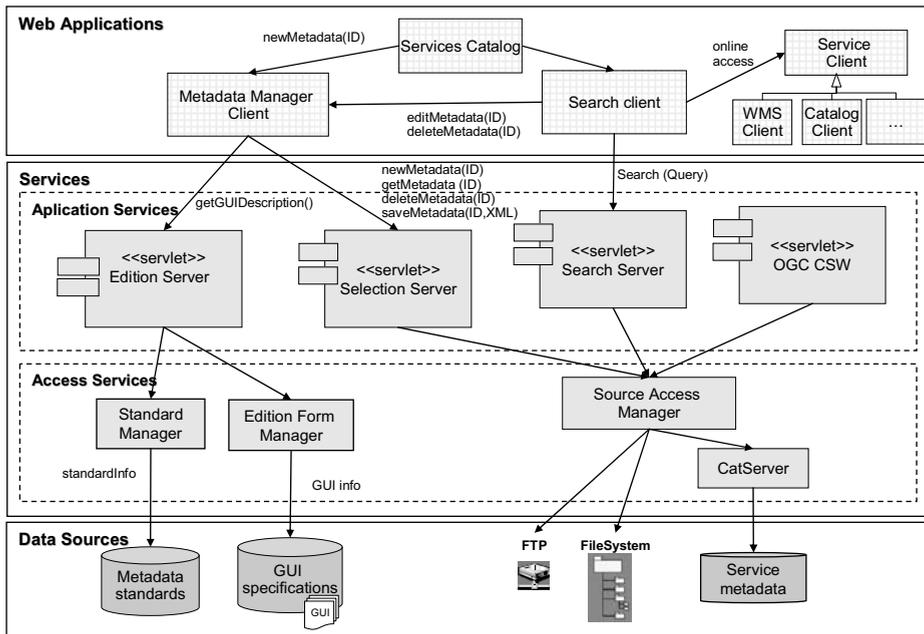
The second category of the *Services* layer includes three main components: *Source Access Manager, Standard Manager*, and *Edition Form Manager*. *Source Access Manager* is the component in charge of retrieving metadata from the final storage device. It provides an abstraction layer over the different types of metadata sources facilitating a uniform access mechanism to the components in the higher layers of the architecture. That is to say, thanks to *Source Access Manager* other components do not need to worry about the storage device, which may be either a file system accessed via FTP, or something more complex such as a XML metadata database implemented on top of a relational database (for instance, see the *CatServer* system described by Tolosana-Calasanz et al., 2005). This component is accessed by three components in the *Application Services* category: *Selection Server* invokes it to perform management operations on the metadata repository; *Search Server* uses this component to find the metadata records satisfying the user queries; and *OGC CSW* provides a standardized wrapper to access the operations offered by this component.

In order to understand the functionality of the other two components that belong to the *Access Services* category, it is necessary to know the mechanism used for metadata edition. This mechanism is the one used in version 4.0 of the *CatMDEdit* desktop application (Zarazaga-Soria et al., 2003; Nogueras-Iso et al., 2008). Using a machine readable definition of metadata standards and a set of rules for GUI layout, this mechanism generates dynamically the edition forms to modify the contents of metadata records in conformance with the correspondent metadata standard. With this purpose in mind, *Edition Server* (in the *Application Services* layer) invokes the *Standard Manager* component to recognize and recover the definition of the metadata standard followed by the metadata record(s) to be updated. Then, *Edition Server* invokes the *Edition Form Manager* component to generate the GUI description of the forms to be displayed by the *Metadata Manager Client*. The *Edition Form Manager* creates the GUI description as a

result of applying GUI production rules to the elements contained in the metadata standard. Each GUI production rule specifies the most appropriate GUI component (e.g., text field, list, text area, choice) for the data type of each metadata element. In the case of the *Services Catalog*, metadata conforms to the INSPIRE metadata implementing rules and their mapping to ISO 19115 and ISO 19119 metadata standards. But this automatic mechanism for metadata edition could be applied to develop other Web metadata editors in conformance to other metadata standards.

Finally, the bottom layer of the architecture (*Data Sources* layer) consists of three different data sources: the data source where the information about metadata standards is stored; the data source with the rules for GUI layout; and the data source for storing service metadata. All these data are encoded in XML format and, although all the data sources used in the IDEE Services Catalog are file systems, the final storage device for service metadata could have been accessed through more complex mechanisms such as FTP or CatServer (accessing an XML database on top of a relational database)*.

3.2 Adoption of the INSPIRE metadata implementing rules for metadata modelling

As mentioned in the introduction, one of the main requirements of this services catalog is to be compliant with the INSPIRE implementing rules, which will dictate the development of national SDIs in Europe in the following years. Therefore, the metadata model proposed will follow the guidelines for mapping the metadata implementing rules to ISO 19115/19119 (EC, 2008b), which is the more mature existent guideline to translate the INSPIRE implementing measures into a particular metadata standard.

As an example of this metadata model, Figure 2 shows the ISO 19119 elements for service identification that must be included in the metadata model to comply with the INSPIRE implementing rules (the correspondent INSPIRE descriptors are shown on the right side). Additionally, it must be noted that we need to include in this metadata model all the mandatory elements of ISO 19119 despite the fact that some of them have no equivalent in the set of abstract descriptors contained in the INSPIRE metadata implementing rules.

With respect to the encoding of service metadata in XML format, it must be noticed that we have followed the guidelines established by the technical specification ISO/TS 19139 (ISO, 2007). This technical specification defines the way to translate the UML models proposed in ISO 19115 (and other related standards) into an XML syntax. In particular, the syntax that has been used is the one proposed in the XML-Schemas (http://schemas.opengis.net/iso/19139/20060504/srv/) accompanying the "ISO Metadata Application" profile (Voges et al., 2007) of the OGC catalog services specification, which adopts ISO 19115/19119 as metadata information model.

3.3 Automatic generation of metadata from capabilities

Metadata are the key element to allow the discovery and reusability of services provided by an SDI, however manual metadata creation is a hard and tedious task. In order to make this process easier, this services catalog includes an automatic method to derive metadata from the capabilities information returned by the services that comply with OGC specifications. The responses to *getCapabilities* operations are usually implemented as an XML file that contains three main sections (the name of these sections may vary according to the service type or the capabilities specification version):

Figure 2: ISO 19119 elements for service identification in compliance with INSPIRE metadata implementing rules, extracted from (EC, 2008b).

```
+ citation [1] : CI_Citation
   + title [1] : CharacterString ..........................................................Resource title (See 2.2.1)
   + date [0..*] : CI_Date ...................................................See note 1
      + date [1] : Date ........................................................Date of publication (See 2.6.2)
      + dateType [1] : CI_DateTypeCode ........................................publication
   + date [0..1] : CI_Date.....................................................See notes 1 and 2
      + date [1] : Date ........................................................Date of last revision (See 2.6.3)
      + dateType [1] : CI_DateTypeCode ........................................revision
   + date [0..1] : CI_Date.....................................................See Note 1 and 3
      + date [1] : Date ........................................................Date of creation (See 2.6.4)
      + dateType [1] : CI_DateTypeCode ........................................creation
+ abstract [1] : CharacterString ....................................................Resource abstract (See 2.2.2)
+ pointOfContact [1..*] : CI_ResponsibleParty .............................See 3.5.1
+ descriptiveKeywords [1..*] : MD_Keywords
   + keyword [1..*] : CharacterString..............................................Keyword value (See 2.4.1)
   + thesaurusName [0..1] : CI_Citation.......................................Originating controlled vocabulary (See 2.4.2)
+ resourceConstraints [1..*] : MD_Constraints...............................See 3.6
+ serviceType [1] : GenericName...................................................Spatial data service type(See 2.3.2)
+ couplingType [1] : SV_CouplingType...........................................Mandated by ISO 19119. See Note 8
+ containsOperations [1..*] :  SV_OperationMetadata ....................Mandated by ISO 19119
   + operationName [1] : CharacterString ........................................Mandated by ISO 19119. Default value is unknown
   + DCP [1..*] : DCPList...........................................................Mandated by ISO 19119. Default value is WebServices
   + connectPoint [1..*] : CI_OnlineResource ..................................Mandated by ISO 19119.
      + linkage [1] : URL ........................................................Mandated by ISO 19119. See Note 5
+ extent [1..*] : EX_Extent .........................................................See Note 4
   + geographicElement [1..*] : EX_GeographicBoundingBox ........Geographic bounding box (See 2.5.1)
      + westBoundLongitude [1] : Decimal
      + eastBoundLongitude [1] : Decimal
      + southBoundLatitude [1] : Decimal
      + northBoundLatitude [1] : Decimal
   + temporalElement [0..*] : EX_TemporalExtent...........................See Note 6
      + extent [1] : TM_Primitive....................................................Temporal extent (See 2.6.1)
+ operatesOn [0..*] : MD_DataIdentification...................................Coupled resource (See 2.2.6 and Note 7)
```

– *ServiceIdentification* (or Service in WMS specification): General metadata including identification information such as the title, the abstract or the keywords.
– *ServiceProvider* (or *ContactInformation* in WMS specification): Metadata about the organisation that provides the specific service instance.
– *OperationsMetadata* (or *Capability* in WMS specification): Metadata about the operations and related abilities specified by the service.

The automated method proposed is able to obtain metadata from services compliant with five OGC service specifications: WMS (Web Map Service), CSW (Catalog Services for the Web), WFS (Web Feature Service), WCS (Web Coverage Service) and WPS (Web Processing Service). This automated method is based on a mapping between the capabilities defined in these specifications and the set of descriptors defined in the INSPIRE metadata implementing rules. This mapping can is shown in Table 1.

As shown in Table 1, the *getCapabilities* response of a WMS has a different structure from the capabilities in contrast to other OGC services. However, there is an easy mapping between this specific schema and the more uniform schema used in other OGC specifications, which is defined in (Whiteside, 2007).

Table 1: Mapping between OGC Capabilities Specifications and INSPIRE metadata profile.

INSPIRE metadata	WMS 1.3.0	CSW 2.0.2	WFS 1.1.0	WCS 1.1.0	WPS 1.0.0
Resource Title	Service > Title	ServiceIdentification> Title	ServiceIdentification> Title	ServiceIdentification> Title	ServiceIdentification> Title
Resource Abstract	Service > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract	ServiceIdentification > Abstract
Responsible organisation	Service > ContactInformation	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact	ServiceProvider > ServiceContact
Keyword	Service > KeywordList	ServiceIdentification > Keywords	ServiceIdentification > Keywords	ServiceIdentification > Keywords	ServiceIdentification > Keywords
	WMS	*CSW*	*WFS*	*WCS*	*WPS*
	infoMapAccessService	*humanCatalogueViewer*	*infoFeatureAccessService*	*infoCoverageAccessService*	
Conditions for access and use	Service > AccessConstraints	ServiceIdentification> AccessConstraints	ServiceIdentification > AccessConstraints	ServiceIdentification > AccessConstraints	ServiceIdentification > AccessConstraints
Spatial data service type	*view*	*discovery*	*download*	*download*	*transformation*
Resource locator	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url	Created from the getCapabilities url
containsOperations *	Capability > Request	OperationsMetadata > Operation	OperationsMetadata > Operation	OperationsMetadata > Operation	OperationsMetadata > Operation

*: Only in INSPIRE metadata implementing rules based on ISO 19115 and ISO 19119

Additionally, it must be noted that, apart from extracting some information from a *get-Capabilities* response, this automatic method can also infer some information which is not directly present in a capabilities response. It adds three elements for the classification of services: the spatial data service type according to the service types in annex D.3 of the INSPIRE metadata implementing rules; a keyword to identify the service type according to the ISO 19119 classification of geographic information services; and another keyword corresponding to the acronym of the OGC specification. For instance, in the case of a WMS, this method will fill: the *spatial data service type* element with the value *"view"*; and two keywords with the values *"infoMapAccessService"* (according to the ISO 19119 classification) and *"WMS"* (according to OGC classification).

4. DEPLOYMENT OF THE SERVICES CATALOG IN THE IDEE

According to the architecture described in the previous section a first prototype of the IDEE Services Catalog has been developed (http://www.idee.es/IDEE-ServicesSearch/ServicesSearch.html?locale=en). This prototype allows search and access to the description of services subscribed to the IDEE. Before this prototype, this description of services was updated manually within the IDEE services directory (http://www.idee.es/show.do?to=pideep_catalogo.EN), a set of static Web pages to report the services offered by the IDEE member organisations. Additionally this prototype facilitates the online connection with OGC Web Map Services (WMS).

Figure 3 shows a screenshot of the prototype. With respect to the construction of search requests, the query specification form (see left part of Figure 3) allows to establish different restriction criteria such as the geographic extension of the data provided by the service, the keywords, the service type, or the services provider. Once the results have been presented, the application enables the online edition of metadata items (see Figure 4). Besides, an important feature of the application is that it allows the automated generation of metadata for services compliant with OGC specifications by

means of their *capabilities* information. This functionality is provided when editing the information of a new service (see *New Service* button in Figure 4). Finally, it must be also noted that the online connection with the services returned by the catalog can be activated through the map icon in the list of results. For instance, Figure 3 (right part) shows how an OGC WMS Client is launched to add the new layers served by the WMS discovered through the catalog.

Figure 3: Search criteria, result presentation and online connection with the services.
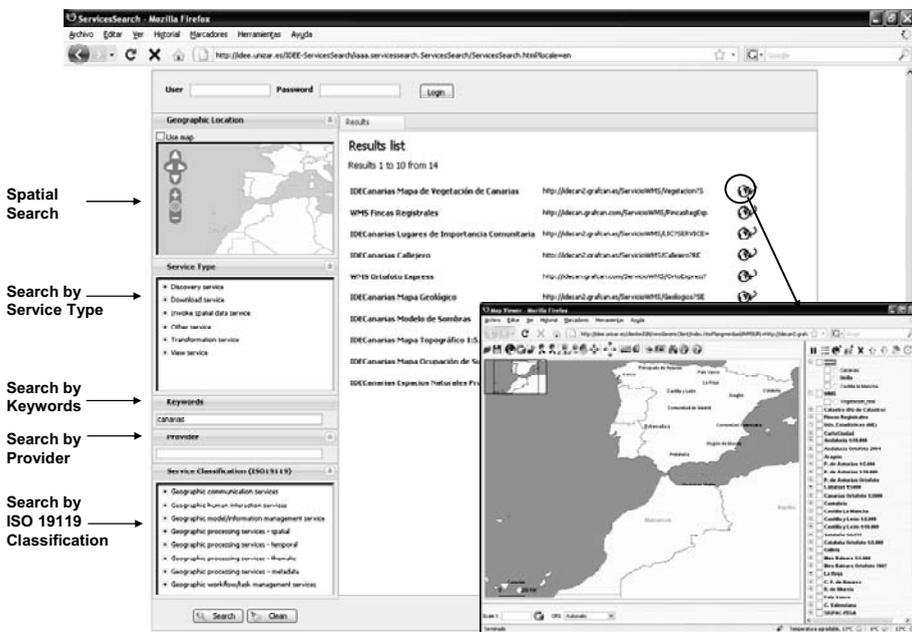
As regards the contents that are accessible through this prototype, we must mention that we have established a process to compile all the service URLs contained in the original static directory (335 services altogether) and apply the automatic method described in section 3.3 to convert the *getCapabilities* response into a metadata record compliant with INSPIRE and ISO 19115/19119 metadata models. Table 2 shows some statistics derived from this upload process.

Table 2: Statistics about the upload of contents to the services catalog (Nr = number).

						Subscribed services							
						Automatic creation				Manual creation		Duplicated cases	
						Correct		Manual review					
Service category	*Service*	*Nr_ss*	*% (=Nr_ss/ tot_ss)*	*Nr_ac*	*% (=Nr_ac/ Nr_ss)*	*Nr_acc*	*% (=Nr_acc /Nr_ac)*	*Nr_acr*	*% (=Nr_acr /Nr_ac)*	*Nr_mc*	*% (=Nr_mc /Nr_ss)*	*Nr_dc*	*% (=Nr_dc/ Nr_ss)*
Discovery	CSW	5	1,5	1	20,0	0	0,0	1	100,0	4	80,0	0	0,0
View	WMS	300	89,6	238	79,3	200	84,0	38	16,0	40	13,3	22	7,3
	WCS	15	4,5	3	20,0	0	0,0	3	100,0	12	80,0	0	0,0
Download	WFS	29 14	8,7 4,2	15 12	51,7 85,7	7 7	46,7 58,3	8 5	53,3 41,7	14 2	48,3 14,3	0 0	0,0 0,0
Transfor-mation	WPS	1	0,3	0	0,0	0	0,0	0	0,0	1	100,0	0	0,0
Invoke		0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
TOTAL				Correct/manual=		207	81,5	47	18,5				
auto./manual/dupl=		335 (=tot_ss)	100,0	254	75,8					59	17,61	22	6,57

Figure 4: Online metadata edition.
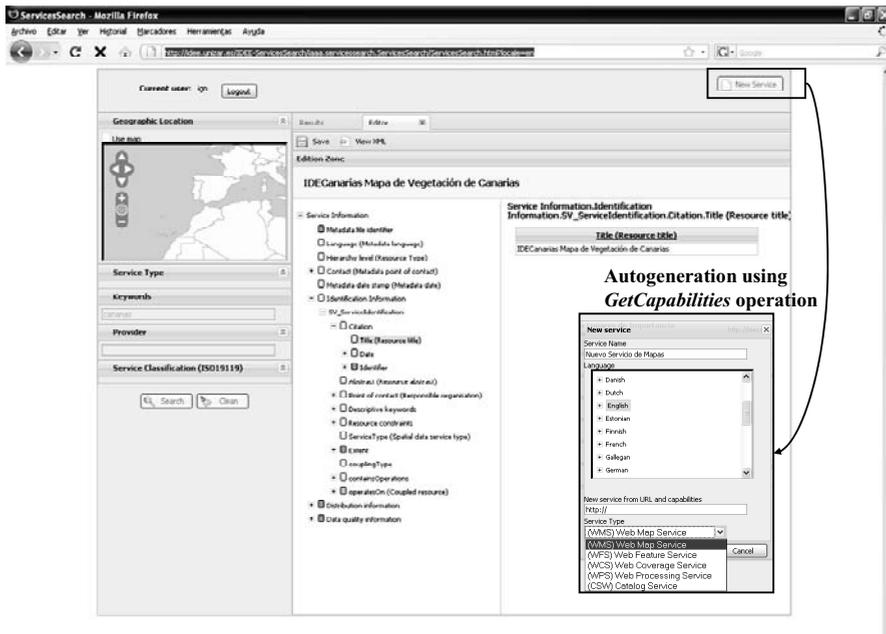
The services subscribed to the IDEE have been grouped in Table 2 according to the spatial data service type defined in annex D.3 of the INSPIRE metadata implementing rules (EC, 2008a): *Discovery*, *View*, *Download*, *Transformation* and *Invoke*. And within a category, there is specific row for each OGC compliant service in this category: services compliant with CSW (Catalogue Service for the Web) protocol binding of the OGC Catalogue Service Specifications within the *Discovery* category; services compliant with the OGC WMS (Web Map Service) specification within the *View* category; services compliant with the OGC WCS (Web Coverage Service) and WFS (Web Feature Service) specifications within the *Download* category; and services compliant with the OGC WPS (Web Processing Service) specification within the *Transformation* category. For each category and OGC specification the table shows statistics about the number and percentage of the services subscribed to the catalog including details about the metadata creation process:

– Column *Nr_ac* (and its associated percentage) under the *"Automatic creation"* header counts the number of records that were automatically created from the *capabilities* responses. And within this category it is possible to distinguish two subcategories: number of records that were generated without any problem (see columns under "*Correct"* header); and number of records that required some kind of manual review (see columns under *"Manual review"* header) due to old versions or capabilities or problems in XML files (incomplete or invalid files). Although not included in the table, it can be derived that 61.8% of records were automatically created without manual review (207 records out of 335 services).

– Column *Nr_mc* (and its associated percentage) under the *"Manual creation"* header counts the number of records that were manually created. This column represents the number of service whose URL was invalid. In this case, metadata records were filled manually with the information stored in the static directory.

– Column *Nr_dc* (and its associated percentage) under the *"Duplicate cases"* header counts the number of records that had been introduced twice.

As a conclusion from the analysis of the figures in this table, we have detected some typical problems to solve while working with services catalogs. Firstly, information becomes rapidly out-of-date: 17.61% of service URLs included in the service directory was not valid anymore. Secondly, interoperability problems between different versions of OGC services persist: 18.5 % of automatically created metadata records were reviewed manually because their capabilities response could not be parsed. Thirdly, there is no control about duplicates: 6.57% of original service descriptions had been introduced twice (but with slight differences). The advantage of using a services catalog with respect to a static directory is that it is easier to establish an automatic mechanism to report these problems.

5. CONCLUSIONS AND FUTURE WORK

This work has presented the development and deployment of a first prototype of Services Catalog in the SDI of Spain, replacing the static content offered until now by IDEE under the name of services directory. Thanks to this new prototype, it is possible to provide online services such as metadata creation, dynamic indexing and metadata search, as well as the connection with generic clients, which allow verifying the correct operation of services.

Thanks to the upload of metadata in the services catalog using the content of the original static service directory, we have detected some problems that could be common to other SDI initiatives: information becomes rapidly out-of-date (it is difficult to maintain service URLs up-to-date); interoperability problems between different versions of OGC services persist (capabilities responses are still very heterogeneous); and there is little control about duplicates. Although the use of a services catalog does not solve these problems directly, it is easier to establish an automatic mechanism to report these problems and mark the records that are not valid at the moment. Additionally, we must remark the low quality, in general, of the metadata that are usually found as a response to the *getCapabilities* request of OGC-compliant services. For instance, most of these metadata do not even contain a citation to the organisation that has set up the service on the internet.

As further lines of this work, we plan to integrate more generic clients to facilitate the online connection with other service types apart from WMS services. Currently, the connection with OGC-compliant catalogs through generic CSW clients is being developed. With respect to the automatic metadata generation, we also plan to infer more metadata elements from the information contained in a capabilities response. For instance, we expect to infer the extent of a WMS from the bounding box of the map layers provided by this service. Last, new functionalities of the application have been proposed to facilitate the monitoring of the services registered in the catalog.

In summary, we think that Services Catalogs will constitute a key element for SDI development. They are essential to facilitate the reusability of services, and in the future, they will probably receive more requests than the data catalogs. We must not forget that if SDIs are based on a Service Oriented Architecture, the services are the basic concept around which an SDI will be conceived and structured. Therefore, the Services Catalogs will play the main role in the exploitation of SDIs.

REFERENCES

Booth D., Haas H., McCabe F., Newcomer E., Champion M., Ferris C. and D. Orchard (eds.) (2004). *Web Services Architecture. W3C, W3C Working Group Note 11 February 2004, at:* http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/.

Doyle A. and C. Reed (ed.) (2001). *Introduction to OGC Web Services*. An OGC® White Paper, Open Geospatial Consortium Inc*., at:* http://portal.opengeospatial. org/files/?artifact_id=14973.

EC (2007). *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*. European Commission (EC), Official Journal of the European Union*, at:* http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L: 2007:108:0001:0014:EN:PDF.

EC (2008a). *Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (Text with EEA relevance), at:* http://eur-lex.europa.eu/LexUri Serv/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF.

EC (2008b). *INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119, v. 1.0*. European Commission (EC), Joint Research Centre*, at:* http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/meta data/MD_IR_and_ISO_20081219.pdf.

Haas H. and A. Brown (eds.) (2004). *Web Services Glossary. W3C, W3C Working Group Note 11 February 2004, at:* http://www.w3.org/TR/2004/NOTE-ws-gloss-20040211/.

ISO (2003a). *Geographic information - Metadata. ISO 19115:2003*. International Organization for Standardization (ISO).

ISO (2003b). *Information and documentation - The Dublin Core metadata element set. ISO 15836:2003.* International Organization for Standardization (ISO).

ISO (2005). *Geographic information - Services. ISO 19119:2005.* International Organization for Standardization (ISO).

ISO (2007). *Geographic information - Metadata - XML schema implementation. ISO/TS 19139:2007*. International Organization for Standardization (ISO).

Lieberman J. (2003). *OpenGIS Web Services Architecture. OpenGIS Discussion Paper OGC 03-025*. Open Geospatial Consortium Inc*., at:* http://portal.opengeospatial. org/files/?artifact_id=1320.

Martell R. (ed.) (2007). *CSW-ebRIM Registry Service - Part 1: ebRIM profile of CSW (Version: 1.0.0)*. OGC 07-110r2, Open Geospatial Consortium Inc*., at:* http:// portal.opengeospatial.org/files/?artifact_id=27092.

Nebert D. (ed.) (2004). *Developing Spatial Data Infrastructures: The SDI Cookbook* v.2.0. Global Spatial Data Infrastructure*, at:* http://www.gsdi.org.

Nebert D., Whiteside A. and P. Vretanos (eds.) (2007). *OpenGIS - Catalogue Services Specification (version: 2.0.2)*. OGC 07-006r1, Open Geospatial Consortium Inc., *at:* http://portal.opengeospatial.org/files/?artifact_id=20555.

Nogueras-Iso J., Barrera J., Gracia-Crespo F., Laiglesia S. and P.R. Muro-Medrano (2008). "Integrating catalog and GIS tools: access to resources from CatMDEdit thanks to gvSIG*". Proceedings 4th International gvSIG conference*. 3-5 December 2008, Valencia*, at:* http://www.jornadasgvsig.gva.es/fileadmin/consel leria/Documentacion/4asJornadas/articulos/Sesion3/CatMDEdit_gvSIG.pdf.

Tolosana-Calasanz R., Portolés-Rodríguez D., Nogueras-Iso J., Muro-Medrano P.R. and F.J. Zarazaga-Soria (2005). "CatServer: a server of GATOS"*. Proceedings of AGILE Conference 2005, Estoril, Portugal,* pp. 359–366 ISBN 972-8093-13-6.

Voges U. and K. Senkler (2007). *OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile*. OGC 07-045, Open Geospatial Consortium Inc.*, at:* http://portal.opengeospatial.org/files/?artifact_id=21460.

Whiteside A. (ed.) (2007). *OGC Web Services Common Specification (Version: 1.1.0 with Corrigendum 1).* OGC 06-121r3. Open Geospatial Consortium Inc.*, at:* http:// portal.opengeospatial.org/files/?artifact_id=20040.

Zarazaga-Soria F.J., Lacasta J., Nogueras-Iso J., Torres M.P. and P.R. Muro-Medrano (2003). "A Java Tool for Creating ISO/FGDC Geographic Metadata*". In: Geodaten- und Geodienste-Infrastrukturen - von der Forschung zur praktischen Anwendung. Beiträge zu den Münsteraner GI-Tagen* 26./27. Juni 2003. Volume 18, Münster, Germany, IFGIprints, pp. 17-30.

Public Sector Geo Web Services: Which Business Model Will Pay for a Free Lunch?

Frederika Welle Donker

OTB Research Institute for Housing, Urban & Mobility Studies
Delft University of Technology, the Netherlands
f.m.welledonker@tudelft.nl

Abstract
Geo-information (GI) is increasingly having a bigger impact on socio-economic bene-fits. Over the last decade, use of GI has shifted from a specialised GIS niche market to serving as a direct input to planning and decision-making, public policy, environmental management, readiness to deal with emergencies, creation of value added products, citizen mobility and participation, and community platforms. The emergence of Google Earth and Google Maps has created a geo-awareness and has catalysed a thirst for custom-made geo-information. Governments possess, often high-quality large-scale GI, primarily created, collected, developed and maintained to support their public tasks. This rich source of GI begs to be used and reused both within the public sector and by society. Both the INSPIRE Directive (2007/2/EC) and the Directive on reuse on Public Sector Information - the so-called PSI Directive - (2003/98/EC) underwrite the philoso-phy of 'collect once, reuse many times'. Web services are an effective way to make public sector geo-information available. They allow information to be accessed directly at the source and to be combined from different sources. However, the costs of web services are high and revenues do not always cover the costs. Assuming that there is no such thing as a free lunch related to public sector GI (Longhorn and Blakemore, 2008), which business models and which financial models form the basis for public sector geo web services? This article explores the different models currently in use and illustrates them with examples.

Keywords: geographic information, public sector web services, business models, fi-nancial models, revenue models.

1. INTRODUCTION

The terms 'geographic information', 'geographic data', 'spatial information' and 'spatial data' are interchangeably used as synonyms. For the purpose of this article, only the term geographic information (GI) will be used. Access to GI is of vital importance to the economic and social development of the nation. Nations around the world are develop-ing geographic information infrastructures (GIIs), also referred to as spatial data infra-structures (SDIs), with access to GI at the core. For more advanced GIIs (re)use is considered the driver of a GII (van Loenen, 2006). One way to facilitate reuse of GI is through web services. The INSPIRE Directive even requires that as part of developing geo-information infrastructures network services should be used. National GIIs are now evolving from first to second generation GIIs. The existence of web services are re-garded as the main technological drivers of second generation GIIs because they can fulfil the needs of users and improve the use of data (Crompvoets et al., 2004; Rajabi-fard et al., 2003). This article will give an inventory of the different models currently in use and illustrate them with examples. In section 2 a description of various types of web services will be provided, including a case study illustrating costs involved setting up a commercial Web 2.0 platform and the potential revenue web services can gener-ate. Section 3 will supply a theoretic framework for business models with a breakdown

of the four essential parts of a successful business model. Section 4 will build on the business model framework with a framework for financial models, including various cost and revenue models and price strategies. In section 5, the summary will show which business model and which financial model will be most suited and robust in a given situation. It will also show some current pricing trends for public sector geographic information (PSGI) in Europe. Section 6 will finish with conclusions and offer recommendations for public sector web services.

2. WEB SERVICES

2.1 Different web services

A web service is a platform that is accessible with open standard protocols such as SOAP and XML. A web services sends a request from the client-computer to a server. The server sends queries to the appropriate source servers and transmits a reply back to the client-computer. The advantage is that data is queried at the original source so it is as current as possible. There are a number of different types of GI web services, which roughly fall into two categories: web services using Open Geo Consortium (OGC) standards and web services using ICT standards.

2.1.1 Open Geo Consortium web services

The main OGC standards used for web services are Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) and Web Integrator Service (WIS). WMS only produces a static image on screen from raster files. Because no actual data is transferred, no information can be downloaded. Therefore, it is easier to comply with protection of intellectual property rights (IPR). With WFS, discrete features (points, lines, polygons) are downloaded in XML to the client-computer. The same applies to a WCS whereby entire coverages (sets of features) are also downloaded in XML. Data from WFS and WCS are suitable for interpretation, extrapolation and other forms of analysis. Because the data itself is transferred from the server, measures to protect data subject to IPR are harder to implement for WFS and WCS than for WMS whereby no data is transferred. Therefore, WFS and WCS are probably more suitable for fee-based web services. A WIS is a service that can horizontally integrate various WMSs. Horizontal integration of WMS means that different WMSs of different organisations are bundled into one new WMS. A WIS allows for instance to integrate all regional WMSs containing planning information to be bundled into one national WMS for planning information. To the end-user, the WIS will appear as one web service (see Figure 1).

WMSs are very popular for 'free' web services as they only produce a static image in a low-resolution format (e.g. jpg, pdf) that allows little to no editing. Often images generated from WMS are embedded into other services such as online route planners or community platforms. However, the images contain an attribution label as part of copyright requirements. If a map is generated from more than one WMS or from a WIS, multiple attribution labels will appear on the image, which may hamper legibility of the image (see Figure 2). In the Netherlands WMSs are the most popular web services used by both the public sector and the private sector. From interviews held for this research, it appeared that to date there is little demand yet for WFSs and WCSs. There are a few WFSs available, which are mainly used within the public sector and by specialised private sector companies such as engineering firms. However, the lack of demand for WFSs/WCSs in the Netherlands may be explained by the fact that potential users of these geo web services may be unaware these web services exist.

Figure 1: Serving geo-information using WMS, WFS and WIS (Geoloketten, 2008).
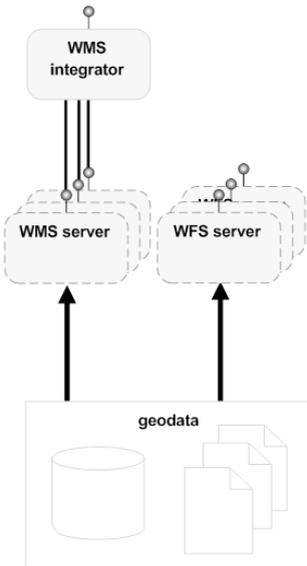
Figure 2: Several source attributions per map image (Bibber, 2008).
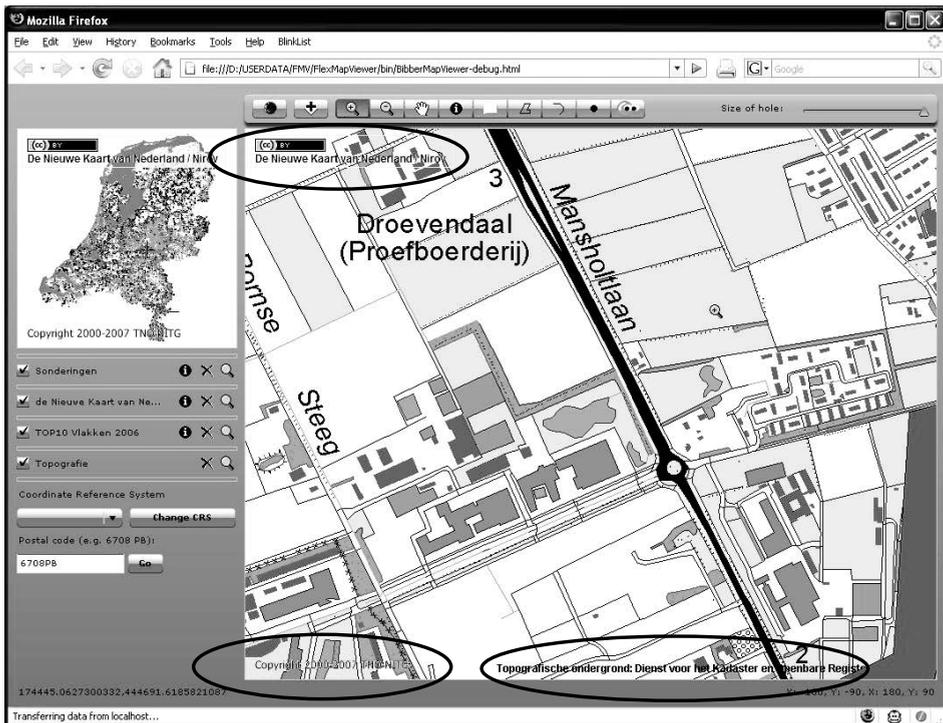
2.1.2 ICT standards web services

For geo web services ICT standards such as SOAP, are actually used more often than OGC standards. The most popular type of web service is a Data Service (DS). The private sector uses DSs because custom-made information is delivered to the client. Furthermore, a DS can combine geo-information with data from other databases. Query tools can then be used to perform analyses on the metadata. Licensed information can be protected with firewalls, although the same firewalls can make it harder to set up query tools. Apart from DSs, there are also Sensor Web Services and Simulation Models. Sensor Web Service is a type of sensor network consisting of spatially distributed sensor platforms that wirelessly communicate with each other. They are most often deployed for environmental monitoring and control. For this research, all ICT standard web services will be bundled into Data Services.

Although the technical specifications and standards used for the various types of web services are different, the economic aspects of them are not so dissimilar. In this article, no distinction will be made between the different types of web services when describing the economic aspects.

2.2. Costs of web services

The costs of setting up and keeping a web service operational are high. To develop a web service one has to invest in hardware, software, legal, technical, sociological and economic expertise, building up know-how, market and target group research, implementation costs, advertising and promotional costs, administrative and project management costs. Then there are the operational expenses such as servers, broadband capacity, licence fees for software and/or (geo) datasets, acquisition costs and personnel costs. During the operational phase of a web service reservations have to be made for future costs such as R&D, equipment depreciation and extra capacity.

The costs of an operational web service are very variable, depending on the type of service. Stieglitz et al. (2008) made a financial analysis of a virtual community as part of a case study. Virtual communities are a group of people sharing a common interest by using internet applications. Web 2.0 platforms are technologies, which enable formation of virtual communities. An increasing number of private sector organisations are using virtual communities to bridge the gap between users and the organisation by including users in the value chain. The financial analysis undertaken by Stieglitz et al. (2008) was conducted for a virtual community of retail investors at the Berlin Stock Market with memberships sold on a subscription base. Stieglitz et al. (2008) distinguish four separate phases in the life of a web service:

(1) the development phase (analysis, design and implementation);
(2) the operational phase;
(3) the adaptation phase (evaluation and evolution), and
(4) the disintegration phase.

Even in the disintegration phase, the web service still incurs costs such as migration costs to another platform, running contract costs and replacement of technology. Only in the operational phase is revenue raised through savings, advertisements and memberships / subscriptions. In their analysis, Stieglitz et al. (2008) noted that the total costs per month were relatively stable during the first year of the operational phase. Only after a critical mass of users and contributions is reached, growth can accelerate. Later in the operational phase, the costs will continue to increase but so will the reve-

nue. With an increasing number of members, the cost per member will decrease until it approaches zero. However, when the number of active members reaches a certain level, the operational costs will step up because of the required extra capacity (servers, broadband, personnel). In addition, this specific virtual community is still in the operational phase. In later phases (adaptation and disintegration), the cost per member will probably increase again.

Although this case study applied to a commercial virtual community, the same principles apply to geo web services. From the various interviews held for this research, the biggest cost item mentioned is sufficient broadband capacity to keep the service operational at all times. Especially for WMS the required server and broadband capacity can be huge if there are many simultaneous users. In addition, it can take some time for an image to build up on the screen of the client-computer. If the build-up time is too slow, the user will abandon the web service. To save building-up time, images can be stored as tiles on the server(s) in advance. However, for large-scale information sets Terabytes of storage capacity is required. Geoportail, the French NGII web service requires 3 Gbps broadband capacity, two 50 Tb caches and a 100 Tb storage capacity (Richard, 2008).

2.3 Web service revenue

Web services are set up by the public sector for several reasons: to share information with other public sector organisations, to inform citizens and the private sector (with or without a legal obligation to do so), or as a way to market public sector information (PSI) for reuse. PS(G)I forms a rich resource for value added resellers (VARs) to create value added products and services. Because the public sector enjoys scale of economies and scales of scope, the costs are relatively low. The benefits may be financial for fee-based services or increased taxation revenue from VARs; or the benefits may be intangible such as a better-informed citizen or increased policy effectiveness. As intangible benefits are harder to measure, cost-benefit analyses tend to be negative. However, end-users of information also incur costs if information needed is scattered all around. These lost productivity costs can be significant when someone has to spend hours searching the Internet for useful information (Bates and Andersen, 2002). The savings made in search costs should be included in cost-benefit analyses when setting up web services for internal use.

3. BUSINESS MODELS

There are many definitions for the concept of business models. Rappa (2003) offers perhaps one of the simplest definitions, that a business model the method is of doing business by which a company can sustain itself - that is, generate revenue. A business model describes the strategies implemented to achieve a goal. A financial model is an essential part of a business model. The financial model describes the cost framework and how revenue will be generated. The simplest business model is producing and selling a good to customers with revenue higher than all costs incurred. Poorly worked out business models and financial models were one of the main causes of the demise of the dot-com companies at the end of the last century (see e.g. Razi et al., 2004).

3.1 Components of a business model

After a comparison of different business model definitions, Bouwman et al. (2008) distinguish four components of a successful business model, namely Service, Technology, Organisation and Finance. Together these components form the so-called STOF-

model (see Figure 3). The four components should be addressed in balance with each other. The starting point is the service domain which addresses aspects such as type of service, intended user group and the value of a service for meeting customer demands. The service domain serves as a guide to the technical design. Some of the aspects addressed in the technical design are architecture, infrastructure, accessibility and payment mechanisms. To develop and market a successful service often requires organisations to collaborate. Collaboration can be as simple as one organisation wanting to launch a web service and needing financial backing from a bank or it can be different organisations bundling information into one web service. The organisation domain describes the value chain required to realise a specific service. A value chain consists of actors with specific resources and capabilities that interact and work together to create value for customers and to realise their own strategies and goals (Faber et al., 2008). The organisation domain has to address the network and actor aspects as well. The last component to be addressed is the finance domain, which is the bottom line of any business model with revenues on one side and investments, costs and risks on the other side.

Figure 3: STOF model (Bouwman et al., 2005).
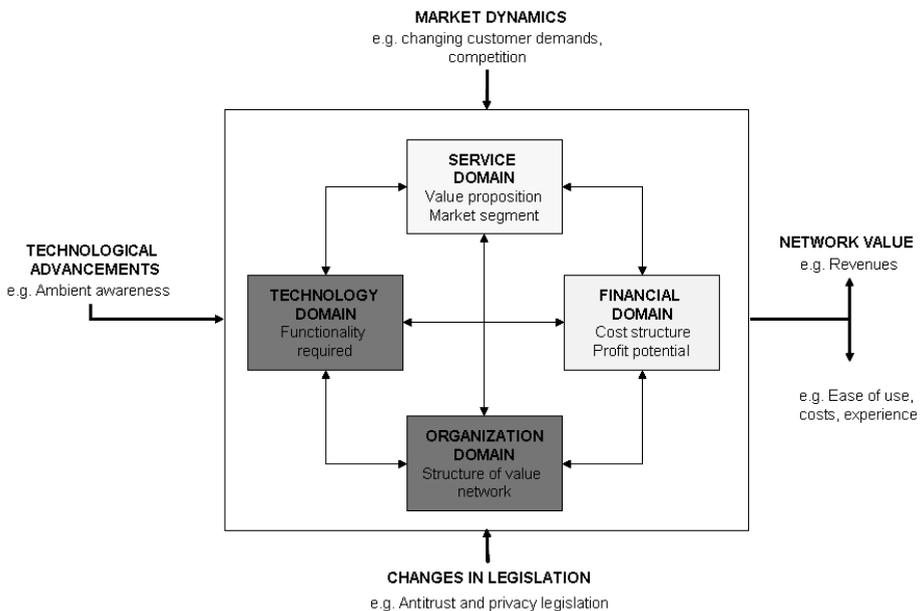
Osterwalder and Pigneur (2002) note that a business model can only be successful if it includes the following three elements: (1) revenue and product aspects; (2) business actor and network aspects; and (3) marketing specific aspects. In their view, a business model should be based on the following columns:

– Product innovation;
– Customer relations;
– Infrastructure management, and
– Finances.

3.2 Types of business models

Malone et al. (2006) designed a simple diagram of 16 types of business models based on two dimensions. The first dimension looks at the type of asset right being sold. These are:

1. a *Creator* buys raw materials or components from suppliers and then transforms or assembles them to create a product sold to buyers;
2. a *Distributor* buys a product and resells essentially the same product to someone else;
3. a *Landlord* sells the right to use, but not own, an asset for a specified period of time;
4. a *Broker* facilitates sales by matching potential buyers and sellers. Unlike a typical Distributor, a Broker does not take ownership of the product being sold, rather only receives a fee from the buyer, the seller, or both.

The second dimension takes into account the type of asset for which rights are being sold. These types are physical (durable goods), financial (e.g. cash, insurance), intangible (e.g. copyrights, knowledge, goodwill), and human (people's time, effort). Combining these dimensions offers the following 16 business models, although effectively there are only 14 as two (human creation and human trade i.e. slavery) will be illegal in most countries.

Table 1: Schema of 16 types of business models (after Malone et al., 2006).

	Creator	Distributor	Landlord	Broker
Physical	Manufacturer	Wholesaler / retailer	Leaser (e.g. real estate)	Auctioneer (e.g. eBay)
Financial	Entrepreneur	Bank, investment firm	Lender / insurer	Insurance broker
Intangible	Inventor	Intellectual property trader	Publisher / brand manager / attractor (e.g. Google)	Intellectual property broker
Human	*Human creation*	*Slavery*	Contractor	Human resources broker

Since information is a physical good, only the business models on the top row are applicable to GI suppliers. GI suppliers are often both 'Manufacturer' as well as 'Leaser' because apart from producing GI, they often only sell the right to use the product rather than transfer ownership. There are some public business organisations trading as 'Broker', such as DataLand brokering municipal GI in the Netherlands. However, most of these organisations also trade as 'Leaser' and the brokerage is often only a secondary business activity. Hence, in this article they are included in the 'Leaser' category. The schema of viable business models can be adapted now for GI suppliers illustrated in Table 2.

Table 2: Schema of viable business models for GI-suppliers (gray) (after Malone et al., 2006).

	Creator	Distributor	Landlord	Broker
Physical	*Manufacturer*	Wholesaler / retailer	*Leaser*	Auctioneer
Financial	Entrepreneur	Bank, investment firm	Lender / insurer	Insurance broker
Intangible	Inventor	Intellectual property trader	Publisher / brand manager / attractor	Intellectual property broker
Human	*Human creation*	*Slavery*	Contractor	Human resources broker

4. FINANCIAL MODELS

4.1 Cost models

Financial models consist of two components: cost models and revenue models. The cost model describes which costs an organisation incurs to run a business. The revenue model describes how an organisation expects to generate income. For public sector organisations supplying PSGI there are two cost model regimes: marginal costs regime and cost recovery regime. With the marginal costs regime only costs of dissemination are taken into account, e.g. cost of a DVD or actual time taken to produce a copy. For web services, the marginal costs are zero if the operational costs of the web service are deemed part of supplying a public service. With the cost recovery regime, all costs that are made by the organisation to create, collect, process and maintain the information are included in calculating the dissemination costs. The PSI Directive even allows a reasonable return on investment.

4.2 Revenue models

All organisations, including public sector organisations, will have to employ a Revenue Model for PSGI web services. In the literature, many revenue models are described. Rappa (2003) distinguishes nine different categories of revenue models. These categories are listed in Table 3.

Table 3: Categories of revenue models (after Rappa, 2003).

Revenue model	Description
Brokerage model	Brokers bring buyers and sellers together and facilitate transactions, usually for a fee or commission
Advertising model	The web site provider provides content (usually, but not necessarily, for free) and services (such as email or blogs) mixed with advertising messages in the form of banner ads.
Infomediary model	Infomediaries collect information, e.g. information about consumers and their consumption habits, or information about producers and their products useful to consumers when considering a purchase. The infomediary then acts as an information intermediary.
Merchant model	Wholesalers and retailers of goods and services. Sales may be made based on list prices or through auction.
Manufacturer (direct) model	The manufacturer or 'direct model' allows a manufacturer to reach buyers directly and thereby compress the distribution channel.
Affiliate model	The affiliate model offers financial incentives (in the form of a percentage of revenue) to affiliated partner sites. The affiliates provide purchase-point click-through to the merchant. It is a pay-for-performance model - if an affiliate does not generate sales, it represents no cost to the merchant.
Community model	The viability of the community model is based on user loyalty. Users have a high investment in both time and emotion. Revenue can be based on the sale of ancillary products and services or voluntary contributions; or revenue may be tied to contextual advertising and subscriptions for premium services.
Subscription model	Users are charged a periodic fee to subscribe to a service. It is not uncommon for sites to combine free content with 'premium' (i.e., subscriber- or member-only) content. Subscription fees are incurred irrespective of actual usage rates.
Utility model	The utility or 'on-demand' model is based on metering usage, or a 'pay as you go' approach. Unlike subscriber services, metered services are based on actual usage rates.

Not all revenue models described by Rappa are suitable to PSGI web services, such as the Brokerage, Advertisement, Infomediairy and Merchant Model. In addition, the term 'Usage Model' may be a better description of the model than the term 'Utility Model'. Public sector organisations with a Marginal Costs regime will not need to charge for their web services at all. Therefore, some extra models are added to the list, including some revenue models out of the creative sector. As most public sector organisations are holders of (semi-)monopolistic data, they employ the Manufacturer Model by definition, therefore this model is further omitted. When the viable business models for PSGI suppliers (see Table 2) of Malone et al. (2006) are combined with the adapted revenue models of Rappa, the following revenue models appear:

1. *Subscription model:* Revenue is raised through periodic fees. This is a popular model for supplying access to a service that is frequently used, e.g. iTunes. The advantage for the web service provider is that revenue is raised in advance and thus providing more certainty of regular income. The advantage for the user is that costs of accessing information are known in advance and access is unlimited within the subscription limit. A disadvantage is that both research and practice show that consumers are reluctant to pay for online services (Schiff, 2003), unless there is a direct relation with their private lives (Reitsma, 2007). Sometimes a basic subscription is offered for free and versions with more features attract a fee (e.g. Google Earth for free, Google Earth Plus $20/year & Google Earth Pro $400/year). Subscription models are best suited to specialist information, or (semi-)monopolistic information, e.g. large-scale base maps.

2. *Usage Model:* Revenue is raised through actual usage of a service. Usage may be measured in time, per bytes, per area or per session. The web service provider has to be able to cope with small amounts of money. The usage model is best suited to ad hoc users whereby *access* to services is more important than *possession.* In addition, the usage model is only suited to web services with geo-data from only a few suppliers, as the pricing structure will become very complicated and intransparent (MICUS 2003, 2008b). Another disadvantage for geo web services is that charging per hectare or bytes will render large-scale area coverage very expensive.

3. *Royalty model:* Revenue is raised through royalties paid after a value added product has been successfully produced. The price of a service is dependant on the results of the user. The price, the royalty, is usually a fixed percentage of the turnover or the revenue of the value added product of the user. The advantage of this model is that a firm only has to pay for the GI after a value added product is successful so there is room for experimenting. The disadvantage of this model is that contracts have to be signed in advance making this model less suitable to click-through licences. Users of the supplied information have to be monitored. In addition, there is no short-term certainty of income.

4. *Free Model:* There is no direct revenue raised through this model, although there will be indirect benefits. Public sector organisations employ this model, either as a legal obligation or for efficiency reasons (no sales staff). The immediate benefits are intangible, e.g. a better-informed citizen or better policy effectiveness, or the benefits may be financial in the long term, e.g. extra taxes when value added products are created. However, making GI available free of charge may be in breach with national Fair Trade Legislation in some countries as it may be deemed an act of unfair trading practices if the private sector already has made vast investments to create similar services. The creative sector also uses the Free Model to achieve name recognition or for altruistic reasons.

5. *Hybrid models:* These are models showing some of the characteristics of the models described above. Below some of the more common varieties are described.
   a. Enticement model: A part of the content is provided free of charge as a lure to entice the user. Revenue is raised from sale of premium content or other related services. This is one of the oldest revenue models first introduced by King Gilette to create a market for his disposable razor blades (Anderson, 2008). Often cross-subsidising is employed, i.e. content is offered for free and revenue is raised from sale of related products such as merchandising (e.g. free mobile phones with revenue from phone calls/text messages; songs downloadable for free and revenue is raised from sale of concert tickets and/or merchandising).

   b. Community model: The viability of the community model is based on user loyalty. Users invest both time and emotions to produce a communal service. Revenue can be raised by sale of ancillary products and services or by donations; or revenue may be tied to contextual advertising and subscriptions for premium services. The best-known example of a Community is Wikipedia. An example of a GI-community is OpenStreetMap (OSM), a project whereby volunteers go out with GPS units to produce open source street maps for distribution free of charge. OSM operates in many countries on six continents. Some private geo companies have donated cartographic information or money to OSM as well in return for their data or as a platform for innovative applications (http://www.opengeodata.org/?p=223). In Germany, the OSM data were used for the development of a 3D Geodata Infrastructure as part of the research project 'Geodata Infrastructure 3D' (GIM, 2009). The private sector uses frequently virtual communities to involve users in the developmental and evaluation phases of services as the users provide useful feedback and ensure quality control.

   c. Street performer protocol: A protocol popular in the creative domain and with software developers. Under this protocol, a producer will release a work (e.g. a book or software application) into the public domain after a certain amount of money has been received in a trust fund. Interested parties pay their donations to this trust fund, which is managed by a publisher. When the work is released on time, both the producer and the publisher are paid from the trust fund. If the work is not released on time, the publisher repays the donors. In some variations, the product is commercially released on the market rather than into the public domain. The producer will repay a return on investment to the donors when the product makes a profit. This protocol is very dependant on the reputation of the producer. This protocol would also be suitable to screened-off web services whereby the users are known in advance. Once the participants have donated their share of development costs and expected operational costs, the service would then be available to the participants to use.

   d. Combination model: Combinations of the above models are quite often employed, e.g. combining the Royalty Model with a start-up fee. The UK Ordnance Survey uses this model for VARs. Another possible combination would the Enticement Model combined with the Subscription Model, e.g. giving away a small sample of the Cadastral Map to consumers. The Dutch Large Scale Base Map combines the Subscription Model with the Utility Model as well as offering a user right for the entire dataset for a one-off fee. Another Model is the Data-For-Data model whereby different public sector organisations participate in a joint program, with or without paying an upfront contribution. They donate their data into this program to produce large-scale geo-information. In return, the organisations receive user rights for this large-scale geo-information, Norge Digitalt in Norway

uses this model to finance large-scale datasets. The Data-For-Data Model can be combined with the Street Performer Model if a participant donates money instead of data.

4.3 Summary of revenue models

Table 4 provides a summary of the various revenue models, their advantages and disadvantages and their suitability to various web services.

Table 4: Revenue models with pros, cons, and their suitability to web services.

Model	Advantages	Disadvantages	Suitable for
Subscription Model	- Certainty of regular revenue - Adaptable to users - Lock-in of users - Suitable for click-through licences	- Not popular with consumers - Only suitable for specialised data that is required frequently	- WMS - WFS / WCS - DS
Usage Model	- User-pay system, only pay for actual usage - Suitable for ad hoc users - Suitable for click-through licences	- Only suitable when access is more important than possession - Need mechanisms to deal with small payments - Pricing may be prohibitive for large quantities - Pricing mechanism complex when combined with other web services	- WMS - WFS / WCS - DS
Royalty Model	- Suitable for experimentation / innovation platform - Low accessibility - May generate long term indirect revenue for VA products	- Uncertainty of revenue (amount, time) - Must monitor progress of experimenters - No revenue from consumers - Nor suitable for click-through licences	- WMS - WFS / WCS - DS
Free Model	- Low accessibility - Indirect revenue (better informed citizen, more effective policy) - May generate long term indirect revenue for VA products - Suitable for click-through licences (if still required)	- No direct or immediate revenue - May be in breach with national Fair Trade Legislation	- WMS - WFS / WCS - DS
Hybrid Models - Community Model	- User is closely involved (feedback, quality control) - Improvement of service / user friendliness - Encourages experimentation / innovation platform	- No direct or immediate revenue (unless combined with another model)	- WMS - WFS / WCS - DS
- Enticement Model	- Lures potential users - Lock-in of users	- No direct or immediate revenue (unless combined with another model)	- WMS - DS
- Street Performer Model	- Financing service is done upfront - Unlimited use for donors / participants	- Donors / participants must be known and willing to donate in advance - Dependant on good reputation of producer	- WMS - WFS / WCS - DS

4.4 Price strategies

Apart from the Revenue Models described above, price discrimination can be applied as well. The British welfare economist A. C. Pigou described as early as in 1920 a pricing theory, which included price discrimination (Pigou, 1920). Price discrimination can only be applied in a limited fashion by the public sector, as the PSI Directive does not allow that a public sector body distinguishes between different groups of users using the data for similar purposes. It may be possible to offer rural GI cheaper than urban GI because the latter is more dynamic and needs to be updated more frequently (Longhorn and Blakemore, 2008). In addition, there may be more need for urban information, i.e. a larger market segment. Another form of price discrimination that may be applied, is offering volume discounts but the volume price is the same for everybody. An example would be to decrease the unit price per hectare when a larger area is selected, e.g. as applied to the *Automatisierten Liegenschaftkarte* (ALK) in North Rhine Westphalia, Germany. Alternatively, a time-based approach could be employed, e.g. charging a higher fee for more timely weather information products, or charging a lower fee for usage outside normal business hours.

In the last couple of years there appears to be a trend that large scale PSGI is coming down in price, because either it was too expensive for the private sector or the prices created barriers to effective reuse within the public sector. With prices being lowered, the number of (re)users is increasing, so the actual total revenue may even go up. Recent examples are found in Austria, Netherlands and Spain. The Austrian Federal Office of Meteorology and Surveying (BEV) have significantly reduced their fees for their PSGI. For instance, the fee for the cartographic model was reduced by 93% and usage went up by 200-1500%, and the digital cadastral map went down by 97% and usage up by 250%. The majority of new users are small to medium enterprises (Schennach, 2008). In the Netherlands, the so-called New Map of the Netherlands (NMN) has been available online with a Creative Commons licence since January 2006 (see www.nieuwekaart.nl). The NMN offers a complete overview of planned spatial developments and functional changes in the Netherlands. Before the NMN became available free of charge, about 20 datasets were sold. Since then, the number of discrete reusers - both from the public and the private sector - downloading the NMN on a regular basis have stabilised to around 200 (Nirov, 2007). The Spanish Cadastre made the complete cadastral map of Spain available on the internet in March 2003. An analysis of the impact of free access to spatial data in Catalonia demonstrated that such initiative is highly profitable to public institutions, by saving a lot of time, simplifying processes and making optimal use of the available information. The impact on private companies is also positive (MICUS, 2008a).

5. SUMMARY OF BUSINESS MODELS

Since the development and operational costs of web services are in general high and the distribution costs low, the underlying business model and financial model must be carefully considered. For public sector bodies the costs of web services will be relatively lower due to their economies of scale. Data often is already available as they are often the holder of such data, and personnel often can be drawn from ICT departments. However, some major aspects still have to be addressed.

The web service should be designed with a clear vision. The STOF Model offers a useful framework to address key components. Firstly, the service component must be addressed. Aspects such as intended users (other public sector bodies, private sector), which functionalities the web service should have, should be considered. Once a type

of web service (WMS, WFS/WCS, WIS, DS) has been selected, technical adaptations may have to be made to cope with data protection and, if needed, payment facilities. Server and broadband capacity should match the expected number of simultaneous users, bearing in mind that new web services often attract many visitors in the first months before the number settles. Web services such as TIM-online in North Rhine Westphalia (Germany), GeoNorge in Norway and Geoportail in France attract millions of visitors per year and their number still increase progressively. It is advisable to design a feedback mechanism for users for quality control.

Developing web services often requires collaboration with other departments or organisations. Therefore, attention must be paid to the actors and networks involved. However, networks are dynamic; changes in policy and legislation will cause actors and their roles to change during the period of collaboration. So, it is important to establish formal and informal agreements on the respective roles and responsibilities within the network. If information is used from third parties, e.g. aerial photography from the private sector, care has to be taken that licence restrictions are complied with. It is vital that when licence agreements with third parties are drawn up, it is made clear in advance that the information will be made available through web services to avoid legal problems afterwards.

Lastly, the financial aspects have to be considered. These aspects include selecting the most suitable revenue model for the type of information made available and which tariff scale, if applicable, will be employed. If fees are to be charged, it is important to set the fees appropriately, as the fee structure is the most visible part of a web service. If the fees are too high, they will form a bar for potential users and insufficient revenue will be raised to cover the costs. Fees that may appear too low to recover costs in the short term may turn out to attract more users that are new and thus actually increase revenue.

The Subscription Model is best suited to web services that offer frequently used information. The user has a clear indication of ongoing fees in return for unlimited use of data within the subscription limit. The supplier has a clear indication of revenue received upfront. The Usage Model is best suited to ad hoc users whereby access to services is more important than possession. However, the Usage Model is only suitable when data is only available from only one or a few sources as the pricing mechanism can become complicated. The Royalty Model is most suited to VARs who need some time to experiment to develop a viable product or service. For the supplier the short-term revenue is uncertain but the long-term revenue may compensate the initial losses. This model is therefore very suitable to public sector bodies that either have an additional source of funding or already have established a steady flow of income out of earlier royalties. The Free Model is best suited to information supplied by public sector bodies funded out of general revenue. It is an open access model, which should remove the current barriers to reuse of PSGI. However, supplying certain PSGI data may be in breach with Fair Trade Legislation if the private sector has already developed similar datasets. The Hybrid Models, either combining aspects of the above models or borrowing elements of revenue models from the creative domain, offer interesting possibilities. The Community Model involves the end-user and thus, provides essential feedback for a successful web service. The Enticement Model can be used in combination with fee-based web services to attract new customers. The Street Performer Model can be adapted for establishing GIIs for the public sector.

6. CONCLUSIONS AND RECOMMENDATIONS

In the last decade, the way GI is used has shifted from only being used in niche applications to becoming embedded everywhere in society. Technological and societal changes have made unlocking PSGI easier. As GIIs are evolving from first generation to second generation GIIs, more and more PSGI web services are set up. However, as technology has changed to make PSGI available, so should the underlying business models and financial models, especially in light of the upcoming INSPIRE implementation. If the only users of a PSGI web service are other public sector bodies, especially when the web service is part of a NGII, then the only viable revenue model is the Free Model or the Data-for-Data Model as variant of the Street Performer Model. Not only is it counterproductive for public sector organisations to invoice each other every time a web service is used, there is also a real risk that public sector organisations will prefer to use (a combination of) alternative 'free' sources such as Google Earth and OpenStreetMap rather than their 'own' public sector geographic information. This contradicts with the spirit of the INSPIRE Directive (see Giff et al., 2008).

If PSGI web services are made available outside the public sector to society, then the only viable revenue model for viewing services such as WMS is the Free Model. The Royalty Model could also be used, as this is effectively a 'free' model since no value added products will be created by just viewing. The private sector, which may need PSGI for their own business processes or to produce value added products, will be prepared to pay for good quality PSGI provided the fees are not too prohibitive. Therefore, for reusers of WFS, WCS and Data Services the Subscription Model, the Royalty Model or Hybrid Models would be suitable. Although the Usage Model is commonly applied, in the long term it is not be viable even for high-quality Large Scale Base Maps. The fees, even with price discrimination discounts, will become too steep for larger areas and the fee structure will become complicated when combined with other data.

To ensure that PSGI is truly shared through web services as envisaged by INSPIRE, national governments will have to provide sufficient funding to guarantee continuous quality. This means that the current cost recovery regime has to be reconsidered. Recent reports in 2008 such as the Cambridge Report (Cambridge University, 2008) and the MICUS Report on Assessment of Re-use of PSI (MICUS, 2008a) support this point of view. While the cost recovery model ensures that a public sector organisation can guarantee that PSGI is maintained at a sufficient level of quality of PSGI (van Loenen, 2009), the model is no longer suited to using web services for PSGI. This is because the specific PSGI data is no longer just accessible from that public sector body but from multiple web service avenues. In the long term, the benefits of making PSGI available free of charge or for lower fees will pay off in the form of intangible benefits and extra revenue raised in the form of taxes when more value added products will be created.

REFERENCES

Anderson, C. (2008). Free! Why $0.00 is the future of business, *Wired Magazine: 16.03 Tech Biz: IT,* February 25, at: http://www.wired.com/techbiz/it/magazine/16-03/ff_free.

Bibber (2008). GeoPortal Networks Working Party, at: https://portal.wur.nl/sites/geoloketten/default.aspx

Bates, M. E., and D. Andersen (2002). Free, fee-based and value-added information services, *The Faciva 2002 White Paper Series* (Dow Jones Reuters Business Interactive LLC), at: http://www.factiva.com/index.asp?node=menuElem0.

Bouwman, H., Faber, E., and T. Haaker (2005). "Designing business models for mobile services: balancing customer and network value", *Proceedings 2nd IEEE International Workshop on Mobile Commerce and Services (WMCS'05), July 19 2005, München, Germany.*

Bouwman, H., Haaker, T. and H. de Vos (2008). *Mobile service innovation and business models,* Heidelberg: Springer Verlag.

Cambridge University [Newbery, D., Bently, L. and R. Pollock] (2008). *Models of Public Sector Information Provision via Trading Funds*. Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007, Cambridge: Cambridge University.

Crompvoets, J., Bregt, A., Rajabifard, A. and I. Williamson (2004). Assessing the worldwide status of national spatial data clearinghouses, *International Journal of Geographical Information Sciences,* 18(7): 665-689.

Faber, E., de Vos, H.,Haaker, T. and H. Bouwman (2008). "About STOF", in Faber, E. and H. de Vos (Eds.). *Creating successful ICT-services. Practical guidelines based on the STOF method,* Enscheda: Telematica Institute, pp. 17-31.

Geoloketten (2008). http://www.geoloketten.nl/wms_integrator_services.html.

Giff, G.A., van Loenen, B. and J.A. Zevenbergen (2008). PSGI policies in Norway and England: are they within the spirit of recent EU directives? *International Journal of Spatial Data Infrastructures Research*, 2008(3): 118-145.

GIM (2009). OpenStreetMap 3D Prototype for Entire Germany, *GIM International*, 20 March, at: http://www.gim-international.com/news/id3688-OpenStreetMap_D_Prototype_for_Entire_Germany.html.

Longhorn, R. A. and M. Blakemore (2008). *Geographic information: value, pricing, production, and consumption,* Boca Raton: CRC Press, Taylor & Francis Group.

Malone, T. W., Weill, P., Lai, R. K., D'Urso, V. T., Herman, G., Apel, T. G. and S. Woerner (2006). Do Some Business Models Perform Better than Others?, *MIT Sloan working paper 4615-06, (*May 2006), at: http://ssrn.com/paper=920667.

MICUS Management Consulting GmbH (2003). *The market for geospatial information: potentials for employment, innovation and value added - Short version*, Study commissioned by Federal Ministry of Economy and Labour (Bundesministerium fuer Wirtschaft und Arbeit)*,* Dusseldorf: MICUS, at: http://www.micus.de.

MICUS Management Consulting GmbH (2008a). *Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report*, Study commissioned by EC in 2007, Dusseldorf: MICUS, at: http://ec.europa.eu/information_society/policy/psi/docs/pdfs/micus_report_december2008.pdf.

MICUS Management Consulting GmbH (2008b). "The value to industry of PSI: the business sector perspective", *Proceedings The socioeconomic effects of Public Sector Information on digital networks: Toward a better understanding of different access and reuse policies, OECD Workshop, February 4, Paris*, *France,* OECD, at: http://www.oecd.org/dataoecd/27/9/40047551.pdf.

Nirov (2007). *Analysis GIS Download 2006-April 2007*, Memo from Jan Kadijk, Project Leader De Nieuwe Kaart van Nederland).

Osterwalder, A. and Y. Pigneur (2002). "An eBusiness Model Ontology for Modeling eBusiness", *Proceedings 15th Bled Electronic Commerce Conference eReality: Constructing the eEconomy, June 17-19 2002, Bled, Slovenia,* pp. 75-91*,* Bled: University of Maribor.

Pigou, A. C. (1920). *The economics of welfare,* London: Macmillan.

Rajabifard, A., Feeney, M.-E. F. and I. P. Williamson (2003). "National SDI initiatives"*,* in I. Williamson, A. Rajabifard and M.-E.F. Feeney (Eds.). *Developing Spatial Data Infrastructures: from Concept to Reality,* London, UK : Taylor & Francis.

Rappa, M. (2003). Business models on the Web, *Managing the digital enterprise. Open educational resource,* North Carolina State University at http://digitalenterprise. org/about.html.

Razi, M. A., Tarn, J. M. and F. A. Siddiqui (2004). Exploring the failure and success of DotComs, *Journal of Information Management & Computer Security,* 12(3): 228-244.

Reitsma, D. (2007). "De consument en zijn behoefte aan geografische data" [The consumer and his need for geographic information], *Proceedings Werkconferentie RGI-117 Geodata, van verstrekking naar toegang, February 13 2007, Den Haag,* pp 35-47. Rotterdam: Bedrijvenplatform Geo-Informatie.

Richard, D. (2008). "Géoportail: The French Geoportal. Access online to public data-sets by means of geographic interfaces", *INSPIRE Drafting Team Network Services Meeting October 15 2008, Amersfoort, Netherlands*. Geonovum.

Schennach, G. (2008). "Reducing charges for national geographic reference data in Austria", *ePSIplus Conference: PSI Re-use - Who takes action next?*, *June 13 2008, Brussels, Belgium,* at: http://www.epsiplus.net/events/epsiplus_conference _psi_re_use_who_takes_action_next/schennach.

Schiff, F. (2003)*.* Business models of news web sites: a survey of empirical trends and expert opinions, *First Monday [Online]* 8(6) June 2 2003 at: http://www.uic.edu/ htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1061/981.

Stieglitz, S., Latteman, C., vom Brocke, J.and C. Sonnenberg (2008). "Economics of virtual communities - a financial analysis of a case study at the Berlin Stock Exchange", *21st Bled eConference eCollaboration: Overcoming boundaries through multi-channel interaction, Bled, Slovenia,* pp. 248-262, Bled: University of Maribor.

van Loenen B. (2009). Developing geographic information infrastructures: the role of access policies, *International Journal of Geographical Information Science,* 23(2): 195-212.

van Loenen, B. (2006). *Developing geographic information infrastructures: The role of information policies,* PhD Dissertation, Delft: OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology.

INTERVIEWS

Jellema, M., DataLand, Reeuwijk, Netherlands, November 20, 2007.

Kooijman, J., TNO Research Centre for Built Environment and Geo Sciences, Utrecht, Netherlands, October 21, 2008.

Murre, L., former director National Cooperation Large Scale Base Map Netherlands (LSV GBKN), Delft, Netherlands, May 21, 2008.

Rijkaart van Capellen, W., director www.buienradar.nl (private sector WMS), Capelle aan den IJssel, Netherlands, April 29, 2008.

Rijnders, J., senior consultant KPN Telecommunications, Utrecht, November 27, 2007.

Schreuder, P., former adjunct director Kadata (former commercial arm of the Dutch Cadastre), Utrecht,, Netherlands, May 8, 2008.

Wormgoor, E. & J. Braaf, Hydrography Service Royal Navy (DHKM), Den Haag, Netherlands, June 25, 2008.

Standard Licences for Geographic Information: the Development and Implementation in Local Government in Italy

L. Garretti[1], S. Griffa[2], R. Lucà[2] and M.T. Lopreiato[2]

[1]Regione Piemonte, Italy, luigi.garretti@regione.piemonte.it
[2]CSI Piemonte, Italy, {griffa.silvana, roberta.luca, mariateresa.lopreiato}@csi.it

Abstract
As technological innovation in the digital frontier advances, many types of content are becoming rapidly available, with no exception in geographic information. Sharing and disseminating geographic information, available by public administration bodies, assume an important role in the geographic information field, above all with the INSPIRE Directive and the Directive on re-use of public sector information. Based on digital rights management principles, a new electronic licensing model was developed by Piedmont Region in Italy, fully adhering to the INSPIRE principles promoting sharing and re-use of geographic information. The model will be effectively implemented in the near future.

Keywords: geolicence, use rights, GeoRM, Intellectual Property Rights (IPR).

1. THE DIFFUSION OF GEOGRAPHIC INFORMATION: ACCESS AND USE POLICIES

As technological innovation in the digital frontier advances, many types of content are becoming rapidly available in digital format, with no exception in geographic information. In fact, the number of uses of geographic information has increased in particular with the evolution of Geographic Information Systems (GIS). New web services and digital media make it easier to access information, so geographic contents can be easily shared, used and reused across the digital domain by various users and devices. This involves particular attention in the use of services and electronic documents, offering greater possibility to share information which, however, at the same time, could allow violation, also involuntary, of third parties' or subjects' rights which share information and metadata. Geographic information moving across digital networks may limit the power of content providers to check the flow of their intellectual property in the geographic marketplace. Increasingly, many public and private information providers wish to protect their information and define distribution policies whilst, at the same time, the users are unaware of the use rights related to a specific resource.

In this context of information sharing, the INSPIRE Directive assumes an important role in the geographic information field. The aim of INSPIRE is to create a European Infrastructure for geographic information to support implementation and monitoring activities related to the environment at European, national and local level. The components of this infrastructure include also agreements on sharing, access and use. the Directive will encourage both the wider sharing of geographic information and services, while respecting intellectual property rights of public administration bodies. It also proposes to implement electronic licences and supports the automated transfer of legal rights.

Currently, public authorities use paper-based licence agreements to protect their intellectual property, with which the acquiring party obtains digital information under restrictions in terms of use, reuse and dissemination rights (Iannella, 2001). Piedmont Re-

gion, aware of the INSPIRE context, is developing a new business model for licensing geographic information. The project, supported by a specific research scholarship, aims to realise a complete policy guideline to regulate the use and dissemination of Piedmont's geographic information, with the addition of a set of standard electronic licences.

In this article, section 2 will describe the project in its regional context whilst section 3 and 4 will treat the different phases of this research. The final section presents the conclusion of this research period and the start of the next implementation phase.

2. DEVELOPMENT AND REUSE OF GEOGRAPHIC INFORMATION

In 2005, Piedmont Region gained experience in legal interoperability, defined as the legal way to reach the capability of information exchange, with the realisation of the 'Discipline of use' for the regional Infrastructure of Spatial Data (SITAD), which in recent years has increased users and information shared. SITAD Rules of Access and Use is a regulatory instrument, which defines roles and responsibilities of users in terms of information use, but it is also open to legislative improvement. However, the Rules refer to generic information and therefore the regional administration developed tools to regulate the dissemination of geographic information available by public administration bodies.

A specific research scholarship 'Development and reuse of geographic information', supported by CSI-Piemonte (the instrumental body of Piedmont Region for Information Systems), started in 2006 to develop new regulatory instruments: 'Guidelines on the use of geographic information' and a set of geolicences (electronic licence agreements related to geographic information).

The main goal of the research was to study in depth methodological and organisational aspects of sharing and disseminating geographical resources (information and services) in the digital environment. The 2–year project aims to provide policy guidelines to help content providers protect their intellectual property (geographic information), using alternative forms of licensing and transmit to end users rules about specific resource uses. The holder of intellectual property rights (information owner) provides specific rights to identified users, within particular constraints, away from an 'all rights reserved' approach towards a 'some rights reserved' approach. The result is a business model for licensing geographic information based on a set of rights associated with specific geographic resources.

The first phase of this research was dedicated to understand the overall context related to use, dissemination and reuse of digital information, in particular geographic resources. The analysis of the international activities in managing use rights of geographic information, focused on Digital Rights Management (DRM) and its application in the geographic field (GeoDRM). Also the European legal context on information sharing was studied. The second phase concerned the development of regulatory instruments for geographic information dissemination. In this stage, a new business model for geographic information was developed and the basis of the practical implementation was founded.

3. PHASE I: EUROPEAN CONTEXT AND ACTIVITIES ON USE AND REUSE OF GEOGRAPHIC INFORMATION

In the first phase of the research the overall context related to the dissemination and reuse of digital information was considered, especially in the geographic information environment.

3.1 Digital Rights Management (DRM) and its application in the geographic context (GeoDRM)

Digital Rights Management (DRM) focuses on the protection of digital content: the term 'Digital' refers to the protected resources, 'Rights' refers to the intellectual property rights linked to the resources and 'Management' applies to defining and enforcing a policy to respect these rights. In this way information owners and distributors can control type of access and use of digital contents.

Geographic Digital Rights Management (GeoDRM) is the application of DRM in the geographic field, and it provides a framework to legislate the use of geographic information. The narrow definition of GeoDRM refers to "electronic licensing of geographic resources to manage and protect intellectual property rights" (Vowles, 2007). In the broad definition, GeoDRM can be taken to cover a large "spectrum of capabilities and underlying technologies supporting description, identification, trading, protecting monitoring and tracking of all forms of use rights for both tangible and intangible (electronic) assets, including the management of rights-holders relationships" (OGC, 2007).

The GeoDRM application needs a set of technologies and a legal framework, where all rights over geographic resources are specified by the information owner, while information users obtain geo information under certified conditions by using an automated or semi-automated implemented geolicence. Piedmont Region founded its project exactly on this assumption.

3.2 The leading role of Geographic Rights Management in the INSPIRE context

The INSPIRE Directive recognises the importance of intellectual property rights for public authorities, using *Geo Rights Management services* (GeoRM), useful to specify electronically distribution policies. A Rights management service is a technology providing additional functionalities to control access, under specified conditions and policies. It represents a filter between the INSPIRE service bus and INSPIRE network services, as described in the draft Implementing Rules for Network Services (INSPIRE, 2008).

In the INSPIRE context, a service bus is a software bus that allows the connection of the geo-portals and their applications to other INSPIRE network services (discovery, view, download, transformation and invoke services). RM services include three different layers: authentication and authorisation, licensing and eCommerce. GeoRM supports the automated transfer of legal rights to final users using electronic licences specifying terms and conditions. The traditional method to protect static content is copyright, using paper-based licences. The Creative Commons concept introduces standard electronic licences for static content for reuse of their intellectual property (Welle Donker and Van Loenen, 2006). GeoRM introduces a different way of managing and protecting intellectual property rights, using electronic licences for dynamic content. The regional project is going to follow this approach, to realise 'click-licences' for its geographic information.

3.3 International activities on managing use rights of geographic information

Worldwide, many activities are focusing on the management of use rights linked to geographic information and on instruments to express use conditions. In recent years, in the GeoDRM environment, the main activities are emerging in the Open Geospatial Consortium, in the ISO TC211 Geographic Information/Geomatic and in the Rights Expression Languages (RELs).

Open Geospatial Consortium (OGC) is a non-profit international organisation "leading the development of standards for geospatial" information and services (OGC web site). By 2004, the OGC established a GeoDRM working group with the mission of coordinating work done on digital rights management. The main objectives are to implement business models on use rights for web-based services (software components accessible over the web for use in other applications) and to develop OGC specifications and technologies required for GeoDRM. During 2007, GeoDRM working group activities had lead to the drawing up of the Geographic Digital Rights Management Reference Model. The document is an abstract specification defining a conceptual model for digital rights management of geo information and requirements for rights management systems (OGC, 2007).

The International Organization for Standardization (ISO) is the largest developer and publisher of International Standards, formed by a network of national standards Institutes. In the ISO context, the ISO/TC211 Geographic Information/Geomatics working group, responsible for the ISO geographic information standards, is developing an Item proposal – "Rights expression language for geographic information" (2007) - on a Rights Expression Language, to be used to compose digital licences for geographic information. A REL is a language that expresses rights related to a specific resource: in the digital rights management sense it is a formal language that can be executed as an algorithm. The GeoREL is an extension of a rights expression language in the GI field, which defines an authorisation model to specify the semantics of a set of rights expressions, given on specified resources. The Item proposal on GeoREL has been created using the Geographic Digital Rights Management Reference Model, drawn up by the OGC GeoDRM working group.

3.4 European legal context of use, diffusion and reuse of digital information

Public authorities produce large amounts of digital information, called public sector information, such as maps, meteorological, legal and business information. Most of this digital information has commercial potential for reuse. In fact, a growing number of companies are reusing public sector information to create new added-value products and services. In Europe, two specific Directives regulate use, sharing and reuse of digital content produced by public authorities: the Directive on public access to environmental information (2003/4/EC) and the Directive on the reuse of public sector information (PSI Directive, 2003/98/EC).

Directive 2003/4/EC requires Member States to make actively available and disseminate environmental information. It guarantees the right of access to environmental information produced or received by a public authority, in order to achieve the widest dissemination. The PSI Directive, instead, establishes a set of rules to regulate the reuse of existing information held by public sector bodies. These documents can be reused for commercial or non-commercial purposes, and public authorities may make available their documents for reuse through a licence. Member States may adopt standard licences, processed electronically and available in digital format.

The reuse of public sector information assumed an important role in the geographic information field, in particular after the enforcement of the INSPIRE Directive, encouraging both sharing of geographic information and services and protection of intellectual property rights, using electronically implemented licences.

4. PHASE II: A NEW BUSINESS MODEL FOR LICENSING GEOGRAPHIC RESOURCES

During the first phase we paid attention to international activities of managing use rights of geographic information, focusing on the application of Rights Management in the INSPIRE context, to understand the regulatory background in Europe. The second phase started with a preliminary analysis of information, uses and conditions of access, to develop the matrix 'Data/use categories/access'. Then the research continued with the realisation of two regulatory instruments on the use and dissemination of geographic information: a set of 'standard geolicences' and the general 'Guidelines on the use of geographic information'.

4.1 The matrix of 'Data/use categories/access'

The preliminary analysis produced a matrix enabling information owners, in this case a public authority, to define possible uses and different kinds of access for different types of users of their geographic information. The first column identifies the 'data set' on which matrix is applied, the second the 'use categories' in term of permitted uses and the third column defines 'access conditions' such as user types, on-screen display, download and so on.

Figure 1: Matrix 'Data/use categories/access'.

DATA SET		USE CATEGORIES			ACCESS CONDITIONS				

DATA SET	A	COMMERCIAL USE		USER TYPES CATEGORIES	DISPLAY	DOWNLOAD	FEE	FORMAT	OBLIGATIONS
Piedmont's geodata	A.1	Pure commercial use	A.1.A non modified product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA CRIMINAL AUTHORITIE	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL
			A.1.B value added product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA CRIMINAL AUTHORITIE	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL
			A.1.C new product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA CRIMINAL AUTHORITIE	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL
	A.2	Professional services	A.2.A non modified product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA CRIMINAL AUTHORITIE	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL
			A.2.B value added product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA CRIMINAL AUTHORITIE	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL
			A.2.C new product	FREE ACCESS PROFESSIONAL MEN UBLIC ADMINISTRATIO EDUCATIONAL AREA HERITAGE AREA	YES/NO	YES/NO	YES/NO	RASTER / VECTOR	PUBLICATIONS, DISTRIBUTIONS TO THIRD PARTIES, COPYRIGHT LABEL

The 'Data set' column defines geographic information resources, probably in the next research phase the definition may include not only data sets but also services. The section 'use categories' includes commercial and non-commercial use. Commercial

use happens when the final user will reuse licenced resources to create a direct or indirect gain, and this category includes 'pure commercial use' and 'professional services'. The first one means a direct financial gain arising from reuse of resources: for example a company that acquires a data set to resell them. 'Professional services' refer to professional use, in which there is an indirect gain, because professionals reuse resources during their activities and receive fees: for instance, an architect using licenced information to create a master plan.

Non-commercial use refers to reuse activities not producing direct financial gain. This category includes internal use, teaching use, spreading use, research use and institutional use. The definition of 'internal use' refers to a reuse of licenced resources for internal use only, in which the customer can manipulate geo information for internal archives but cannot disseminate the resultant products in the public domain. 'Teaching use' includes the reuse for educational activities and 'spreading use' refers to spreading activities such as meetings, exhibitions or free publications. Reuse activities related to research refer to 'research use' and 'institutional use' including reuse in public authorities activities and projects.

For each use category (commercial and non-commercial), including sub-categories, three levels of reuse exist:

1. *non modified product*: end user displays and makes query on licenced product without modifying it;
2. *value added product*: the product is developed by the end user and contains information from the licenced product in addiction to other new information;
3. *new product*: the product is developed by the end user and does not contain information from the licenced product, but only other new information.

The third column describes '*access conditions*' such as user types or use rights related to display and download. User types category includes 'free access', when end user accesses and obtains resources without personal identification, and 'registered access', in the case of personal identification of the end user. Registered access contains the following sub-categories: professional, public administration bodies, heritage and educational area, law enforcement agencies. Other kinds of access conditions refer to on-screen display, download, format and distribution to third parties, which content provider decides to permit or not to final user. Concerning resource price content providers may opt for:

(a) a *cash fee*: data owner may establish a cash fee for each allowable information item;
(b) a *service fee*: data owner may establish a service fee instead of a cash fee. Data owner asks the final user to return updated information item;
(c) a *mixed fee*: data owner may establish both a cash and a service fee.

Concerning constraints, the access conditions sub-category defines user obligations in terms of publications and diffusion via web. In this case, users reusing licenced information for publications or diffusion via web must respect strict obligations:

(a) *copyright labels*: data users must use a copyright label referring to the licensor's copyright;
(b) *communications concerning publication conditions*: data users must inform information providers about information publication methods, specifying the amount and the typology (publications, domain addresses or applications);

(c) *free copies or privileged accesses*: data users must inform data owners of free copies or privileged accesses to servers.

4.2 Developing 'standard geolicences'

The elaboration of this matrix is the background for the development of 'standardised geolicences' and 'Guidelines on the use of geographic information'. It represents the conceptual model useful to implement digital rights management capabilities for geographic applications. Enforcing the GeoRM application, the holder of intellectual property rights (information owner) agrees to provide specific permission (use rights) to identified parties (users), under various constraints. Thanks to the definition of these use cases, a first draft of the licence was drawn up in this second research phase. The draft licence will be used in the near future to compile automatically electronic geolicences for information distribution, considering specific restrictions and requirements that information owner, in this case Piedmont Region but also the other regional public administrations, establishes for its information.

Geolicence defines terms and conditions for the end user which acquires a set of rights (permitted uses, display, download, format, distribution) under specific constraints. The licensor and the licensee play primary roles: information owner (Piedmont Region) as a licensor, defines policies to be applied when information and services are distributed among the stakeholders and the end user, as a licensee, accesses geographic information under specific terms and conditions (OGC, 2007). Successful negotiations create standardised geolicences, expressed through two forms:

1. *legal code*: representing a legal expression of terms and conditions of the licence;
2. *user friendly*: a simplified, readable version for the web, expressing terms and conditions.

The implementation phase, starting from now on, will enable end users to access or download licenced geographic information under specific terms and conditions, electronically composed by a web form in a '*click geolicence*'. This is a common form of agreement, mostly found on the web, as part of information download. Click licence contents and forms vary widely, but the main activity requires end users to subscribe to the licence or not by clicking an 'agree button' in a dialog box. Piedmont Region's geographic information will be disseminated using a click geolicence system, in which end users will enter the regional Spatial Data Infrastructure, will search information using a metadata catalogue and obtain selected geographic information, after subscribing to a geolicence agreement. The result may permit display or download.

The web form of Piedmont click licences, now in a draft version, is composed of three sections:

1. *user information*: this part refers to user information such as name, surname, e-mail, phone number, user type class, fiscal code number, etc;
2. *use categories*: this section contains selected use categories (commercial or non - commercial);
3. *access conditions*: what licensee 'can do', after obtaining licenced resources.

Click licence is a user friendly form: the end user will receive also a 'legal code' by e-mail. In the section use categories and access conditions, the licensee can select only the permitted use, previously defined via the matrix 'data/use categories/access', whilst other unauthorised uses are frozen.

Geolicences implementation has drawn attention to verification of the user's identity. During this first licence implementation, the system requires the name, surname, e-mail and phone number of the end user. Whilst a working e-mail address is the sufficient condition to obtain a licenced product the information concerning name and telephone number cannot be easily verified. In the following research phase, we will study and define new ways to make a positive identification of users such as initial registration, fiscal code number or login with password.

4.3 Guidelines on the use of geographic information

The final step of this research is the formulation of the 'Guidelines on the use of geographic information'. To apply the Geolicence model to public administration in Piedmont, the management of geographic information use rights needs general regulatory instruments concerning sharing and dissemination of geographic information.

Guidelines define general aspects on sharing Piedmont's geographical knowledge system and implementing public and private activities reusing geographic information. Guidelines regulate public and private access to Piedmont's geographic information, specifying general policies and harmonised geolicences.

Guidelines establish some common policies:

− all territorial information will be available to public and private bodies, in order to avoid information dispersion, duplication and non-authorised use;
− access to geographic information will be under impartial and non-discriminatory conditions;
− to allow free access to Piedmont's geographic services: users can access without paying fees;
− to adopt several information access profiles: users can access to geographic information and services, using different level profiles;
− to disseminate information always using geolicences specifying use rights;
− geolicences establish use rights in terms of display, download, permitted reuse, distribution to third parties and publication.

The definition of strategic, organisational, technical and operating guidelines is a priority objective. Piedmont's geographic knowledge will be shared using the regional Spatial Data Infrastructure, connected with the regional Geographical Information System, a unitary system shared among local public administrations and citizens that use common information, policies, procedures and infrastructures.

5. CONCLUSIONS

The two year project expects to define a new way of sharing geographic information: away from an 'all rights reserved' approach towards a 'some rights reserved' approach. The realisation of a business model licensing for geographic information, based on a set of rights associated with specific geographic resources, is a specific aim of Piedmont Region, above all with the definition of the regional Geographical Information System. The results coming from this project are the preparation of Guidelines, as a management instrument to define general aspects on sharing the regional geographical knowledge system, and the implementation of click geolicences; an electronically implemented agreement which defines terms and conditions for information owner and end user.

The 'Guidelines on the use of geographic information' were recently submitted for enforcement by the regional government and we are now testing a geolicence prototype. However, the main aim in the near future will be the implementation of regional *Rights Management Services*, as described in the INSPIRE context, useful to manage access to regional geo-services.

ACKNOWLEDGEMENTS

REFERENCES

Bishr M. (2006). *Geospatial Digital Rights Management with Focus on Digital Licensing of GML Datasets*, Thesis submitted to the International Institute for Geo-information, Netherlands.

Bishr M., Wytzisk A. and Morale J. (2006). "GeoDRM- Towards Digital Management of Intellectual Property Rights for SD", *GSDI – 9 Conference Proceedings, November 6-10 2007, Santiago, Chile,* at: http://gsdidocs.org/gsdiconf/GSDI-9/papers/TS8.4paper.pdf.

Cary T. (2007). Geospatial Digital Rights Management: more than making money, *Geoworld*, 2007(1): 32-35.

European Parliament (2003). Directive 2003/4/CE of the European Parliament and the Council of January 28, 2003 on public access to environmental information.

European Parliament (2003). Directive 2003/98/CE of the European Parliament and the Council of November 17, 2003 on the reuse of public sector information.

European Parliament (2007). DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

Herring J. (2005). Digital Rights Management for the Geospatial Community, *Geoplace*, 2005 (1), at: http://www.geoplace.com.

Iannella R. (2001). Digital Rights Management Architecture*, D.Lib Magazine*, 7(6), at: http://www.dlib.org/dlib/june01/iannella/06iannella.html.

INSPIRE Network Services Drafting Team (2008). Draft Network Services Architecture (Version 3.0), at: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/D3_5_INSPIRE_NS_Architecture_v3-0.pdf.

Joffe B. and Bacastow T. (2005). *Geodata access while managing digital rights: the Open Data Consortium's Reference Model,* at: http://www.opendataconsortium.org/.

National Research Council (2004). *Licensing Geographic Data and Services*, Washington D.C.: National Academies Press.

OGC (2007). Geospatial Digital Rights Management Reference Model (GeoDRM RM), at: http://www.opengeospatial.org/standards/as/geodrmrm.

Queensland Spatial Information Office (2006). *Government Information and Open Content Licensing: An Access and Use Strategy*, Queensland, Australia.

Spatial Information Council (2008). *The Value of Spatial Information: the impact of modern spatial information technologies on the Australian economy*, Australia.

Vowles, G. (2007). Geospatial Rights Management, (December), at: http://www.eu-orchestra.org/docs/ORCHESTRA-Day/11_Geospatial_Rights_Management_Graham_Vowles-OS.pdf.

Welle Donker, F. and B. van Loenen, (2006). "Transparency of accessibility to government-owned geo-information", *Proceedings of the 12th EC-GI&GIS Workshop*, Innsbruck, Austria, 21-23 June 2006, pp. 89-96.

Legal Simcity; Legislative Maps and Semantic Web Supporting Conflict Resolution

Rob Peters[1], Rinke Hoekstra[2], Tom van Engers[3] and Erik Hupkes[4]

University of Amsterdam / Leibniz Center for Law
[1]Rpeters@vrk.nl, [2]Hoekstra@uva.nl, [3]T.M.vanEngers@uva.nl
[4]ehupkes@leibnizcenter.org

Abstract
Participative decision-making may promote the quality and the support of regulations. This also applies to regulations applying to a location. To date it has been very difficult for citizens to participate in legislative debates since this domain requires a level of expertise which is not widely available. Traditional approaches providing access to these regulations are not satisfactory to citizens since they are confronted with vast amounts of often contradicting regulations. Questions like "where will I be able to do this kind of activity" or "will this activity be allowed here" are hard to answer in traditional web-based service environments. There are many attempts to create one-stop-shop front-ends to eGovernment, but these are seldom built from the perspective of the user. Developing more sophisticated visualization tools allows for a future in which legal planning is an important part of modern democracies. More accessible interfaces will mean that people can engage in a dialogue between interests, possibilities and regulative impact as a form of balanced system management rather than voting for or against a proposal set by experts. Improving the access to the legal planning process implies that legislation can become part of the democratic debate rather than the territory of experts. This article describes a number of prototypes that have been iteratively built and resulted in the Legal Atlas approach. Legal Atlas seems to provide the required supporting environment for public authorities that govern complex issues that require a participative policy- and decision-making strategy. The Legal Atlas system described here is designed to support INSPIRE environmental policy implementation. Qualified map layers and dynamic legal comparison using Simcity-like manoeuvrability can help to avoid conflict polarisation and result in conflict resolution.

Keywords: INSPIRE, geographic information, regulations, semantics, participatory GIS, Europe, Feed project.

1. INTRODUCTION

Governments have recognised the need to take advantage of 'the wisdom of the crowds', and the need for finding solutions to complex problems that in the end are supported by as many stakeholders as possible. This requires new ways of interacting with stakeholders. The complexity of the interaction between many different existing legal sources having constraints on solutions to problems at hand make it difficult for both experts and laymen to understand the consequences of their proposed solutions. Balancing the interests of stakeholders involved make it even more complex.

Governments realise that participative approaches to policy- and decision-making are helpful to create better regulations that are supported by a majority of stakeholders. Recently, different web-based support tools have been developed in various e-participation projects, allowing citizens participation and building their support for decisions to be made. Legal pluralism and the huge volumes of sometimes conflicting regulations

make it hard for citizens and civil servants to oversee the consequences of proposed solutions.

In this article we will show the Legal Atlas as an eGovernment approach that helps citizens, businesses and civil servants to cope with a vast number of rules in their day-to-day processes. Standardised legal knowledge representations and maps enable stakeholders to cope with legal pluralism and the complex interaction between European and national legislation and local regulations. We advocate that the Legal Atlas approach helps users to access relevant regulations concisely and comprehensively. The map-based interface providing access to the applicable regulations offers a powerful means to ensure usability and client satisfaction.

This article explains in section 2 the need for such an approach, and the research set-up. Sections 3 and 4 describe the development of the prototype towards the Legal Atlas application. This is followed by a section assessing the impact of INSPIRE for local authorities. Finally, conclusions are provided and further research recommended.

2. CITIZEN'S NEEDS AND LEGAL ATLAS

The Legal Atlas systems should provide the ultimate transparency of local regulations so that citizens could ask 'where can I do this?', instead of 'can I do this in place such and such?' (Wilson and Peters, 2004). Such an application requires a unique coding for all local and regional legal functionalities affecting citizens and a mechanism for semantically mapping (Boer et al., 2002) between the legal terms and legal procedures affecting those required functionalities. In that way the citizen could view and monitor the consequences of decisions and opposition. The Legal Atlas uses the IMRO XML/GML coding scheme (the legally binding Dutch information model for spatial planning) to enable the user to query the underlying legal system using detailed coloured maps representing legal constraints and possibilities (supporting explanation and understanding by users).

Our research followed an iterative design approach resulting in Legal Atlas prototypes I, II and III. For Legal Atlas I (see Figure 1), we conducted empirical studies (Peters and Van Engers, 2004) in decision making-processes reflecting their complexity including the stakeholders involved in such processes. Users from government and businesses tested a number of prototypes, screen mock-ups and real life systems. Most testing groups were between 10 and 20 individuals. Different forms of performance measuring provided the necessary feedback on the adequacy of our solution, and different interviewing techniques were applied to obtain thorough feedback of the users. This resulted in Legal Atlas I which contained navigation techniques that helped to navigate through four levels of laws and regulations: European, national, regional and municipal. Early results from using and demonstrating the system in the Netherlands (Wilson and Peters, 2004) showed high user satisfaction and gave confidence that this type of solution would be feasible in other European countries.

3. LEGAL ATLAS II

A second version of the Legal Atlas builds on the experiences from Legal Atlas I. In Legal Atlas II, we wanted to extend the retrieval possibilities of law from text to map. By clicking on a regulation the user would be able to identify all objects or map areas where this regulation applies through a visualisation mechanism. While this may seem a small extension of the first version, the reader should realise that there is a big difference between the question: "what will be allowed here?" and "where will this be al-

Figure 1: Result set showing protected coastal area with socio economic map layers and geocoded Legal document retrieval on terms like "alcohol".
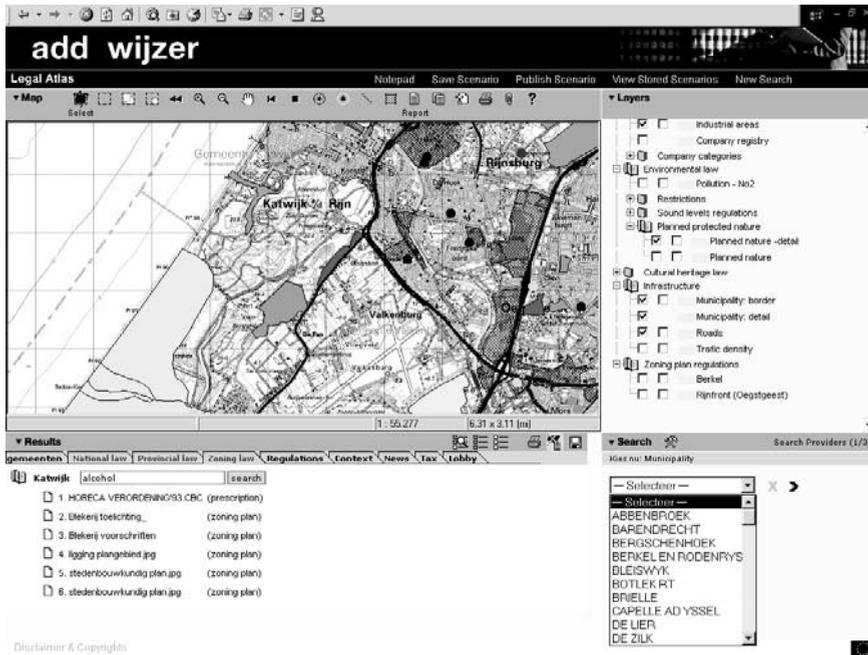
lowed?" The second inverse question demands a complete semantic definition of the world and complete legal coverage of that world (Boer et al., 2007).

The prototype of the second version demonstrating this functionality is shown in Figure 2. The darkened contour in the picture is selected by searching for 'housing' in the legal texts that apply to the area. The application seeks for 'housing' among all legal texts and returns hits with those regulations that apply to 'housing'. When the regulation is selected it returns all relevant map contours where this regulation applies.
With additional legal reasoning (Deontic) arithmetic, such a functionality enables questions like: return all contours on the map where regulation 'X' and 'Y' apply with exception of rule 'Z' (Boer et al., 2003).

The codes are now translated to GML syntax rules, which helps to create a more flexible and maintainable codification system. Legal Atlas I already had demonstrated the usefulness of Geographic-Legal Information systems, but the legal sources were selected manually. Legal Atlas II was based on an open and extensible architecture using semantic web technologies (Winkels et al., 2007).

4. USER CONSULTATION AND EVALUATION OF LEGAL ATLAS II

To create and test the Legal Atlas II system we included consultation and user evaluation activities in our development project. From March to May 2008, 12 in-depth interviews were conducted with building experts, development experts, environmental specialists and government officials. The purpose of the prototype was to engage the interviewees in a design dialogue before building a real life system. The second struc-

Figure 2: Prototype to test navigating from text to map or from regulation to application area.
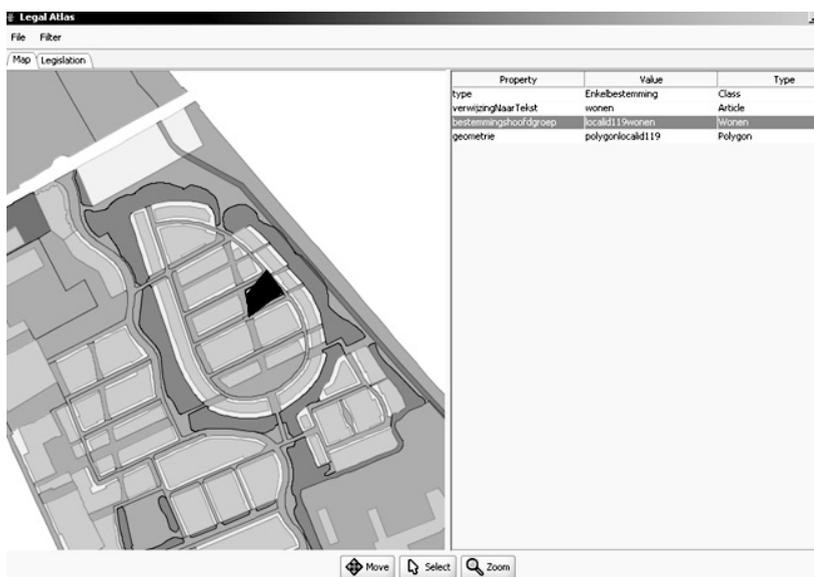
tured range of tests were based on Legal Atlas II. Both sets were recorded on video-tape, transcribed and analysed. These tests followed the Software usability measurement inventory (SUMI) approach (Human Factors Research Group, 1994). This approach is more catered for usability testing than for affordance testing in the line of Gibson's school of functionality design.

After this limited prototype of text-to-map search of legal constraints, we developed a realistic prototype for which we used data of the Dutch province of Flevoland. This prototype uses IMRO combined with GEMET ('GEneral Multilingual Environmental Thesaurus'). The regional data layers and most of the content are provided by Flevoland that maintains a number of databases about the economy, the environment, our cultural heritage and the public infrastructure.

The dark areas are text-to-map areas that can be activated using keywords out of the controlled vocabulary. The next step is to concentrate the retrieval possibility on a much more narrow scale (see Figure 3).

For the purpose of European transferability and usability Google maps technology is used next to GIS technology to support 'contour editing' by visitors of the website. The visitor can indicate an area by drawing a line on the map to discuss and to support his case with documents, research reports, legislation and plans that can be uploaded, stored and downloaded for that drawn area.

The interface of the prototype is linked to the Oracle database of the province of Flevoland. All map layers can be projected on this Google map screen. The conflict resolution system will maintain dynamic input from legal constraints and map layer updates. The document retrieved in this particular screen dump example shows the relevant research report as 'hit' upon entering the keyword *Algemeen_Rec* (recreation_general) in the search field. The hits are restricted to the square area drawn on the map by the user. In the same way, resources can be uploaded and downloaded.

Figure 3: Detailed text-to-map retrieval with response based on IMRO codification.
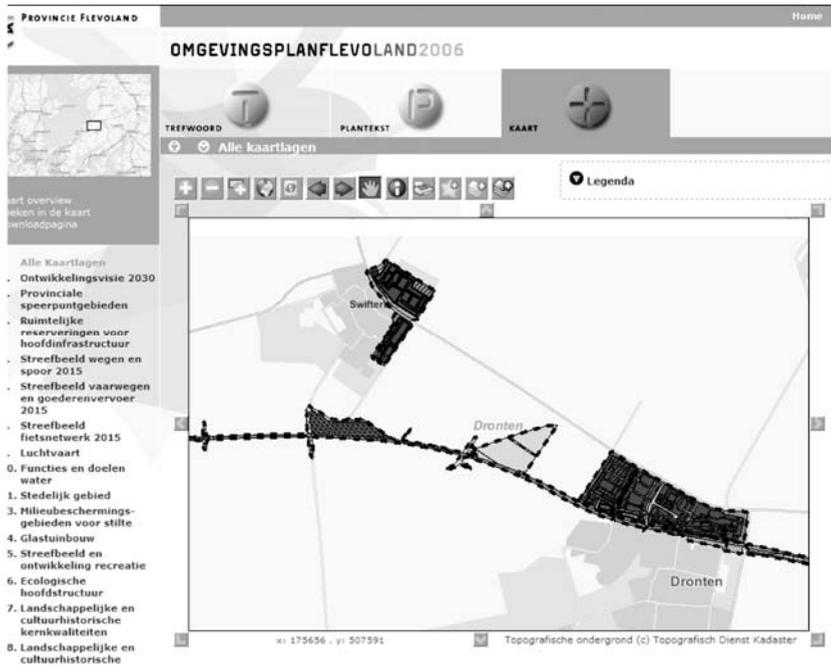
5. CONFLICT RESOLUTION AND THE USE CASE INSPIRE IN FLEVOLAND

For the real life trial we used the example of INSPIRE annex III layers 11 and 18: areas with legal constraints, regulated areas and Habitat/Natura2000 areas. The Lake Markermeer requires numerous ecological measurements to protect the area from further determination. The water management suffers from sediment that creates uninhabitable murky waters. The ecological balance of the region is even more threatened by the obligation to host another 60 000 houses with accompanying recreational and transport infrastructure.

Calling the region 'locked under Natura2000' as indicated by the standard map layering is a simplification that would bring people into the courtroom needlessly and too early. The groups involved in protection of birds and those involved in building the infrastructure have been in court for the last 20 years. The chairman of the bird protection stated that the courtroom does not help 'his' food chain to develop in a natural way. Pressure groups require more insight in each other's negotiation space. Another reason to have dynamic maps fuelled by dynamic constraint analyses is the fact that no protected area is fixed in the way some maps tend to indicate. For example, the ministerial website shows on a map small contoured areas which represent protected breeding areas for a certain type of duck. In reality these regions may shift a few miles in each direction, depending on the best trade-off in view of the NGOs involved. A more dynamic design of the relation between the map and the constraint it represents, would support the legal case much better.

Inspire (INSPIRE, 2008) has an explicit limitation to its scope as support mechanism for transparency. It states in article 13.1(h):

"Member States may limit public access to spatial data sets and services [...] where such access would adversely affect [...] the protection of the environment to which such information relates, such as the location of rare species."

We argue here that map-layers represented in static ways powered by static legal- and policy constraints do exactly that. They do not support conflict resolution and negotiation, but rather suggest more inflexibility and legal rigidity. INSPIRE based static map layers may be counterproductive to conflict resolution because of the tendency to be too general or too specific with no dynamic adjustment possibilities based on flexible regulative parameters. The interesting difference between a screen showing map layers already there and those that 'turn up' while moving a qualified cursor (like the one seeking space for open sailing areas) is the affordance of opportunity finding in contrast with the annotation of an area that is 'locked up'. The provinces call this functionality a 'seeking area'. They have created the unusual legal term 'seeking area' to obtain legal degrees of freedom in development plans that do not occur with fixed parameterisation.

Feedback on the tests with the Flevoland regional development plan prove that questions like: return all contours on the map that fulfils the legal constraints 'X' and 'Y', but not 'Z' are answerable. How does one provide such type of opportunity finding for the user in a meaningful representation that allows for more flexibility? We have argued that INSPIRE Maps showing Natura2000 areas or sites should enable the functionality of 'seeking area'. Simcity game developers who created manoeuvrability using a cursor and 'tiles' with fixed business rules may have developed the answer already. This was in 1985 when Simcity was still called 'Micropolis' (Wright, 2004a).

6. CONFLICT ANNOTATION ENGINE FOR LEGAL ATLAS III

Simcity-like functionality (see Wright, 2004b) resembles the required flexibility or 'seeking area'. It is mostly based on Semantic Web Technology. The knowledge models about the legal constraints and the domain knowledge of the working scenarios are all described with Resource Description Framework/ Web Ontology Language (RDF/ OWL). We chose RDF/OWL to infer and reason with these models. To publish this information as a service we use the OpenRDF Sesame server. This server has an SPARQL (*SPARQL Protocol and RDF Query Language*)-endpoint, which is an access point to which SPARQL queries can be sent. The SPARQL-endpoint is accessible through the web. The RDF that is stored within the OpenRDF Sesame server is processed with OWLIM. OWLIM is a high-performance semantic repository. It is packaged as a Storage and Inference Layer (SAIL) for the Sesame RDF database. It reasons about the RDF data and propagates this by means of rule-entailment. The SPARQL-endpoint is used to fill the Legal Atlas III with information. The Legal Atlas III is as an interface for the OpenRDF Sesame server, and the SPARQL-endpoint is the interface between them.

The SPARQL queries are based on the schemata of the RDF/OWL models (van de Ven et al., 2007). This means that they are independent on the content. This ensures that different content is annotated with the RDF/OWL models to ensure that the SPARQL queries are able to retrieve the content. The return is a gigantic list of all the concepts that can be used for annotation. Such a huge list might not be convenient in a user interface. Therefore it is better to replace this list with a practical list. Pruning this list down to a domain is one way to limit the amount of concepts. The following example shows the SPARQL query for a specific domain, namely the IMRO2006 (*Informatiemodel Ruimtelijke Ordening*, Dutch information model for spatial planning) SKOS

(Simple Knowledge Organization System) vocabulary. These queries can be specified further and can become complex.

Figure 4 shows the architecture of the interaction between the representation portal and the Annotation engine. The representation portal containing Legal Atlas gets its information by sending the appropriate queries to the SPARQL-endpoint. The SPARQL-endpoint is an access point to the OpenRDF Sesame server. The OWLIM is an inference layer within the OpenRDF Sesame server. This is filled with the Cross Border Deliberation models and the domain models and the content that is part of the representation portal.

Figure 4: Architecture of the annotation engine of a map based conflict resolution portal.
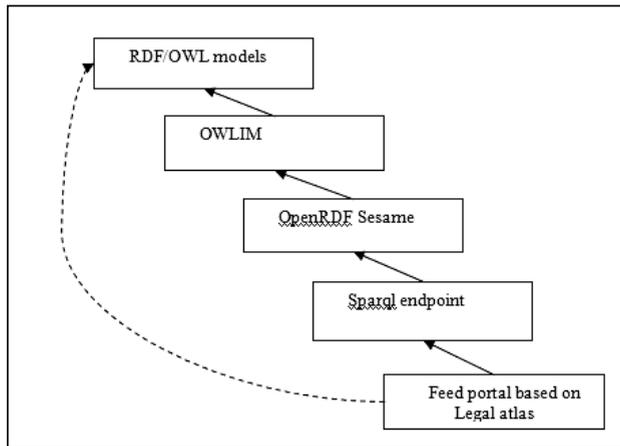
The portal also provides an interface for adding data. It adds RDF/OWL to the models. The OWLIM module then processes these and the result is propagated to the Feed Portal by the same SPARQL-endpoint.

The 'seeking area' functionality of Legal Atlas III is about solving conflict situations. Figure 5 describes the architecture. A description of the plan in the middle top box is combined with the normative descriptions in the regulations. Both the regulations and the plans are about SKOS vocabulary and regions, and it is possible to trigger particular situations that are allowed or disallowed. The OWL models do not become inconsistent, but the allowed and disallowed situations are also OWL classes. It is now possible to represent the disallowed situations and their regions as conflicts on the map (Hoekstra et al., 2009).

7. EARLY RESULTS OF LEGAL ATLAS III

The last and third set of structured interviews focused on Gibson's (Gibson, 1979) affordance needs in the INSPIRE (EC-INSPIRE, 2003) production chain. 30 different experts were interviewed in this way. Figure 6 shows one of the early results of our research. The most ergonomical lay-out for a use case would be, that the user drags the cursor (the black circle in Figure 6a) and the surrounding area highlights problem areas (striped circle, in this case a 5 000 ha. swamp that is envisioned to be the ecological motor for the lake) and areas up for negotiation (see oval in Figure 6b). When the cur-

sor is dragged further into the protected area it turns red to indicate the legal absolute negative (i.e. the oval in Figure 6c).

Figure 5: Conflict situation architecture.
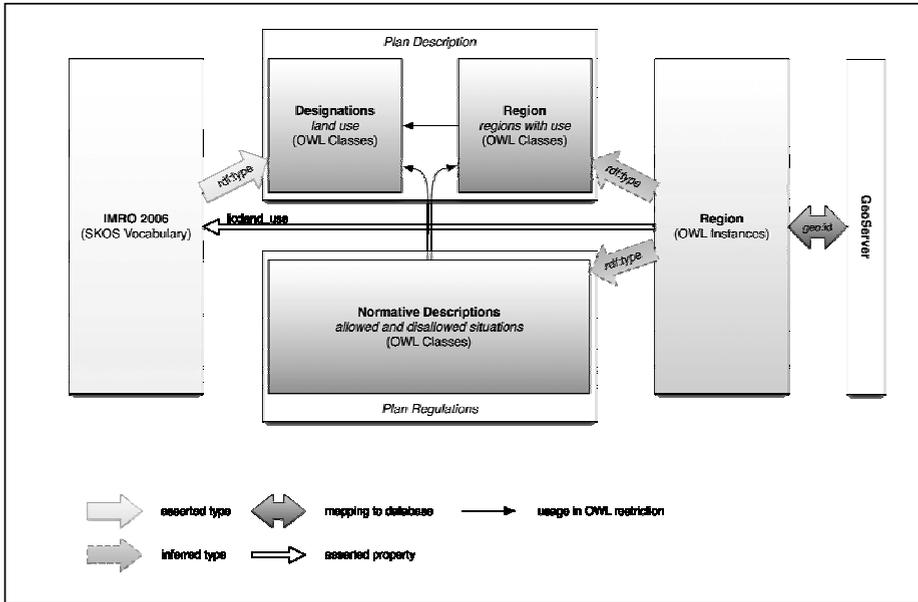
The upload and retrieval functionality has been judged to be part of the Atlas behind the decision process. The legal annotation engine is currently being embedded behind the spatial context (*Omgevingsplan*) of Flevoland.

Figure 6: Simcity approach to conflict resolution and opportunity finding (a, b, c).
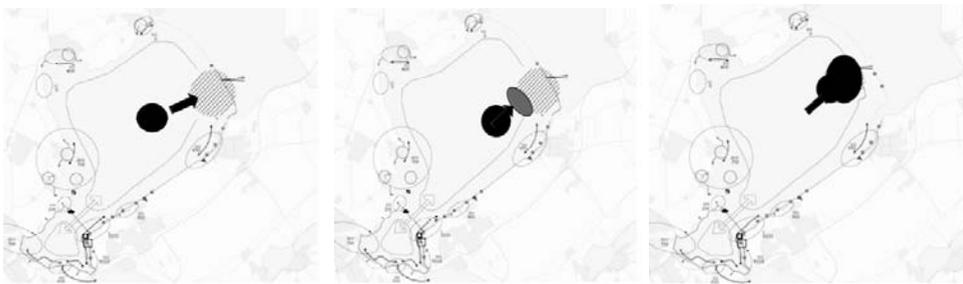
a   b   c

The map layer turning up while approaching with the qualified cursor simulates the 'seeking area' degree of freedom the authorities require to state intentions with a certain level of flexibility, like moving 2 km to the right or left if this helps with the original ecological goals. We call this the Simcity approach of Legal Atlas, as it allows for playing around until a solution is found fit for purposes which is similar to the Simcity game. The English partners in the project have focused on the Petitioning upload and retrieval based on maps. Their legal infrastructure and the culture are not catered yet for this

Legal Atlas approach, except for petitioning. The Greek partners in the project also seem to be faced with a less rule-oriented culture. The standardisation efforts in Europe may result in decreasing these differences when it comes to enforcement and comparison of policies in Europe.

8. RELEVANCY FOR INSPIRE MAPPING FOR DUTCH PROVINCES

The Dutch provinces have carried out an impact analysis of INSPIRE (Woudenberg et al., 2009). This impact analyses mentions several ambiguities with the new infrastructure, starting with the definition of 'Protected sites' according to the International Union for the Conservation of Nature (IUCN):

"A Protected Site is an area of land and/or sea especially dedicated to the protection and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means".

The scope of this definition raises a discussion among policy advisors how to apply the INSPIRE directive wisely. If too many 'protected sites' are defined it tends to weaken enforcement possibilities and reporting to the European Commission and the Shared Environmental Information System (SEIS) may become tedious. If too few areas are being identified as protected the province will lose those 'sites' it wants to protect. The latter occurs because of the hierarchical legal effect of European directives. 'Lesser' legal regimes become less useful to oppose against economic interests. The Ecological main infrastructure (*Ecologische Hoofd Structuur*) in the Netherlands, that was meant to become the vehicle to ensure less fragmentation of green areas in an urban country, may well become negligible in it's effect given the greater impact of Natura2000 areas. The Dutch provinces have several examples of non-intended effects. The regulation of salt and sweat water balance, for example, indicated that one province was obliged to subsidise farmers for lack of sweet water while those farmers had one of the best crops in Europe. The Dutch provinces do not want to lose this need for flexibility due to inflexible standardisation. These findings are relevant for our research on Legal Atlas. INSPIRE is an interesting problem area creating many challenges for the governments. INSPIRE not only states what environmental indicators have to be measured but it also states how these environmental indicators should be measured. As a consequence INSPIRE has implications for governments since it may impact future spatial plans.

It appears that often the visualisation of map layers for environmental purposes, like with INSPIRE, creates a legal challenge; the feeding time of geese in the Netherlands is six months at most. Then they fly to Scandinavia. The map showing areas that are protected for the sake of those geese should convey the temporal factor which is unusual in legal constraints mapping.

For Dutch provinces, the biggest problems of European standardisation efforts like INSPIRE concern the harmonisation of calculating methods behind the layers. The Dutch provinces are not sure that all environmental indicators are calculated in the same way in Europe. This envisioned level of variety requires more explanation, flexibility and enough administrative room for local interpretation. Even within the same country it takes time and effort to understand the different ways to calculate sound level measurements, visualisations and mappings. Water quality differs with temperature. Ecological complexity defies simple arithmetic's. There is a relationship between legislation and methods of calculation; changes in calculation methods at the European level may

cause changes in Dutch and regional legislation. Therefore, INSPIRE may well be one of the directives with the highest impact on local and provincial authorities.

9. CONCLUSIONS

This article described three prototypes that have been iteratively built and have resulted in the current Legal Atlas approach. Legal Atlas seems to offer the required supporting environment fulfilling a clear need of those public authorities that govern complex issues that require a participative policy- and decision-making strategy aimed at conflict resolution rather than polarisation.

The participative and iterative building approach of Legal Atlas has resulted in a series of working prototypes that could be tested in settings that are similar to the problem areas involving the stakeholders working on those problems. This practice-oriented design and development approach resulted in a useful solution; something that probably would have been much harder if we would have limited ourselves to traditional requirements oriented design methods. The participative design method we have chosen made it possible to better understand user needs and helped us to develop a effective Legal Map oriented solution.

We have highlighted problems based on INSPIRE impact research among Dutch provinces. The most important problem concerns the harmonisation of calculation methods of indicators. We showed that we succeeded to build a tool that now becomes part of the core business of the province of Flevoland. It enables further query articulation and precision for citizens and businesses that seek negotiation space, opportunities and problems while the political decision process is underway.

This article described an interaction approach similar to the one used in the well-known Simcity game. This allows users to detect easily the negotiation space in conflicts of interests in complex policy- and decision-making processes.

Despite the fact that our research already shows a solution for handling the complexity caused by legal pluralism, one could argue that this approach is limited to much rule-oriented spatial planning problems that are typical for small countries like the Netherlands. We argue that the need for weighing and balancing various interests in densely populated regions such as the Netherlands may be higher than in less populated regions, but this weighing and balancing is essential for and core to the rule of law and our approach may very well be beneficial to other countries.

It should be stressed that although most of our tests have been conducted in the domain of spatial planning our approach can also be applied to other legal domains. All laws have a jurisdiction and are consequently location bound. While a legislative maps approach has advantages in the domain of spatial planning it is not limited to that domain. Spatial planning is an area where many different interests have to be weighed and balanced, like many other areas. It for the same reasons neither is typically suited for dense regions or countries like the Netherlands. Every regulation has location related elements if only jurisdiction and many problems that require legal reasoning require a smart combination of legal sources, constraint satisfaction and conflict resolution. It is also true that not just spatial regulations require other that text based representations and combining non-textual representation with textual legal sources can be realized using exactly the technology demonstrated here.

Furthermore, we pointed out that the enabling technology, based on Semantic Web technologies including RDF tuples and SPARQL queries allows us to create a truly flexible and dynamic solution. It enables the user to approach any piece of information either starting from the map side or from the text side, and even more important, it allows for intelligent support, i.e. automated reasoning. In this paper we showed how this can be used for constrains satisfaction and conflict resolution, but more complex reasoning about regulations related to spatial objects is also possible using the same underlying technology. For a more detailed description of this technology we refer to Hoekstra et al. (2009).

Further research

Further research should concentrate on other European areas of legal and semantic implications for policy making and balancing interests. Within the spatial domain a future research area could be on balancing environmental with economic and recreational interests in the context of INSPIRE.

The massive volumes of data required to provide simple services is another area for further research One of the prototypes, for example, already contained 3.5 gigabytes of content that was collected from 10 municipalities, five large infrastructural sources and the chamber of commerce. This data was used to 'fill' the prototype and was additional to the already vast amount of content obtained from the province. The issue of content volume becomes more important if we model legal knowledge-based systems at European scale (because of the jurisdictions of national laws).

We will further develop and test new versions of Legal Atlas in 2009. In a new version we will further improve the capacity to handle conflicting interests, to visualise and interact with policy intentions and use automated 'seeking area'.

ACKNOWLEDGMENTS

REFERENCES

Boer, A., R. Peters, T. van Engers, and Radboud Winkels (2007). Separating Law from Geography in GIS-based eGovernment Services, *Artificial Intelligence and Law*, 15(1): 49-76.

Boer, A., R. Hoekstra, R. Winkels, T. van Engers, and F. Willaert (2002). "META*lex*: Legislation in XML", in T. Bench-Capon, Aspassia Daskalopulu, and R.G.F. Winkels (eds.), *Legal Knowledge and Information Systems (Jurix 2002)*, Amsterdam: IOS Press, pp. 1-10.

Boer, A., T. van Engers, and R. Winkels (2003). "Using Ontologies for Comparing and Harmonizing Legislation". *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, Edinburgh (UK).

EC-INSPIRE (2003). Spatial Data Infrastructures in Europe: State of play Spring 2003 commissioned by EUROSTAT & DGENV, K.U. Leuven.

Gibson, J. (1979). *The Ecological Approach to Visual Perception,* Hillsdale, New Jersey: Lawrence Erlbaum Associates Inc.

Hoekstra, R., Winkels R.G.F. and Hupkes, E. (2009). "Reasoning with spatial plans on the semantic web". *Proceedings of the Twelfth International Conference on Artificial Intelligence and Law (ICAIL 2009)*. IAAIL, ACM Press, June.

Human Factors Research Group (1994). SUMI. *Software Usability Measurement Inventory* Cork: University College.

INSPIRE (2008). Work Programme Transposition Phase 2007 - 2009, d.d. 16 mei 2007.

Peters, R. and T. van Engers (2004). The legal atlas: Map-based navigation and accessibility of legal knowledge sources, *Lecture Notes in Computer Science*, 3035: 228-236.

van de Ven, S., R. Hoekstra, R. Winkels, E. de Maat, and Á. Kollár (2007). MetaVex: Regulation Drafting meets the Semantic Web. *Lecture Notes in Computer Science,* 4884: 27-32.

Wilson, F. and R. Peters (2004). "Mapping the law: Knowledge support for business development enquiry". *Proceedings of the eChallenges 2004 Conference*, IST Programme, Vienna.

Winkels, R., A. Boer, and E. Hupkes (2007). "Legal Atlas: Access to Legal Sources through Maps". *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL 2007)*.

Woudenberg, B., R. Peters and M. Hoogwout (2009). Inspire begint vandaag; Wat de EU-Kaderrichtlijn INSPIRE voor provincies gaat betekenen. IOG Geo.

Wright (2004a). http://simcity.ea.com/community/events/will_wright_01_08_04.php.

Wright (2004b). http://simcity.ea.com/playsimcity_classic.phpor.

Power and Privacy: the Use of LBS in Dutch Public Administration

Charlotte van Ooijen and Sjaak Nouwt[1]

Tilburg Institute for Law, Technology, and Society (TILT), Tilburg University, the Netherlands, C.W.vanOoijen@uvt.nl

Abstract

There are in the policy fields of traffic management as well as public order and safety in the Netherlands new applications of location-based services (LBS) such as the public transport chip card and the use of mobile phone location data in policing. Combining citizens' location information and personal data is essential for the provision of LBS. We explored three cases of LBS in Dutch public administration and argue that LBS may affect the balance between the roles citizens can have in their relationship with government: subject, client and citoyen. Consequently, we discuss the concept of privacy in public places and relate this to European case law. It is important for government to be aware of the powerful inherent logic of LBS and how this may shape government-citizen interaction.

Keywords: Location-based Services (LBS), public administration, citizenship, privacy.

1. INTRODUCTION

"As Dutch citizens we are well taken care of by our state. Just look at our beautiful road infrastructure. Even the tiniest village in the outskirts of the country can be easily reached. Diverse access ways have been constructed into our big cities. As clients of our government we are entitled to use this well-maintained system, at a fair price of course, which is determined by our road tax system. And now the very good news is that a new pricing system is going to be developed which will be even more efficient and better tailored towards the individual situation of every citizen. How is this possible? Luckily, our government is always keen to look at new technological developments and investigate how we, as citizens, may benefit from these. Consequently, innovative policy makers have suggested to implement a satellite-based road pricing system which will be able to tax us based on our actual usage of the Dutch roads. So, we will only pay for the products we use. A more honest and fairly divided system can hardly be imagined, or can it?"

This could be the testimonial of a government promoted advertisement for the new Dutch road pricing system which is due to be implemented starting 2011 (Ministerie van Verkeer en Waterstaat, 2008). We would almost be inclined to forget that, as clients of our government, we are not just consumers like in the private market. Citizenship is shaped by and shapes itself through power structures as vested in our democratic institutions. Law, politics and administration as well as the civil society and the media determine the multi-faceted nature of the government-citizen relationship. Consequently, citizenship has to do with constitutional and democratic rights and duties along with mutual dependencies of state and society.

---

[1] From 1985 to February 2009, Sjaak Nouwt was an assistant professor at the Tilburg Institute for Law, Technology, and Society (TILT) of Tilburg University. Currently, he is a policy advisor at the Royal Dutch Medical Association.

The use of location-based information about citizens in public administration, such as the new road pricing system may affect the meaning of citizenship by shifting the information and the power relationship between government and citizens. Not only can LBS increase the governments' knowledge of who is where at what moment, the underlying technologies may also instigate the exertion of control of who goes where at what moment (Dobson and Fisher, 2003). Here, we address the inherent logic of LBS-applications in the context of Dutch public administration and aim at demonstrating the importance of the issue of privacy in public places.

In the next section we discuss in more detail the notion of citizenship in the information age. In the third section we focus on three LBS-applications in Dutch public policy, each of which are based on different positioning technologies. First, the introduction of a Radio Frequency Identification (RFID)-based public transport card is considered. Second, the plans for the new satellite-based road pricing system will be reviewed. Thirdly, we consider the use of mobile phone location data in crime fighting. We then present a model of the actors and data streams applying to location-based services in public administration and use this to reflect on the three cases. In section four of our contribution we discuss the concept of privacy in public places in relation to LBS as well as some relevant European case law.

2. CITIZENSHIP IN THE INFORMATION AGE

2.1 Citizens as subjects, clients and citoyens

Several scholars have argued that ICTs have the potential of altering the balance in the government-citizen relationship. The literature shows three extreme positions, which each indicate a different direction in which this relationship may evolve as a result of the ICT-revolution. The first two positions can be found in the Orwell-Athens debate as set out by Van de Donk and Tops (1992). The authors describe the scenarios of a powerful, Orwellian state on the one hand and that of democratic Athens-inspired society on the other. In each scenario different values regarding citizenship are dominant matching a different role the citizen can fulfil when interacting with government. In the Orwellian scenario, the citizen is predominantly treated as a *subject* of the state. Government has the power to set limitations on the behaviour of citizens and make sure that they obey the will of the state. The increased transparency of the citizen, caused by ICT, only helps government in steering and controlling society. Any deviant behaviour can be detected and acted upon using technological devices. Government thus uses ICT as a weapon to exercise power over citizens. In the opposing Athens scenario citizens seem to become more powerful thanks to ICT. They manifest themselves as *citoyens* in their relationship with government. In this role, citizens are treated as partners in the process of public policy making. Their opinions matter and now they can be consulted more easily than ever before thanks to ICT. The ideal of direct democracy is realised in Athens. As *citoyens* citizens have the right to get involved in public matters. In the Athens scenario, this right becomes reality.

Another debate in e-government literature shows us a third extreme position. Taylor et al. (2009) distinguish studies of the 'surveillance state' and studies of the 'service state'. The first type of studies are generally inspired by the fear of an Orwellian state, we already presented earlier as a first extreme position. The latter add a third perspective, which Frissen (1998) labelled 'Soft Sister'. As a Soft Sister, government emphasises the role of the citizen as a *client*. Government's top priority is to provide multiple and excellent services to its clients. As a *client,* the citizen expects and demands a certain service level when interacting with government. ICT opens new ways to cater

for these needs by increasing the quantity and quality of public services as well as tailoring these to the desires of the individual *client*. Government uses ICT as a tool to act in favour of citizens. At the same time, Soft Sister appears to be a different presentation of Big Brother.

When we use government services, there is no option to choose a different service provider. Moreover, usually we cannot withdraw from using government service because it is obligatory by law. In the Netherlands, for example, all citizens above the age of fourteen must, when asked by a police officer, identify themselves with an official identification document. In order to comply with this duty, Dutch citizens are obliged to use a government service to provide them with documents like passports or ID-cards. We are not clients of government in the same sense as being customers of private enterprises, because we are not free to choose whether we want to use a service or who we want to enjoy the service from. Consequently, we are always both *subjects* and *clients* of the state at the same time. Therefore, a holistic perspective in both research and practice of information-intensive government is highly desirable (Taylor et al., 2009).

2.2 The conscience of technology

The three citizen roles of *subject*, *citoyen* and *client* each emphasise different aspects of citizenship which are all important for governance in a democratic society. The three scenarios of Orwell, Athens and Soft Sister are extreme positions, which show the conflicting values government has to deal with when using ICT in interactions with citizens. Government ought to respect all three citizen roles and address citizens as such. However, we cannot count on technology itself to make the proper judgement and adapt its functionalities to the context it is used in. In other words, citing Davis (2003, in Michael et al., 2008), "technology has no conscience of its own". This does not imply, however, that technology would be a neutral tool, serving as a mere means to reach governmental goals. We would rather state that the characteristics of a particular technology convey an inherent logic which ought to be taken into account when applying it in interactions with citizens. What then, is the inherent logic of location-based services? This question will be explored in the next section by discussing three applications of LBS in Dutch public administration. We will demonstrate that despite the different underlying technologies, a similar pattern can be distinguished revealing the inherent logic of LBS.

3. LOCATION-BASED SERVICES IN DUTCH PUBLIC ADMINISTRATION

3.1 Three cases of LBS in Dutch public policy

3.1.1 Satellite technology for road pricing

In December 2007, the Dutch Ministry of Transport, Public Works and Water Management announced the Cabinet's decision to implement a new pricing system for the use of public roads (Ministerie van Verkeer en Waterstaat, 2007). According to the plan, by 2012, car drivers will be charged a price per kilometre. To implement this new system, the Dutch government will be using the latest satellite technology to collect location information about every car. Even though the legal, political and technological specifications have not yet been entirely determined, it is evident that the gathered location information will need to be connected to personal data in order to be able to send the right bill to the right person. Of course, the underlying report mentions that considerations about people's privacy will be taken into account when developing the system. Already, a private company has offered the responsible Minister a technological solu-

tion which does not measure a person's exact route, but just the number and type of roads he or she drives on instead, thereby realising a lesser invasion of privacy (Pieper, 2007). Nevertheless, a rich database with up-to-date, accurate, precise information about activities on the Dutch roads will be available.

3.1.2 RFID in public transport

The introduction of a new chip-based system for public transport is a politically sensitive topic. The Dutch government has decided to push forward the plan to implement one public transport card which can be used nation-wide on the infrastructures of all suppliers (Teepe, 2008). In this system, at the start of a journey the traveller checks in by bringing the RFID-enabled card close to the card reader. At the place of destination, during check-out the appropriate fee will be charged to the electronic wallet in the card. One reason for implementing this system is to get a more honest distribution of income between the various public transport companies. At this moment, the system has already been implemented by local public transport providers in parts of the Netherlands.

3.1.3 Mobile phone localisation in criminal investigation

The local law enforcement of the Dutch city of Nijmegen used cell phones in March 2006 to find possible witnesses to a criminal act. Three-thousand people received a text message asking them to contact the authorities if they could provide information about the murder of activist Louis Sévèke (Nu.nl, 2006). These people were not selected because they lived near the crime scene or because they were acquainted with the victim. They were merely selected because of their location or, to be more precise, the location of their cell phones at the time of the murder. This technological possibility instigated their involvement in the police investigation. The local law enforcement requested these data from the telecom providers and consequently, was able to send the request for information to the given phone numbers, resulting in several reactions. Little is known, however, about the psychological and social effects of this type of message. What were reasons to respond or not respond to this police call? Did people feel curious, scared, spied on or perhaps important or even appreciated? More research into this matter is desired in order to be able to assess the wanted and unwanted consequences of this kind of LBS-application.

3.2 LBS logic

3.2.1 LBS actors and data streams

Location-based services are "IT services for providing information which has been created, compiled, selected or filtered taking into consideration the current locations of the users or those of other persons" (Küpper, 2005). Following this definition, generating information is the essence of LBS. Different positioning technologies can be used to do this. It follows that the LBS-user receives information based on his or her own location or that of other individuals. In the cases described in the previous section a government organisation places itself in the role of the LBS-user receiving information based on the location of citizens. For the technical realisation of LBS private parties play an important role as well. In addition, data streams from several other actors are also required for LBS. Based on Küpper's LBS supply chain (2005), Figure 1 is a possible visualisation of this situation.

78

Figure 1: LBS actors and data streams.
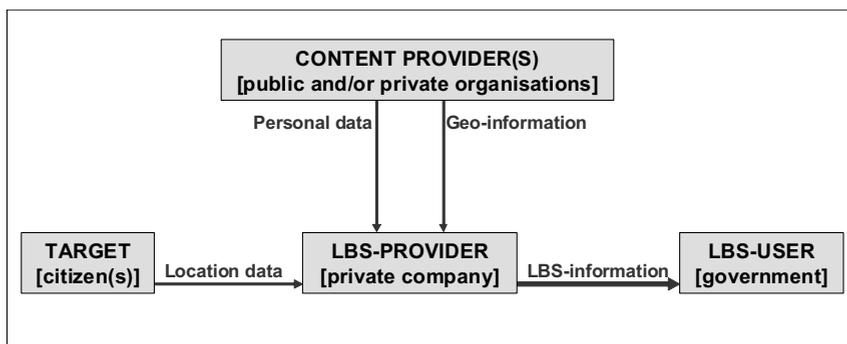
3.2.2 Citizens as targets

In this LBS supply chain, citizens are labelled as the *target*. The target is the actor whose position is determined by means of a positioning device. Technologically speaking, this actor is the starting point for LBS from the moment *location data* are generated. In the case of the public transport card information is transmitted as soon as the traveller's card finds itself near a RFID card reader. At that point, the card ID, the location of the reader and the time is first stored in a database belonging to the particular public transport provider and subsequently copied to a central database, thus containing the location data coming from all providers (Teepe, 2008). The road pricing system will use drivers' coordinates obtained from a satellite receiver placed in the vehicle. At this point it is uncertain whether these data will be stored in a device inside the vehicle or in an external database (Hoepman, 2008). In the Sévèke case location data in the shape of cell IDs were obtained from the potential murder witnesses as soon as their mobile phones were connected to the antennas in the area and these data were stored in the telecom providers' databases.

3.2.3 Personal data and geo-information as content

Another important data stream in the LBS supply chain consists of *personal data* about the 'targeted' citizens. In all three cases this is a relevant data stream. Regarding the public transport card, we see that people who currently have subscriptions or discount cards for their train or bus journeys automatically receive so called a personalised public transport card. This card's unique ID is linked to the subscribers' personal data as stored in, for example, the Dutch Railways' customer database. This organisation can thus serve as a *content provider* in the LBS chain. The argument is that this information is needed for authentication purposes (Teepe, 2008). A particular person's rights (to a discount) have to be linked to a particular card, literally opening the gates to the section that person is allowed to travel. In the case of road pricing, a similar reasoning is applicable, only in this case connecting a location device to a person's road tax paying duties. The third LBS case also requires personal data; this time it is a phone number so that they can be approached by a text-message. The telecom providers are the content providers of these data.

Other content providers, like Tele Atlas or Google may provide a stream of geo-information about the target's surroundings. Currently we do not know whether this will be the case in the examples mentioned here. However, we are aware of government and commercial LBS-applications where the use of GIS and geo-information plays an

important role (Dobson and Fisher, 2003; Ahas and Mark, 2005; Raper et al., 2007). GIS containing maps of the Rotterdam metro system, the Dutch road infrastructure and the neighbourhood of the murder could be the geo-information streams used in our examples.

3.2.4 Providing LBS-information to the government

The *LBS-provider* is the actor who integrates the aforementioned data streams. This actor collects location data about one or more targets, makes spatial analyses and combines this with other (geographical) data. The produced LBS-data are sent to the *LBS-user*. As such, raw location data are translated and enriched into LBS-information which can be read by the LBS-user. At this stage the previously collected personal data can be anonymised or pseudonymised and the location data may be aggregated if the LBS-user has no need for personalised detailed location data. This is the plan regarding road pricing. Even though the technological details on how to create the desired LBS-information are yet unknown, it is clear that the Dutch government has expressed its interest in this kind of aggregated data (Hoepman, 2008). As far as the public transport card is concerned, government interest in the collected data may come down to anti-terrorist or for crime fighting purposes (Van 't Hof, 2007; Teepe, 2008). However, it is yet unknown whether actual interest in this direction has been shown. The telecom providers in the murder case play the role of content provider as well as LBS-provider. The list of phone numbers belonging to the mobile phones which were localised in the neighbourhood of the crime scene around the time of the murder was obtained from LBS-information which the Nijmegen police force (LBS-user) received.

3.2.5 LBS-information: more and less

It is important to emphasise that LBS-information does not consist of the sum of the relevant location data, personal data and geo-information. It can be more and less at the same time. More, because of the added value to the government as the LBS-user, opening possibilities for analyses of large quantities of data and monitoring of citizens (Snellen, 2000). The result may also be less, because government may end with less detailed location information than initially collected or with (pseudo)anonymised data. In this respect private parties who play the role of LBS-provider are key actors because of the technological realisation of LBS. At the same time, government holds the power and responsibility to arrange for appropriate regulation deciding in which situations the LBS-information will be more or less than the sum of data streams.

3.3 Conclusion

We've seen that the core characteristics of LBS consist of collecting and synthesising data about citizens. These data are location data, personal data and geo-information connected to citizens. The possibilities the underlying technologies offer seem to be tempting to government organisations, especially in the field of intelligence. Whether government decides to use LBS-information to provide new services, create new democratic arrangements or monitor and control citizens, in all cases the initial data streams need to be acquired. Citizens always need to be targeted first in order to potentially benefit from new services or citizen participation. Looking at the aforementioned Sévèke murder case, it seems that the government wants to cooperate with the citizens – those who are witnesses – to solve the crime. Therefore, the government has a relationship with the citizen-citoyen. However, the same government also has the powers to relate the obtained information to possible suspects, for example when citizens do not cooperate for whatever reason. In that case, the government has a rela-

tionship with the citizen-subject. In other words, citizens are vulnerable in principle when LBS is concerned and find themselves in a weak position towards both the involved private parties and government. Therefore, at the starting point of LBS citizens are placed in the role of subject of the state. The government has legal powers to obtain location-based information about citizens. Obviously governments should be careful when collecting and using location-based information about citizens because it puts citizen trust in government at stake. Legal norms concerning a person's right to privacy in public places and data protection principles can help guiding governments through this pitfall.

4. PRIVACY IN PUBLIC PLACES

4.1 Definition of 'public place'

What distinguishes a private place from a public place seems obvious, but in 2004 this led, at least in the Netherlands, to discussions in parliament when the Bill on Camera Surveillance in Public Places was discussed (Kamerstukken 2004/5a). The (present) Act is only applicable to camera surveillance for the prevention of public disorder in municipalities. The explanatory memorandum defines 'public place' as "a place that is open to the public, according to its function or regular use". 'Open to the public' means that there are no barriers to enter the place, like a duty to report, preceding permission, or levying an admission ticket. As a result, stadiums, post offices, department stores, restaurants, and hospitals are in this respect not considered as public places.

'Function' refers to the nature given to the place. The nature of a place may follow from a decree or from the purpose that follows from the functionality of the place. A place becomes a public place through 'regular use' when this is used for this purpose, and the rightful claimant allows the place being used as such. Therefore, a public place is a place where people come and go, like for example:

− the street;
− the (public) road and in the continuation thereof:
    o public gardens;
    o playing fields;
    o parks, and
    o open sections of indoor shopping centres and arcades.

Shops, discotheques, parking garages, town halls, churches and mosques, public sections of a railway stations (if private property) are private, not public places.

4.2 The opinion of the Dutch government

In the discussion of the Dutch Camera Surveillance Bill, the Christian Democratic Party (CDA) stated that, in their opinion, it is impossible to have a right to privacy in a public place because whoever exposes themselves in a public place relinquishes the right to see this as a private part of their lives. Therefore, there is no question of interference of an individual's private life in a public place (Kamerstukken, 2004/5b).

As we will show hereafter, this opinion obviously differs from that of the European Court of Human Rights (ECtHR), but also from the opinion of the Dutch government. According to the Dutch government, the right to privacy is not spatially limited. The government refers to a judgment by the Dutch Supreme Court in 1991, about the seizure of videotapes from a public demonstration (Hoge Raad, 1991), and concludes that

camera surveillance on a public road can interfere with the right to one's private life. However, according to the Dutch government, the more public a citizen's behaviour is, the less the right to privacy will be an issue. So, according to the Dutch government, a citizen's behaviour can be less or more public. This also means that a citizen can expect less or more privacy.

4.3 Article 8 ECHR

The right to privacy protects our 'private and family life, home, and correspondence' (Article 8, paragraph 1 European Convention on Human Rights). These are the basic elements of our privacy. Therefore, it seems that the right to privacy is especially applicable to private places. However, nowadays these private places are not so private anymore: we store a lot of our personal data on our personal hard disks, laptops, Blackberries, or iPods; also a lot of personal information is stored on the servers of our Internet service provider or on Google's servers. It seems that Big Brother is not only watching us, but he also knows where we are, where we have been and probably even where we are going to. Thanks to the technical solutions for large scale collection and analysis of personal data, including geo-information (location data, whereabouts), law enforcement agencies can compare these data with so called risk profiles. As a result, the privacy of citizens (subject, client and citoyen) will come under pressure because they are becoming more transparent to law enforcement and intelligence agencies. It also enhances the risk of mistakes being made because criminal investigations could then be extended to cover everyone. There is a big difference between legitimising the preventive monitoring of everyone and the limited application of a means of coercion against specific suspects (Vedder et al., 2007).

The second paragraph of Article 8, ECHR allows public authorities to interfere with the exercise of the right to privacy of an individual. Such interference is only allowed on two conditions, namely that (1) interfering is in accordance with the law, and (2) interfering is necessary in a democratic society in the interests of:

– national security;
– public safety;
– the economic well-being of the country;
– the prevention of disorder or crime;
– the protection of health or morals, or
– the protection of the rights and freedoms of others.

An interference is in accordance with the law when it is allowed by legislation or by case law, as long as it is transparent for the citizen. An interference is necessary in a democratic society when there is a pressing social need to reach a certain goal while interfering with the right to privacy. "Necessary" means that the measure is appropriate to reach that goal (proportionality) and that no alternative measures are available that could also be appropriate (subsidiarity).

In the following subsection, we will analyse how the European Court of Human Rights (ECtHR) has recognised the right to privacy in public places. We will discuss two important cases in this respect: *Rotaru v. Romania* (ECtHR, 2000) and *P.G and J.H. v. The United Kingdom* (ECtHR, 2001). Other cases with regard to privacy in public places are e.g. *Peck v. The United Kingdom* (ECtHR, 2003a) and *Perry v. The United Kingdom* (ECtHR, 2003b).

4.4 European case law

4.4.1 Rotaru v. Romania

In 2000, the European Court of Human Rights passed an important judgment on the difference between private and public places in the case of *Rotaru v. Romania* (ECtHR, 2000). In this case, the ECtHR confirmed their earlier judgments by recognising that information about the applicant's life, in particular his studies, his political activities and his criminal record, when systematically collected and stored in a file held by agents of the State, falls within the scope of 'private life' for the purposes of Article 8, ECHR (Ibid, § 44.). The Court disagreed with the Romanian government that this information is related to the applicant's public life, and therefore did not fall within the scope of 'private life'. With regard to public information that can fall within the scope of the right to private life, the Court made an interesting remark:

"Moreover, public information can fall within the scope of private life where it is systematically collected and stored in files held by the authorities. That is all the truer where such information concerns a person's distant past" (Ibid, § 43).

The Court recognised that a right to privacy exists when a government agency systematically collects and stores personal information, even when this is public information.

4.4.2 P.G. and J.H. v. The United Kingdom

In the case of *P.G. and J.H. v. The United Kingdom* (ECtHR, 2001), the Court dealt with the scope of privacy in public places. The applicants complained that covert listening devices were used by the police to monitor and record their conversations in an apartment, that information was obtained by the police concerning the use of a telephone at the apartment, and that, while they were at the police station, listening devices were used to obtain voice samples.

The most relevant domestic law consisted of the Telecommunications Act 1945 and the Data Protection Act 1984. Section 45 of the Telecommunications Act prohibits the disclosure by a person engaged in a telecommunications system of any information concerning the use made of the telecommunications services provided for any other person by means of that system. However, section 28(3) of the Data Protection Act 1984 reads: "Personal data are exempt from non-disclosure provisions in any case in which – (a) the disclosure is for any of the purposes mentioned in subsection 1 above; and (b) the application of those provisions in relation to the disclosure would be likely to prejudice any of the matters mentioned in that subsection." Subsection 1 refers to data held for the purpose of: "(a) the prevention or detection of crime; (b) the apprehension or prosecution of offenders; or (c) the assessment or collection of any tax or duty." In this case, the Court concluded that the disclosure to the police was permitted under the relevant statutory framework where necessary for the purposes of the detection and prevention of crime (Ibid, § 47).

However, in the Court's opinion, there is also an area, even in public space, where people may have interactions, which are protected by the right to privacy:

"There is therefore a zone of interaction of a person with others, even in a public context, which may fall within the scope of 'private life'" (Ibid, § 56).

Furthermore, the Court gave a number of elements that are relevant to the consideration of whether a person's private life is concerned by measures effected in public places:

"Since there are occasions when people knowingly or intentionally involve themselves in activities which are or may be recorded or reported in a public manner, a person's reasonable expectations as to privacy may be a significant, although not necessarily conclusive, factor. A person who walks down the street will, inevitably, be visible to any member of the public who is also present. Monitoring by technological means of the same public scene (for example, a security guard viewing through closed-circuit television) is of a similar character. Private-life considerations may arise, however, once any systematic or permanent record comes into existence of such material from the public domain. It is for this reason that files gathered by security services on a particular individual fall within the scope of Article 8, even where the information has not been gathered by any intrusive or covert method (…)" (Ibid, § 57).

The Court concluded that the recording of the voices of the suspects at the police station was an interference with their right to respect for private life. In this case, the Court recognised that personal information collected in a public place, falls under the scope of the right to privacy when this information has been collected and stored systematically, for example by a government agency. This conclusion can also be applied to geo-information, when that information is related to an identified or identifiable natural person. Systematically collecting, storing, and analysing geo-information must be considered an interference with the right to privacy of the individual. The next question is whether the interference is legitimate.

4.5 Data protection principles

In the context of LBS, most of the time location data will be collected that can be related to identified or identifiable citizens. In such cases, the data protection legislation will be applicable. Apart from the question whether citizens can have a reasonable expectation of privacy, the 'controller' (in this case: the government) will have to comply with these data protection rules. The data protection framework is based on a number of general data protection principles. The basic data protection principles were formulated in the OECD Privacy Guidelines in 1980 and in Council of Europe Data Protection Treaty (Convention 108) in 1981. These traditional data protection principles still determine the framework for the fair and lawful processing of personal data, also with regard to LBS. The following general principles can, for example, be found in the Dutch Personal Data Protection Act (Hooghiemstra and Nouwt, 2007):

– processing of personal data must be fair and lawful and in accordance with the law;
– personal data are collected only for specified, explicit and legitimate purposes;
– processing of personal data must be based on legitimate legal grounds (e.g. the consent of the data subject or necessary for the performance of a contract);
– further processing of personal data must be compatible with the purposes for which the data were originally collected;
– personal data shall be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed;
– the controller must implement appropriate technical and organisational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access;
– personal data are kept no longer than is necessary for the purposes for which they were collected or for which they are further processed.

Without prejudice to the general data protection legislation, it is possible that special data protection legislation is applicable, for example for police data. However, this has no influence on the applicability of the general data protection principles.

Technical possibilities for large scale collection and analysis of personal data, including spatial data and telecommunications data make it much easier, for example, for law enforcement agencies to compare these data with so called risk profiles. As a result, the privacy of groups of citizens is at stake because they are becoming more transparent to law enforcement and intelligence agencies. It also enhances the risk of mistakes being made because criminal investigations could then be extended to cover everyone (Nouwt, 2008; Vedder et al., 2007).

5. CONCLUSIONS

Legal norms are important for governments to demarcate the borders for collecting and using location-based information about citizens without interfering with their right to privacy. For some politicians, it is obvious that citizens have less reasonable expectations of privacy in public places. However, from ECtHR case law, we can conclude that the right to privacy also exists in public places where citizens can be monitored and information about them can be collected. From a legal perspective, governments are only allowed to collect location-based information about citizens when the powers to do so are in accordance with the law and there is a pressing social need to collect this information. Furthermore, collecting and further processing of personal data must be in accordance with the general data protection principles. Legal guidelines, however, are not sufficient to give direction to socially desirable applications of LBS in public policy (Onsrud, 2008). Governments should be critical towards the policy and societal goals they wish to attain by using LBS. When interacting with citizens, they should be aware of the conflicting values of the subject, citoyen and client role in order to avoid the extremes of Orwell, Athens and Soft Sister.

REFERENCES

Ahas, R. and U. Mark (2005). Location Based Services-New Challenges for Planning and Public Administration? *Futures*, 37(6): 547-561.

Dobson, J.E. and P.E. Fisher (2003). Geoslavery - Society Must Contemplate a New Form of Slavery, Characterized by Location Control, *IEEE Technology & Society Magazine: a Publication of the IEEE Society on Social Implications of Technology*, 22(1): 47-52.

Council of Europe (1981), Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, at http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=108&CM=8&DF=8/27/2008&CL=ENG.

Davis, B. (2003). "Technoism: will loss of freedom unleash the voice of dissent?", *International Symposium on Technology and Society: Crime Prevention, Security and Design, September 26-28 2003,* pp. 19-24.

ECtHR (2000). Rotaru v. Romania: European Court of Human Rights, Judgment of 4 May 2000, no. 28341/95, at: http://www.echr.coe.int/echr/.

ECtHR (2001). P.G. and J.H. v. The United Kingdom: European Court of Human Rights, Judgment of 25 September 2001, no. 44787/98, at: http://www.echr.coe.int/echr/.

ECtHR (2003a). Peck v. The United Kingdom: European Court of Human Rights, Judgment of 28 January 2003, no. 44647/98, at: http://www.echr.coe.int/echr/.

ECtHR (2003b). Perry v. The United Kingdom: European Court of Human Rights, Judgment of 17 July 2003, no. 63673/00, at: http://www.echr.coe.int/echr/.

Frissen, P.H.A. (1998). "Public Administration in Cyberspace: A Postmodern Perspective", in Snellen, I.Th.M. and W.B.H.J. van de Donk (Eds.), *Public Administration in an Information Age. A Handbook*, Amsterdam/Berlin/Oxford/Tokyo/Washington DC: IOS Press, pp. 33-46.

Hoepman, J.-H. (2008). Follow that Car! Over de Mogelijke Privacygevolgen van Rekeningrijden, en hoe die te Vermijden, *Privacy en Informatie*, 11(5): 225-230.

Hoge Raad (Dutch Supreme Court), 19 February 1991, *NJ* 1992, 50.

Hooghiemstra, T. and J. Nouwt (2007). *Tekst en Toelichting Wet Bescherming Persoonsgegevens*, Den Haag: Sdu.

Kamerstukken (2004/5a). "Wet tot Wijziging van de Gemeentewet en de Wet politieregisters in verband met de invoering van regels omtrent het gebruik van camera's ten behoeve van toezicht op openbare plaatsen" (*camera surveillance in public places*). *Kamerstukken II*, 2004/05, 29 440.

Kamerstukken (2004/5b). *Kamerstukken II*, 2004/05, 29 440, nr. 6, p. 10.

Küpper, A. (2005). *Location-based Services: Fundamentals and Operation,* Chichester: Wiley.

Michael, M.G., S.J. Fusco and K. Michael (2008). A Research Note on Ethics in the Emerging Age of Überveillance, *Computer Communications*, 31(6): 1192-1199.

Ministerie van Verkeer en Waterstaat (2008). Wet op de Kilometerprijs naar Raad van State, at: http://www.verkeerenwaterstaat.nl/actueel/nieuws/wetopdekilometer prijsnaarraadvanstate.aspx.

Nouwt, J. (2008). Reasonable Expectations of Geo-Privacy? *SCRIPT-ed*, 5(2): 375-403.

Nu.nl (2006). Politie Blij met Resultaat SMS-Bom Sévèke. *Nu.nl/internet*, at http://www.nu.nl/news.jsp?n=697945&c=50&rss.

OECD (1980), Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, at: http://www.oecd.org/document/18/0,3343,en_2649_34255_18151 86_1_1_1_1,00.html.

Onsrud, H.J. (2008). Implementing Geographic Information Technologies Ethically, *ArcNews Fall 2008 Issue*, at: http://www.esri.com/news/arcnews/fall08articles/implementing-gi-technologies.html.

Pieper, R. (2007). Anders Betalen voor Mobiliteit met Mobimiles. Brief aan Minister Eurlings 18 September 2007, ROAD Group.

Raper, J., G. Gartner, H. Karimi (2007). A Critical Evaluation of Location Based Services and their Potential, *Journal of Location Based Services*, 1(1): 5-45.

Snellen, I.Th.M. (2000). Territorialising Governance and the State: Policy Dimensions of Geographic Information Systems, Information Infrastructure and Policy: an International Journal on the Development, Adoption, Use and Effects of Information Technology, 6(3): 131-138.

Taylor, J.A., A.M.B. Lips and J. Organ (2009). Identification Practices in Government: Citizen Surveillance and the Quest for Public Service Improvement, *Identity in the Information Society,* 1(1).

Teepe, W.G. (2008). In sneltreinvaart je privacy kwijt, *Privacy en Informatie*, 11(5): 217-224.

Van de Donk, W.B.H.J. and P.W. Tops (1992). "Informatisering en Democratie: Orwell of Athene?", in Frissen, P.H.A., A.W. Koers and I.Th.M. Snellen (Eds.), *Orwell of Athene?: Democratie en Informatiesamenleving*, Den Haag: Sdu, pp. 31-74.

Van 't Hof, C. and R. van Est (2007). *RFID: Meer Keuze, Gemak en Controle in de Digitale Publieke Ruimte*, Den Haag: Rathenau Instituut.

Vedder, A., J.G.L. van der Wees, B-J. Koops and P. de Hert (2007). *Van Privacyparadijs tot Controlestaat? Misdaad- en Terreurbestrijding in Nederland aan het Begin van de 21ste Eeuw*, The Hague: Rathenau Instituut, Study 49 (summary in English), at: http://www.rathenau.nl/showpageBreed.asp?steID=1& ID=3814.

Harmonising and Integrating Two Domain Models Topography

Jantien Stoter[1], Wilko Quak[2] and Arjen Hofman[3]

[1]ITC, the Netherlands, stoter@itc.nl
[2]TU Delft, the Netherlands, c.w.quak@tudelft.nl
[3]Logica, the Netherlands, arjen.hofman@logica.nl

Abstract
This article presents a case study on harmonising and integrating two domain models topography into a global model that have been established for different purposes; IM-Geo defines topography to serve municipalities in maintaining public and built-up area; TOP10NL defines topography for visualisation at 1:10k scale. The study identifies problems and proposes solutions to accomplish integration, which is required to provide the datasets within the principles of the national Spatial Data Infrastructure. At first the types of differences between the current models are analysed. Secondly, the article formulates recommendations to harmonise the differences which may be random (i.e. easy to solve) or fundamental (to be addressed in the integration). Finally, the article presents modelling principles for an integrated model topography based on two conclusions of the comparison study: two domain models are necessary to meet the specific demands of the two domains and secondly, TOP10NL cannot be derived from IMGeo because differences in perspective proved to be more dominant than scale differences did. Both the recommendations for harmonisation and the modelling principles are illustrated with prototypes which show the problems and potentials of harmonising and integrating different local (national) data models into global models.

Keywords: Information modelling, data harmonisation, integrating domain models, information model topography.

1. INTRODUCTION

An important objective of the INSPIRE Directive is to reduce duplicated data collection (INSPIRE, 2007). An absolute necessity for 'collecting data once, and use it many times' is harmonising specifications of datasets to fully integrate data from various sources. This is both valid for different datasets covering one state, but also for datasets of different states that touch at borders in order to "ensure that spatial data relating to a spatial feature the location of which spans the frontier between two Member States are coherent. Member States shall, where appropriate, decide by mutual consent on the depiction and position of such common features" (INSPIRE, 2007).

EuroGeographics has responded to the INSPIRE Directive by launching the EuroSpec project (EuroSpec, 2009). This project is the collective contribution of the National Mapping Agencies (NMAs) to build the European Spatial Data Infrastructure (ESDI), in line with the concepts of INSPIRE. An important activity of the project focuses on harmonised pan-European and cross-border specifications for large scale topographic data that goes beyond the successes of harmonised pan-European small scale products, such as EuroGlobalMap (scale 1:1million) and EuroRegionalMap (scale 1:250k) for topography and EuroBoundaryMap (scale 1:100k) for administrative units (Euro-Geographics, 2009).

Although it may seem straightforward that topographic datasets within and between countries will contain similar types of objects, many (small) differences occur between

these datasets. Afflerbach et al. (2004) studied the differences between four national topographic datasets within the context of the GiMoDig project (Geospatial Info-Mobility Service by Real-Time Data-Integration and Generalisation (GiMoDig, 2009): Germany, Finland, Sweden and Demark and found many differences. For example, different geometries for similar concepts (e.g., road centre lines representing individual lanes or road centre lines representing the whole road construction); differences between classifications for water (different types of water), for roads (using either widths or official administrative categories as classification criteria) and for recreational areas (which may have been further classified into 'amusement park', 'campground', 'parks', but not in all datasets); different minimum size criteria to collect area objects such as parks and forests; different collection criteria for hydrographical networks resulting in different densities for same types of hydrography.

Because of these differences in the definition of concepts several problems occur. Firstly, national, local data models are not easy to map to global (e.g., European) data models that realise the harmonisation without information loss. Secondly, a question such as "what is the total forest coverage of Europe?" is not easy to answer, because different conditions are used to identify an area with trees as forest. For instance what minimum density of trees is required to identify forest? What is the minimum dimension of the area to identify it as forest? What are, apart from presence of trees, criteria to identify forest, i.e. the function of the area (recreation or hunting), the maintenance characteristics of the area or the type of land-cover?

Besides cross-border differences between topographic datasets which is caused by lacking agreement of the national topographic data producers, differences may occur between topographic datasets covering the same country. The reason for these differences is also originating from different data organisations being responsible for producing and maintaining the different datasets with their own goals in mind. For example, municipalities collect and maintain large scale topographic data in support of the management of public and built-up area, while National Mapping and Cadastral Agencies (NMCAs) collect and maintain topographic data for the same area to represent maps at different scales (also at large scale). The last few decades these datasets are being translated into object oriented datasets to support database applications and GIS analysis. Key question for providing these different datasets within an Spatial Data Infra Structure (SDI) how they relate to each other to enable collecting data of same objects once in the future.

This article studies the feasibility of harmonising and integrating two independently established information models topography, expressed in UML (Unified Modelling Language) class diagrams. UML diagrams are often used to describe the content and meaning of datasets. The term 'harmonising' is used in this article as 'agreeing on thematic concepts' and 'integrating' as 'defining how objects in one dataset can be derived from objects in another dataset'. The article will provide insight into generic problems and solutions to accomplish such harmonisation and integration.

The case study contains two datasets representing topography at different scales for different purposes in the Netherlands. For both datasets information models have been established that describe the content and meaning of the data. The first dataset is the object oriented Large-scale Base Map of The Netherlands (*Grootschalige Basiskaart Nederland*: GBKN; LSV GBKN, 2007) defined in the Information Model Geography (IMGeo, 2007). Municipalities are the main providers (and users) of this dataset. The second dataset is the topographic dataset at scale 1:10k provided by the Netherlands' Kadaster and defined in the TOP10NL information model (TOP10NL, 2005). Harmonis-

ing and integrating these two domain models have become an important issue now 'key registers' are being established to support the Dutch SDI (VROM, 2009a). Legally established key registers contain authentic base data and their use is mandatory for all public organisations. As long as harmonisation and integration is not realised, municipalities and the Kadaster need to collect data of the same objects in parallel to meet the requirements of the specific domain in which they operate. This requires two key registers topography:

− *Basisregistratie Grootschalige Topografie* (BGT), 'key register large scale topography'. The feasibility of BGT as key register defined in IMGeo is being studied (VROM, 2009b);
− *Basisregistratie Topografie* (BRT), 'key register topography', in force since 2008. Currently BRT only contains topographic data at scale 1:10k. From 2010 the smaller scales will be added to this register (VROM, 2009c).

This article aims at formulating recommendations and proposing modelling principles, illustrated with UML examples, for an integrated information model topography that serves both domains. This integration at conceptual level can be used to move towards 'collect once, use many times' in the future. Although the case study is limited to the Netherlands, it presents common needs and problems for harmonising core national topographic databases within and across countries as well as solutions to establish global (e.g. European) data models.

Section 2 identifies the differences that need to be addressed in the integrated model topography. These differences are categorised based on the results of Hofman et al. (2008) and Stoter (2009), who studied the differences and commonalties between the two information models in detail. Section 3 proposes the integrated model topography based on results and conclusions of the comparison study. The proposal consists of two parts: a) recommendations to harmonise differences as much as possible, b) modelling principles for the integrated model topography. For both parts representative solutions are presented not limited to the case study. Consequently the proposed solutions can serve as recommendations and modelling solutions for harmonising and integrating information models established for different purposes. The article ends with conclusions in section 4.

2. DIFFERENCES TO BE ADDRESSED

Both IMGeo and TOP10NL are domain models extending the abstract data model NEN3610. The ISO compliant version of NEN3610 (Basismodel Geoinformation) was established in 2005 (NEN3610, 2005). This data model provides the concepts, definitions, and relations for objects which are related to the earth surface in the Netherlands. Domain models extend NEN3610 by defining their classes as subclasses of the NEN3610 GeoObject. Therefore these classes inherit all properties of the NEN3610 GeoObject. Examples of domain models are information model for physical planning (IMRO), information model for cables and pipelines (IMKL), information model for soil and subsurface (IMBOD), and information model for water (IMWA) (Geonovum, 2008). ISO19109 defines such a domain model as "application schema" (ISO, 2005): 'a conceptual schema for data required by one or more applications'.

The idea behind NEN3610 and the extended domain models was that inheritance of the same NEN3610 GeoObject would assure harmonisation of the domain models. However, when comparing the information models TOP10NL and IMGeo, which both extend NEN3610, we observe that many differences need to be addressed before inte-

gration can be realised. This section lists the types of differences to be addressed which is a result of analysing the studies of Hofman et al. (2008) and Stoter (2009). The types of differences that we will present in this section are:

− Differences in perspective (section 2.1);
− Differences in main classes (section 2.2);
− Differences in object demarcation (section 2.3);
− Differences in attribute values (section 2.4);
− Different classes for same concepts (section 2.5);
− Same attribute name for different concepts (section 2.6);
− Differences in amount of information (section 2.7), and
− Differences in class definitions (section 2.8).

2.1 Difference in perspective

IMGeo and TOP10NL model the same geographic extent and same types of objects from a different perspective. The difference in perspective is due to differences in objectives, source data, scale, application domain, providers, acquisition method and rules, see Table 1. These differences have resulted in different contents of the datasets.
An example is how topology is implemented in the datasets (see Table 1). Terrain, water and road objects in TOP10NL that are visible from above form a planar partition (i.e. no overlap or gaps); whereas IMGeo models the planar partition at ground level. Consequently in IMGeo objects can be located above the planar partition (indicated with relativeHeight > 0). In contrast, in TOP10NL no objects can be located above the planar partition: objects with heightlevel=0 are located at ground level or on top of a stack, for example in case of infrastructural objects at crossings, *and* they are part of the planar partition.

2.2 Differences in main classes

Table 2 lists the main, non-abstract classes that occur in either IMGeo, TOP10NL or in both models. Also the corresponding NEN3610 classes are listed. As can be seen in the table a few classes start with 'part of'. This is to model the division of whole objects into several geometries in an object oriented approach.
A comparison of the main classes provides the following insights. Six classes occur in both models. In addition, Geographical Area (used to link toponyms to objects), Functional Area (used to group objects of different classes) and Relief are only modelled in TOP10NL (Relief not available in NEN3610). Furthermore, IMGeo and NEN3610 distinguish Engineering Structure which is not available as separate class in TOP10NL. Another observation concerns the classes related to buildings. NEN3610 models Building Complex (*Gebouw*), Building (*Pand*) and Living Unit (*Verblijfsobject*). IMGeo only models Building and Living Unit in accordance with the Building and Address Register (BAG, 2006). TOP10NL only models Building Complex, which also includes single buildings when they are larger than a minimum size.

A final observation of comparing the main classes, is a similar granularity of NEN3610 classification at the one hand and IMGeo and TOP10NL at the other hand, i.e. they contain more or less the same number of classes. However, since IMGeo and TOP10NL extend the abstract data model NEN3610 to define content of specific datasets, one would expect refinement of the classes, i.e. more classes in the domain

Table 1: Differences in background between IMGeo and TOP10NL.

	IMGeo	TOP10NL
Objectives	enabling and standardising exchange of object oriented geographical information, IMGeo should be a framework of concepts for all organisations that collect, maintain and disseminate large scale geographical information	object oriented semantic description of the terrain for TOP10vector, according to requirements of internal and external users of the TOP10vector dataset
Source data	Object oriented GBKN	TOP10vector
Scale	1:1k in urban area; 1:2k in rural area	1:10k
Application domain	management of public and built-up area	- visualising objects in map at scale 1:10k. - Large scale GIS analyses
Providers	municipalities, water boards, provinces, manager of Dutch railway infrastructure (Prorail), Department of Public Water (Rijkswaterstaat), Kadaster	Kadaster
Acquisition method	terrestrial measurements	aerial photographs completed with terrain acquisition
Acquisition rules	no generalisation is applied	little generalisation is applied, e.g.: - only buildings with minimum area of 3x3 meter are acquired - buildings are merged when the distance is closer than 2 meters - roads smaller than 2 meters are represented as lines
Topology	all objects of any class with polygon geometry and relativeHeight '0' divide the terrain into objects without any gaps or overlaps	all objects of classes PartOfWater, PartOfRoad and Terrain and heightlevel '0' form a complete partition without any gaps or overlap.
	- '0' means 'part of the terrain' - possible values are .., -1, 0 ,1 etc - all objects at ground level form planar partition	- '0' indicates that the object is on top of a stack of two or more objects - only values smaller than 0 are allowed (-1, -2 etc) - objects visible from above form a planar partition
	buildings are part of terrain	buildings are located on top of terrain

models. Compared to NEN3610, TOP10NL only further specialises the Relief class in five subtypes; IMGeo only further specialises Layout Element in eleven subclasses and Registration Area in nine subclasses. A major consequence of the limited number of classes in both models is heterogonous classes which are hard to harmonise: it is easier to agree on the definition of a lamp post or a recreational area than on the definition of layout element respectively terrain.

2.3 Differences in demarcation of objects

IMGeo and TOP10NL differ in how they demarcate objects during acquisition. The demarcation of objects is only limitedly defined in the models, but becomes clear when comparing the underlying datasets.
We use the example of road to illustrate two main differences in the demarcation of area objects. Because of a minimum width of 2 meters, many TOP10NL road areas are assigned to neighbouring objects. Examples are parallel roads (cycle paths and footpaths) and parking areas and verges along a road. Sometimes these areas are assigned to the neighbouring terrain and sometimes to the neighbouring roads. The small

Table 2: Main classes in NEN3610, IMGeo and TOP10NL.
Dutch translations added in italics.

Class	NEN3610	IMGeo	TOP10NL
(PartOfRoad (*Wegdeel*))	Yes	Yes	Yes
Terrain (*Terrein*)	Yes	Yes	Yes
(part of)Water (*Waterdeel*)	Yes	Yes	Yes
(PartOf)Railway (*Spoorbaandeel*)	Yes	Yes	Yes
Layout Element (*Inrichtingselement*)	Yes	Yes	Yes
Registration Area (*Registratief Gebied*)	Yes	Yes	Yes
Building (*Pand*)	Yes	Yes	No
Living Unit (*Verblijfsobject*)	Yes	Yes	No
Engineering Structure (*Kunstwerk*)	Yes	Yes	No
Building Complex (*Gebouw*)	Yes	No	Yes
Geographical Area (*Geografisch gebied*)	Yes	No	Yes
Functional Area (*Functioneel gebied*)	Yes	No	Yes
Relief (*Reliëf*)	No	No	Yes

roads themselves are represented with line geometry. In contrast, IMGeo does model these small area objects as specific types of roads. These differences become clear in Figure 1. In this figure TOP10NL road objects only cover the roadways (thus adding parking areas and footpath areas to terrain class), while IMGeo road objects cover the full construction of roads.

The second type of differences in the demarcation of area objects is the division of objects into 'part of' objects. This is well defined in TOP10NL, i.e. Road is divided into PartOfRoads at crossings. IMGeo will most probably follow the division as applied in GBKN. In this dataset roads are divided into parts based on maintenance characteristics such as paving type and administrative boundary. The difference can be seen in Figure 1, where the division of IMGeo roads is different than the division of TOP10NL roads.

Besides differences in demarcation of area objects, also linear objects may contain different geometries in both models as for rails. The line geometry for rails assigned to class Rails in IMGeo represents the middle of the rails (specialisation of Layout Element). In contrast TOP10NL models the geometry of rails as centre lines representing the whole railway body, assigned to class Railway.

A last example of differences in demarcation of objects concerns whether buildings are included in the dataset or not. IMGeo models all buildings independent of their size. TOP10NL only models buildings that meet a minimal area (3x3 m).
The way objects are demarcated is mostly only available in acquisition rules and not in the models. Consequently for harmonising differences in demarcations, it is most important to formalise this information in the domain models.

2.4 Differences in attribute values

The attribute values in both models differ slightly in many cases. The differences between the paving types and railway types, as shown in Table 3, are representative for such differences. The differences are small and may not be significant, such as open pavement (IMGeo) and partially paved (TOP10NL).

Figure 1: Examples of IMGeo data (courtesy of municipality of Almere) and TOP10NL data (courtesy of Kadaster).

IMGeo roads	TOP10NL roads shown transparently on top of IMGeo roads. This shows that TOP10NL roads cover smaller areas than IMGeo roads
IMGeo test data	*TOP10NL data*
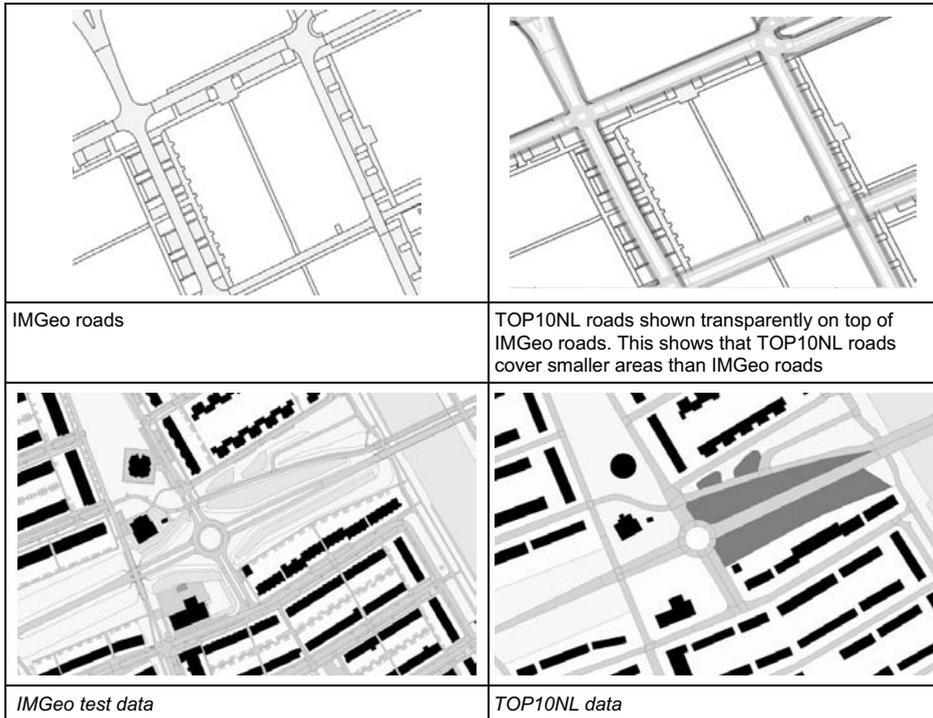
Another important difference to be solved for integration are the different values for terrain type (also shown in Table 3). None of the twelve IMGeo types has exactly the same label as one of the nineteen TOP10NL types. A few types are presumably the same (grass and grass-land; 'nature and landscape' and heather). In addition IMGeo contains coarser classifications for two types of terrain: forest and green object (see Table 3). A last difference in terrain types worth mentioning is 'built-up' area in TOP10NL to identify terrain on which buildings are located. Since buildings cause a gap in the terrain (see Table 1), IMGeo does not have 'built-up' terrain.

A last noticeable example of slightly different attribute values (not shown in Table 3) are Layout Element types. Of the eighty identified types in both models, only nine have exactly the same name, examples are tree, hedge, wall, and sign post. Ten types that differ in name presumably model the same concept, examples are road closing (TOP10NL) and barrier (IMGeo); hectometer stone (IMGeo) and milestone (TOP10NL). All other 60 types cannot be mapped. The IMGeo types are mainly originating from the utility sector or required for the management of public area. The TOP10NL types are needed for orientation or are related to Defense (the original application domain of TOP10vector).

Table 3: Slightly different attributes and attribute values pointing at same concepts.

Concept	IMGeo			TOP10NL		
	Class	Attribute	Value	Class	Attribute	Value
Paving	Part Of Road Terrain	typeOfPaving	closed pavement open pavement unpaved	Part Of Road	pavingType	paved unpaved partially paved unknown
Railway types	Rail	typeofRail	crane, metro, tram, train, fast tram/lightrail			
	Railway	typeofRailway	crane, metro, tram, train metro, fast tram/lightrail, railway verge, to be defined	Railway	typeofRailway	mixed metro tram train
Terrain type	Terrain	typePartOfTerrain	**forest** grass nature and landscape culture land ***other green object*** industrial terrain uncultivated terrain courtyard area with plants recreational area sport terrain embankment	Terrain	landUseType	**mixed forest** **brushwood** **deciduous wood** **coniferous wood** grassy area heather arable land ***orchard*** ***tree cultivation*** ***poplar*** graveyard ***graveyard in forest*** jetty sloped stones built-up area fruit cultivation loading bay area for railway sand

2.5 Different classes for same concepts

Several concepts are modelled with different classes, as shown in Table 4. An example is the concept 'verge' which is terrain of type 'grass' in TOP10NL and a specific type of road in IMGeo. Another example are sport area, recreational area and industrial area. These are considered Terrain in IMGeo and Functional Area, i.e. a collection of objects of different classes, in TOP10NL. Lastly, Engineering Structure is a specific class to model infrastructural structures such as viaducts, bridges, locks and dams in IMGeo. In TOP10NL these objects are modelled as a specific type of infrastructural objects (PartOfWater, PartOfRailway or PartOfRoad) or as a Layout Element.

Table 4: Same concept, differently modelled.

Concept	IMGeo	TOP10NL
	Modelled with class	
Verge	PartOfRoad	Terrain
Industrial Area	Terrain	Functional Area
Recreational Area	Terrain	Functional Area
Sport Area	Terrain	Functional Area
Engineering structure	Engineering Structure	Layout element PartOfRoad/PartOfWater/PartOfRailway

2.6 Same attribute name for different concepts

Table 5 shows how the attribute typeOfRoad assigned to PartOfRoad is used in a different manner in the two models. IMGeo uses the attribute to distinguish different parts of a road; TOP10NL to define a hierarchy of roads required for visualisation. As shown in Table 5, NEN3610 even uses the attribute in a third way.

Table 5: Different use of attribute typeOfRoad.

	IMGeo	TOP10NL	NEN3610
Used for	identifying different parts of a road	defining hierarchy for roads for visualisation	defining hierarchy for function of roads
Example values (not complete)	parking area public transport footpath verge roadway cycle path pedestrian area residential area	highway main road regional road local road street	continuous road access road access road to residential areas other roads facilities

2.7 Differences in amount of information

In general, TOP10NL models more information than IMGeo. For example TOP10NL contains more attributes for its classes than IMGeo. Most likely this is because GBKN, the underlying dataset of IMGeo, contains less attributes than TOP10NL data. Another example of less information in IMGeo compared to TOP10NL is that IMGeo identifies 'forest' versus four types of forest in TOP10NL and one 'green object' versus four types in TOP10NL (see Table 3). A last example of less information in IMGeo refers to sport area and recreational area. IMGeo classifies sport and recreational areas as single objects where TOP10NL identifies different types of objects (roads, buildings, terrain) that constitute the areas (see also section 2.5).

2.8 Differences in class definitions

Differences in application domains have led to different classifications for the same objects, for example when is an object 'forest', 'grass', 'recreational area' and/or 'area with plants'? We can observe these differences when comparing the datasets, i.e the instances of the classes, but they are not appearant from the models. In TOP10NL a wooded area may be split in two areas: deciduous wood and coniferous wood to identify the type of land-cover. The same area can be split in a different way in IMGeo depending on whether the wood is maintained (area with plants) or not (forest). Figure 2 (Almere) shows another example of such differences. A forest object is identified within

grassy area in TOP10NL data (right). IMGeo data does not identify this forest area (left). From the information models the reason for this difference does not become clear.

Figure 2: Differences in instances of same class (i.e. Terrain).

Terrain-grass in IMGeo	Terrain-forest and Terrain-grass in TOP10NL
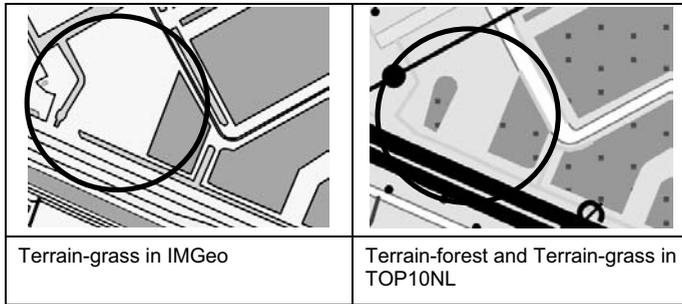
3. INTEGRATED INFORMATION MODEL TOPOGRAPHY

Although TOP10NL models topography at scale 1:10k and IMGeo at scale 1:1k in urban area and 1:2k in rural area, we can conclude from the previous section that TOP10NL cannot be considered as a derivation of IMGeo. This is not surprisingly since none of the models used the other model as starting point. Also the differences in history, objectives, providers, source data and stakeholders explain the differences of two domain models topography: one supports management of public and built-up area and one visualises topography at scale 1:10k.

Because of these differences, topographic data of the same objects is currently collected twice to serve two application domains. To collect topographic data that meets the requirements of both domains in the future, the domain models need to be integrated. Such integrated model will assure consistency when users (or applications) move from one dataset to another. Based on the conclusions of the comparison study, this section presents recommendations to accomplish the integration.

Starting from the current differences, two main steps are required to build the integrated information model topography. Firstly harmonisation, i.e. agreeing on definitions of concepts, should be accomplished as much as possible. Section 3.1 describes recommendations for harmonising the differences identified in section 2. The result will be two better aligned domain models topography. Section 3.2 presents and motivates the modelling principles for the integrated model topography. This model defines how topographic data on real-world objects can be collected once and used in both the IMGeo and TOP10NL domain, starting from the harmonised versions of IMGeo and TOP10NL.

3.1 Recommendations for harmonising

The first main step for harmonising is to study whether the differences between the models are principal or random: which differences in modelling can be harmonised based on agreement of concepts without having significant consequences for one of the models? To support this harmonisation, this section formulates recommendations for harmonising the differences identified in section 2.

Some differences are not clear from the models, but were identified by comparing the datasets. Therefore our first recommendation, before harmonising the models, is to model any information about the content and meaning of datasets that is currently not included in the domain models, e.g. acquisition rules. This allows harmonising and integrating this information as well. For example, the current TOP10NL acquisition rules state that railway banks are not measured, despite the presence of the class Railway in the model.
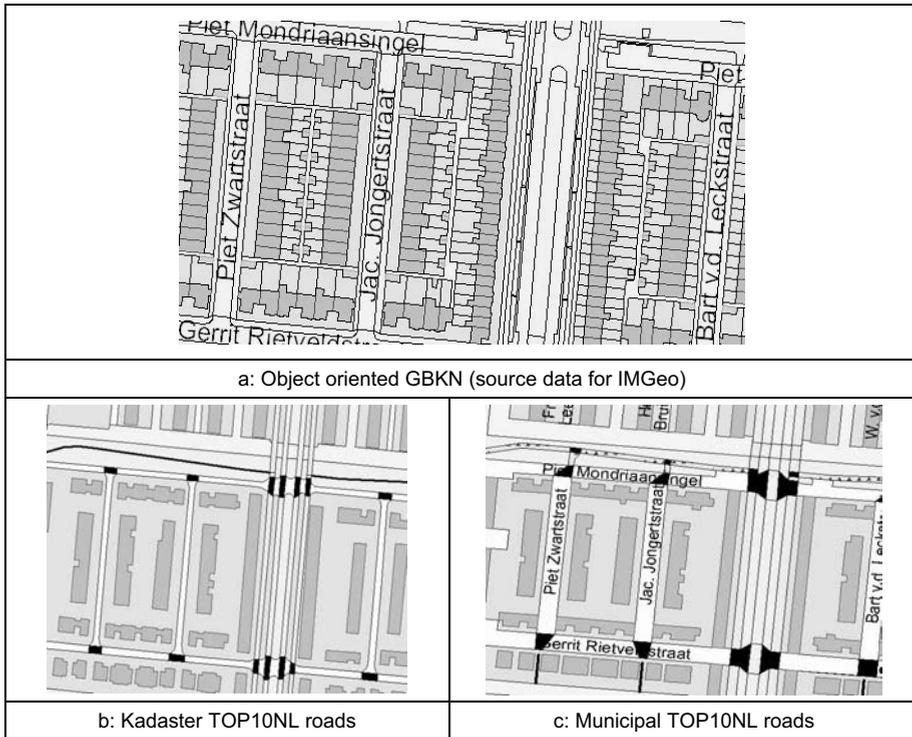
Many differences between the information models are caused by *difference in perspective* (section 2.1). These different perspectives cannot easily be harmonised because they are justified by differences in objectives, source data, application domain, providers, and acquisition methods. However, the law on key registers provides potentials for harmonising some parts of the perspective. Specifically two aspects of the law enforce municipalities to inform the Kadaster about updates for TOP10NL data. The first aspect is that any user of the key register must inform the provider when (s)he notices an error. Secondly, municipalities are obliged to use the Kadaster's TOP10NL data, updated every 2 years, instead of their self-produced 1:10k datasets, updated more frequently. To effectuate updates in TOP10NL data as soon as possible, some large municipalities will send TOP10NL updates based on their 1:10k datasets. For this purpose they are currently converting their 1:10k datasets into TOP10NL compliant datasets. This practice makes the integration issue of the two topographic datasets, i.e. IMGeo and TOP10NL, relevant within municipalities.

Because not all information on the data is laid down in the TOP10NL model, TOP10NL data can be generated from a municipal perspective without violating the model. An example is that the Kadaster often assigns road areas that are too small to be area objects (smaller than 2 meters) to terrain. However, assigning these areas to neighbouring roads better supports the municipal maintenance task of public area and fits better with the definition of PartOfRoad in the TOP10NL model. Consequently municipal TOP10NL roads may cover the full construction of all IMGeo road objects which solves the *differences in object demarcation* of roads (section 2.3). To illustrate this, two TOP10NL road implementations, one generated by Kadaster and one generated by municipality of Rotterdam, are compared with IMGeo roads in Figure 3.

Obviously this poses new research questions, since now the differences in perspective do not occur between IMGeo and TOP10NL, but within one dataset, i.e. TOP10NL. In conclusion, to solve the *differences in object demarcation* it is most important to make these demarcations unambiguously explicit in the models. In a next step it can be studied whether differences can be aligned and how.

A first step in harmonising *differences in main classes* (section 2.2) is to model more specialisations (i.e. subclasses). The result will be more homogenous classes on which it is easier to agree. Figure 4 shows an example. The left part of the figure shows the current Terrain class in TOP10NL with its different attribute values for different types of terrain. Integrating IMGeo and TOP10NL requires agreeing on the concept of Terrain. The alternative modelling with subclasses for different types of terrain (Figure 4, right) requires only agreeing on the definition of types of terrain, for example, Farmland or Forest.

Figure 3: IMGeo roads (a) and TOP10NL implementations of roads by Kadaster (b) and municipality of Rotterdam (c).

a: Object oriented GBKN (source data for IMGeo)

b: Kadaster TOP10NL roads	c: Municipal TOP10NL roads

Figure 4: Left: Terrain class in current TOP10NL model. Right: Subclasses for different types of Terrain result in homogenised classes.
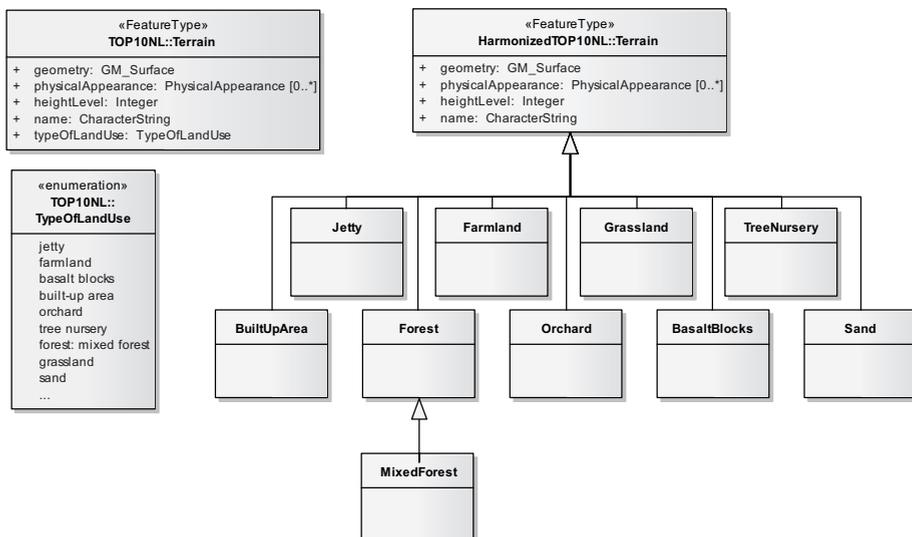
Such homogenous classes will also avoid that *different classes are used for the same concepts* (section 2.5)*.*

Differences in attribute values (section 2.4) can be harmonised through lists of common types completed with harmonised values in case of non-significant differences, for example for the paving types (Figure 5) and railway types. Some values may remain information model specific, example is the built-up area for TOP10NL terrain types which is lacking in IMGeo because buildings cause gaps in the terrain.

Figure 5: Harmonised values for paving types.
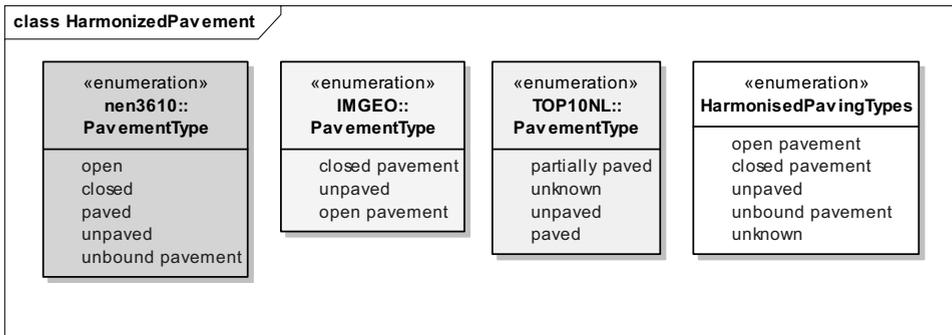
The *same attribute name for different concepts* (section 2.6) can only be harmonised by agreeing on common use of attributes. To avoid such differences in the future, attribute names should be used that have less ambiguous semantics.

To solve *differences in amount of information* (section 2.7), information required at the smallest scale (TOP10NL), but not available in the largest scale (IMGeo) can be either moved down to the largest scale or be removed from the smallest scale. Moving down information to IMGeo is only of interest for municipalities when it is relevant for their application domain.

To solve *differences in class definitions* (section 2.8) new objects at cross sections of classifications could be generated. However, a class for every possible combination makes the models more complex. An example are the four possible combinations for *area with plants/forest* (IMGeo) and *deciduous/coniferous wood* (TOP10NL): DeciduousWoodAreawithPlants etc. A better option is therefore to keep the classes from the original models. This will result in overlapping polygons in the datasets, but since two different concepts are registered for the same area (maintenance and land-cover) this reflects the real-world situation. In any case the exact definitions of classes should be unambiguously defined in the information models. In the current situation such differences only become clear when comparing the data (i.e. instances of classes).

Harmonising the information models using these recommendations will result in better aligned IMGeo and TOP10NL models. The more harmonisation can be achieved, the more straightforward the integration of the two domain models will be.

3.2 Recommendations for integrating: Base Model Topography

We propose an information model topography that integrates the two information models. The modelling principles that we present here are motivated by two important conclusions of the comparison study (section 2).

Firstly, two datasets defined in two information models topography are necessary to meet the specific demands of the two domains, i.e. IMGeo for maintenance of public area and TOP10NL for visualisation at scale 1:10k. Secondly TOP10NL cannot be derived from IMGeo, because the application domain proofs to be more dominant on the perception of topography than scale is.

Starting from these two conclusions, we propose an intermediate domain model between NEN3610 at the one side and IMGeo and TOP10NL at the other side: Base Model Topography (BMT). The motivation for this intermediate layer instead of solving the integration within NEN3610 is that NEN3610 is meant to integrate at a higher level of abstraction. The two conclusions that direct the modelling principles of BMT are invalid for all domain models under NEN3610. Consequently it is better to solve the integration of these two topographic domains outside NEN3610.

BMT is an information model defining scale-independent topographic classes where both IMGeo and TOP10NL can be derived from. The BMT classes respect both the IMGeo and TOP10NL perspectives on topography. However, they do not necessarily have the same label (see further). For the moment BMT defines how concepts in IMGeo are related to concepts in TOP10NL. This provides consistency for users (and applications) when moving from one dataset to the other. However, the data is still separately collected until an organisation has interest to collect data to serve both domains. In that case the 'collect once, use many times' principle will be realised through collecting data on BMT classes. Therefore they contain all information that becomes relevant in any dataset that needs to be derived from BMT.

The modelling principles of our approach are based on the multi-scale Information Model TOPography (IMTOP, see Stoter et al., 2008). IMTOP, which integrates topographic data at scales 1:10k to 1:1000k for the Netherlands' Kadaster, proposes an abstract super class for every topographic class. These super classes have subclasses at all scales and only contain attributes and attribute values valid for all scales. The super classes are abstract and the data is collected for the largest scale dataset, while smaller scale datasets are derived from the next larger scale dataset.

Similar to IMTOP, we define IMGeo and TOP10NL classes that are derived classes from BMT classes. An example is shown in Figure 6 where we model the derivation of the PartOfRoad object. Constraints defined in Object Constraint Language (OCL) can define how objects in IMGeo and TOP10NL can be derived from BMT.

Although we follow the main principles of IMTOP, the proposed BMT differs on a few fundamental aspects.

Firstly, the name of the BMT classes and the derived classes can be different because of different perspectives on concepts (see Figure 7). In contrast, every IMTOP super class occurs as subclass with the same name in each scale. For example the properties of an IMTOP road super class are inherited by road class at 1:10k scale, by road class at 1:50k scale, by road class at scale 1:100k. Secondly, we define association relationships between BMT classes and the derived classes, instead of a generalisation/specialisation relationship as in IMTOP. The reason is that BMT classes and the derived classes do not necessarily represent the same concepts. Thirdly, the BMT classes (comparable to super classes in IMTOP) are non-abstract. The reason for this is that objects in both domains (IMGeo and TOP10NL) are derived from instances of

BMT classes, which contain all information required to derive both IMGeo and TOP10NL data. Fourthly, we recommend moving all information down to BMT to avoid extra data acquisition for derived datasets. This implies that all attributes of the IMGeo and TOP10NL objects are derived, except the Identifier and other system attributes not shown in Figure 6. In IMTOP the classes at smaller scales can have extra attributes which are only valid (and collected) for the specific scale. Our BMT approach also slightly differs with the i(integration)-classes as identified in the multi-representation approach of Friis-Christensen and Jensen (2003). The i-classes only contain attributes that are valid in the corresponding classes, as the super classes for IMTOP.

Figure 6: Concept of PartofRoad, modelled in integrated model topography.
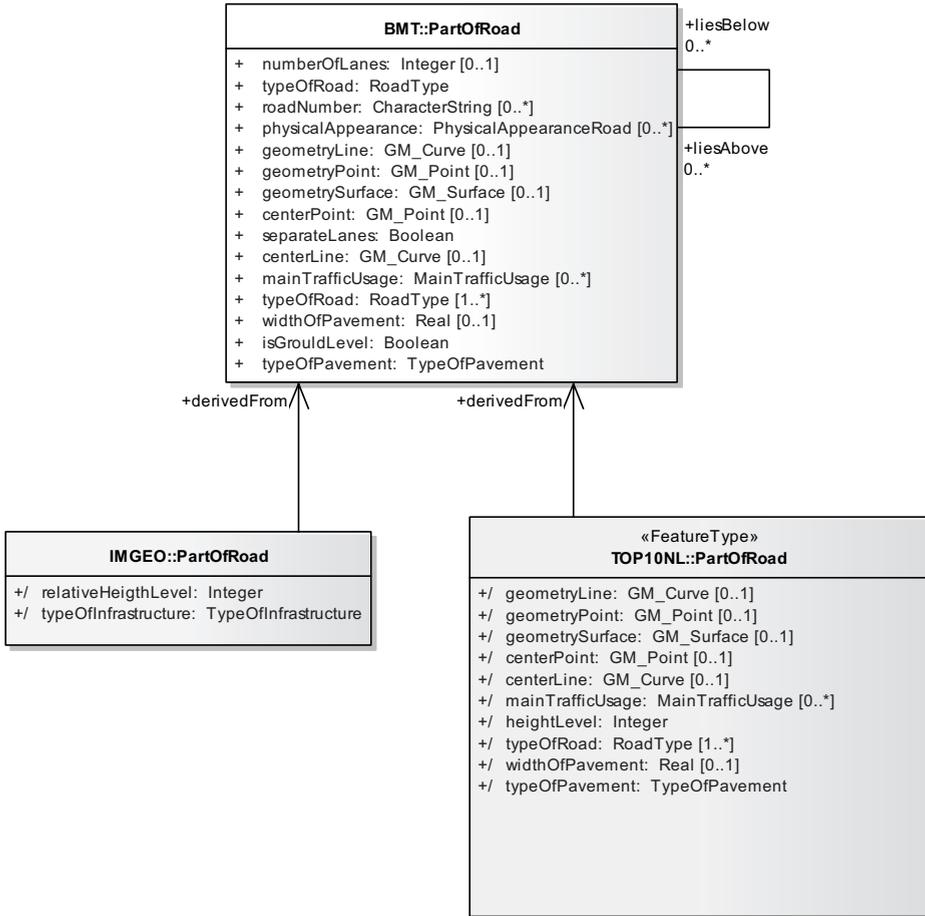
Finally, we recommend relationships (liesAbove and liesBelow) and the Boolean attribute IsGroundLevel to every BMT class (as shown in Figure 6) to derive both the relativeHeight and heightlevel attributes required for IMGeo respectively TOP10NL. Consequently both the IMGeo and the TOP10NL implementation of topology can be derived from BMT.

Figure 7: Modelling sport area as BMT class and derived classes in IMGeo and TOP10NL.
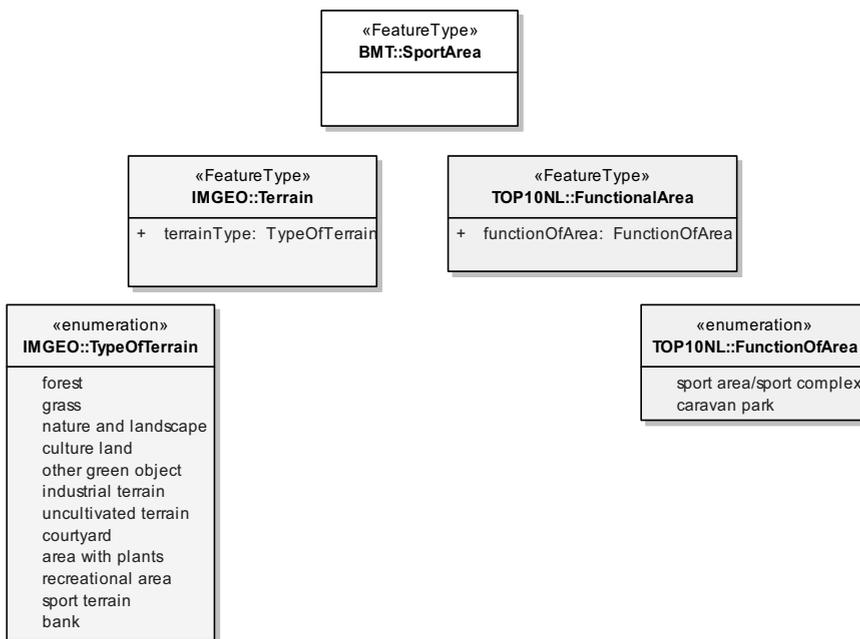
4. CONCLUSIONS

In this article we studied requirements and possibilities to harmonise and integrate two independently established information models topography. The harmonisation and integration consists of several steps.

At first we identified types of differences that have to be addressed. Apart from these differences, four other conclusions from this comparison study are important for harmonising and integrating the two models. Firstly, for many differences it is not clear whether they are random (i.e. easy to harmonise) or fundamental (i.e. to be addressed in the integration). This requires further study. Secondly, because not all information is defined in the models, datasets compliant to the models may be implemented with different 'flavours'. These ambiguities are unwanted when reusing data of other domains in SDIs. Therefore an important recommendation is to make information on the content and meaning of data as much as possible explicit in the information models. Thirdly, two information models topography are necessary to meet the specific demands of the two domains, i.e. maintenance of public area and visualisation at scale 1:10k. Finally TOP10NL cannot be derived from IMGeo, because the application domain of these two large scale data sets determines the different perspectives on topography rather than scale does. At large scales (also valid for scale 1:10k) objects can be represented with their true geometries, and therefore harmonisation and integration is mainly a schema matching problem. At the smaller scales, symbolisation causes objects to be altered with respect to reality. Consequently at smaller scales harmonisation and integration becomes merely a multi-scale problem, i.e. how can a dataset be converted into a dataset with fewer details.

Based on the conclusions, the article formulated recommendations to harmonise the differences and presented modelling principles to define an integrated model topogra-

phy, both illustrated with UML examples. These recommendations, principles and illustrations show the problems and potentials of harmonising and integrating different data models into global data models to enable data provision within national and international SDIs. First the proposed integrated model formally defines how concepts in one dataset relate to concepts in another dataset. In a future step the results of this study can be further developed to move towards 'collect data once, maintain it at several domain databases, and use it multiple times'. Comparing similar developments in other countries, for example aligning Teknisk Korte and TOP10 for the Danish SDI, can be very useful here.

REFERENCES

Afflerbach, S., Illert, A. and T. Sarjakoski (2004). "The Harmonisation Challenge of Core National Topographic Databases in the EU-Project GiMoDig". *Proceedings of the XXth ISPRS Congress*, July 12-23, 2004, Istanbul, Turkey, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXV (B4:IV): 129-134.

BAG (2006). Catalogus Basis Gebouwen Registratie (2006), versie 4.0, Ministerie van VROM, Projectbureau Basisregistraties voor Adressen en Gebouwen (BAG), at: http://bag.vrom.nl/ufc/file/bag_sites/a9c332aa4bb9765d34c1be8581949dd1/pu/BGR_cataloguc_2006_def.pdf. In Dutch.

EuroGeographics (2009). http://www.eurogeographics.org/eng/00_home.asp.

EuroSpec (2009). http://www.eurogeographics.org/eng/01_EuroSpec_goals.asp.

Geonovum (2008). Standaarden voor Geo-informatie, the National Spatial Data Infrastructure (NSDI) executive committee in the Netherlands, at: http://www.geonovum.nl/standaarden.html. In Dutch.

Friis-Christensen A. and Jensen C.S. (2003). "Object-relational management of multiply represented geographic entities". *Proceedings of the Fifteenth International Conference on Scientific and Statistical Database Management*. Cambridge, MA, USA, July 9–11, pp 183–192.

GiMoDig (2009). http://gimodig.fgi.fi/.

Hofman, A.M., P. Dilo, P. van Oosterom and N. Borkens (2008). "Developing a varioscale IMGeo using the constrained tGAP structure. Using the constrained tGAP for generalisation of IMGeo to Top10NL model", *Workshop on Generalisation and Multi-representation of the International Cartographic Association*, Lyon, June, 2008. http://aci.ign.fr/montpellier2008/papers/27_Hofman_et_al.pdf.

IMGeo (2007). Informatiemodel Geografie (IMGeo); beschrijving van het model, IMGeo versie 1.0, Werkgroep IMGeo, at: http://www.geonovum.nl/informatiemodellen/imgeo/, 2007. In Dutch.

INSPIRE (2007). http://inspire.jrc.ec.europa.eu/.

ISO (2005). ISO 19109:2005 Geographic information - Rules for application schema.

LSV GBKN (2007). GBKN Handboek, versie 2.1, Stichting LSV GBKN, Document number 07.05/052, at: http://www.gbkn.nl/nieuwesite/downloads/07.05.065%20GBKN%20handboek%20VIPU2.1.pdf. In Dutch.

NEN3610 (2005). Basic scheme for geo-information - Terms, definitions, relations and general rules for interchange of information of spatial objects related to the earth's surface. Normcommissie 351 240 "Geo-informatie". In Dutch.

Stoter, J.E., Morales, J.M., Lemmens, R.L.G., Meijers, B.M., van Oosterom, P.J.M., Quak, C.W., Uitermark, H.T. and van den Brink, L. (2008). "A data model for multi - scale topographic data", *Lecture Notes in Geoinformation and Cartography*.

Stoter, J.E. (2009). "Towards one domain model and one key register topography", in: P. van Oosterom (ed.). *Core Spatial Data*, Delft: Netherlands Geodetic Commission.

TOP10NL (2005). TOP10NL versie 2.3, februari 2005, http://www.kadaster.nl/top10nl/ gegevensmodel_top10nl_2%203.pdf. In Dutch.

VROM, (2009a). Ministry of Housing, Spatial Planning and the Environment http://www. vrom.nl/pagina.html?id=12050. In Dutch.

VROM (2009b). Ministry of Housing, Spatial Planning and the Environment http://www. vrom.nl/pagina.html?id=36586. In Dutch.

VROM (2009c). Ministry of Housing, Spatial Planning and the Environment http://www. vrom.nl/pagina.html?id=36688. In Dutch.

An Analysis of Technology Choices for Data Grids in a Spatial Data Infrastructure

Serena Coetzee and Judith Bishop

Department of Computer Science, University of Pretoria, Pretoria, South Africa
{scoetzee, jbishop}@cs.up.ac.za

Abstract

The concept of grid computing has permeated all areas of distributed computing, changing the way in which distributed systems are designed, developed and implemented. Data grids enable the sharing of data in a virtual organisation and are typically implemented for data federation in data-intensive environments. So far, they have been applied to traditional data (text, image, sound). We present a scenario that describes for the first time how data grids can be applied to enable the sharing of address data in a spatial data infrastructure (SDI). Consolidating spatial data from distributed heterogeneous sources into a single centralised dataset requires, amongst others, a considerable human coordination effort. A data grid consolidates data directly from the distributed sources, thereby eliminating the effort. We present a reference model called Compartimos (Spanish for 'we share'), that is based on the Open Grid Services Architecture (OGSA) but is customised for sharing address data in an SDI, and we analyse existing technologies, such as the Globus Toolkit, ISO 19100 standards and Open Geospatial Consortium (OGC) web service implementation specifications, for Compartimos. This article advances the mutual understanding between data grids and SDIs and sheds light on a future technological solution that could overcome some of the data sharing impediments that are experienced in SDIs today. Finally, results from the analysis and future directions for research are discussed.

Keywords: spatial data infrastructure (SDI), data grid, data sharing, service-orientation, web service, grid computing, address data, geographic information, GIS.

1. INTRODUCTION

A grid is a system that is concerned with the integration, virtualisation, and management of services and resources in a distributed, heterogeneous environment that supports virtual organisations (collections of users and resources) across traditional administrative and organisational domains (real organisations). A data grid is a special kind of grid in which data resources are shared and coordinated (OGF, 2007b). How virtual organisations collaborate and share resources in order to achieve a common goal is described as the 'grid architecture' in *The Anatomy of the Grid* (Foster et al., 2001) and *The Physiology of the Grid* (Foster et al., 2002). This has subsequently evolved into the Open Grid Services Architecture (OGSA) published by the Open Grid Forum (OGF, 2006), a vision of a broadly applicable and adopted framework for grids. The OGSA data architecture (OGF, 2007a) describes the interfaces, behaviours and bindings for manipulating data within the broader OGSA.

The work reported in this article is part of a research project on 'Distributed Address Management', which has the objective of establishing whether the data grid approach is an option for national address databases in an SDI and if so, what the design imperatives for such an approach are. In earlier work, we used a novel evaluation framework for national address databases to evaluate existing information federation models, as well as the data grid approach, for the use in address databases for national

SDI. This evaluation, as well as an analysis of address data in an SDI, confirmed that there are quite a few similarities between the data grid approach and the requirement for consolidated address data in an SDI. The evaluation further showed that where a large number of organisations are involved, such as for a national address database, and where there is a lack of a single organisation tasked with the management of a national address database, the data grid is an attractive alternative to other models (Coetzee and Bishop, 2008). In this article we present Compartimos, a reference model for an address data grid, which is based on the OGSA data architecture. Compartimos was developed to gain a better understanding of the components involved in a data grid approach.

Due to their service, infrastructure and land administration responsibilities, it is commonly found that it is the local authority that establishes and maintains address data for its area of jurisdiction (Levoleger and Corbin, 2005; Williamson et al., 2005; Coetzee et al., 2008). When address data is required for an area that extends across these jurisdictional boundaries, the data has to be collated from the various local sources. Currently, many national address databases of the world follow the centralised approach where address data is loaded into a single centralised database (Paull, 2003; Fahey and Finch, 2006; Nicholson, 2007). The novel data grid approach proposed in this article deviates from this centralised approach.

Reports on grid computing for spatial data in general are found in Hua et al. (2005), Aloisio et al. (2005b), Goodenough et al. (2007), Koutroumpas and Higgins (2008), Xue et al. (2008) and Aydin et al. (2008). First research reports on Grid computing technologies in SDI environments are found in the papers by Zhao et al. (2004), Aloisio et al. (2005a), Shu et al. (2006), Wei et al. (2006) and Di et al. (2008), as well as the recently launched *Geodateninfrastruktur-Grid* (GDI-Grid) project (http://www.d-grid.de/index.php? id=398&L=1), which is part of D-Grid, a long-term German strategic initiative in Grid computing. It is expected that the recently initiated collaboration between OGC and the OGF (OGC OGF, 2007) will start adding to the momentum of such publications. The initial focus of the collaboration is to integrate OGC's OpenGIS Web Processing Service (WPS) Standard with a range of "back-end" processing environments to enable large-scale processing, or to use the WPS as a front-end interface to multiple grid infrastructures, such as TeraGrid, NAREGI, EGEE and the UK's National Grid Service. Results from our work suggest that grid-enabling spatial data integration in an SDI environment should also be explored, i.e. grid-enabling other web services specified by OGC, such as the Web Feature Service (WFS). The OGC-OGF collaboration proves that the international geospatial community is increasingly interested in utilising grid technology as a solution to its problems, while the grid community has found other users that can benefit from its technology.

In the position paper by Craglia et al. (2008), a group of international geographic and environmental scientists from government, industry and academia present the vision of the next generation Digital Earth and identify priority research areas to support this vision, which include information integration and computational infrastructures. Both these priority areas are addressed in our research.

Thus the research community, as well as industry, recognises the importance of grid computing for SDIs and geospatial data in general. The related work confirms that our approach of the data grid as enabler for sharing address data in an SDI is innovative and new, and it proves that the work is extremely relevant at this point in time, both in Computer Science (grid computing) and in Geographic Information Science (SDI). Our work is unique because Compartimos is designed for address data.

The objectives of this article are 1) to present a scenario that describes how data grids can be applied to enable the sharing of address data in an SDI; 2) to present Compartimos, a reference model for an address data grid; 3) to analyse technology choices for Compartimos objects; and 4) to present results and discuss future research in this new area.

2. SCENARIO: INTERNATIONAL PROPERTY VALUATION

An airline rewards company, AirMiles, wants to introduce an AirMiles credit card to its estimated ten million international customers. They have contacted FinBank as the provider of the credit card. FinBank is interested, but wants to evaluate the customer base before finalising the terms and conditions and signing an agreement. This evaluation includes a valuation of the property at each AirMiles customer's residential address. The property valuation comprises geocoding the customer's address and comparing it to other datasets such as credit rating per suburb, crime statistics in the area, probability of sinkholes in the area, and proximity to public transport. Neither AirMiles nor FinBank are experts in these areas and have contracted ConsultCo to do the property valuation.

The AirMiles customer base spans more than one country and therefore the geocoding has to be done against address data collected from different countries, including local authorities within these countries. In some countries this data is available for free, in others the data has to be purchased. Since customers are randomly spread across the country, it is not known which parts of the country are needed for the geocoding, and therefore the dataset for the whole country has to be purchased, where applicable, at a steep fee. The AirMiles customer database in itself is a valuable asset that has to be protected and it includes personal information that requires protection for privacy reasons. AirMiles would prefer employees from ConsultCo doing the valuation on-site at the AirMiles offices where stringent security measures are in place. This implies that ConsultCo has to fly in experts from different offices, adding to the travelling costs. Finally, the licensing of the sophisticated geocoding software package that ConsultCo uses, does not allow ConsultCo to install the geocoding software on AirMiles machines. The property valuation is quite simple, but the geocoding depends on an address dataset spanning more than one country without which the rest of the valuation cannot continue.

If this scenario is projected into a future world where an address data grid is a reality, the valuation process could be simplified as follows.

AirMiles configure their customer database as a Grid resource for which they set a strict policy that allows ConsultCo access to a customer's address for purposes of geocoding only, and one or two attributes of a customer into which they can write geocoding and valuation information. ConsultCo queries an online directory of address data providers who have set-up their address datasets as Grid resources and provide access to their data through standardised Grid services. A specific address data provider could supply data for an area ranging from a local authority's jurisdiction to a province or state, a country or even an international region. The online directory includes pricing and quality of service information so that ConsultCo can pick the best offer available. The Grid services are standardised to eliminate differences resulting from data stored in underlying DBMSs from different vendors. The Grid services are further standardised to exchange address data in a standard format that the ConsultCo geocoding software understands.

ConsultCo now executes their geocoding software from machines at their offices, which reads the customer address from the server at AirMiles offices, matches it to the address data providers from the relevant country, and writes the resulting coordinate into the geocoding attribute in the AirMiles customer database. Thus, from the point of view of the geocoding software a single integrated address dataset is used for geocoding. When it is time for the property valuation, ConsultCo access the customer geocoding attribute (coordinates) and compare it to the other datasets (e.g. public transport, which in turn could each come from a different grid resource). The resulting valuation information is written into the valuation attribute on the AirMiles customer database. Refer to Figure 1 for this scenario.

Figure 1: Geocoding a customer database.
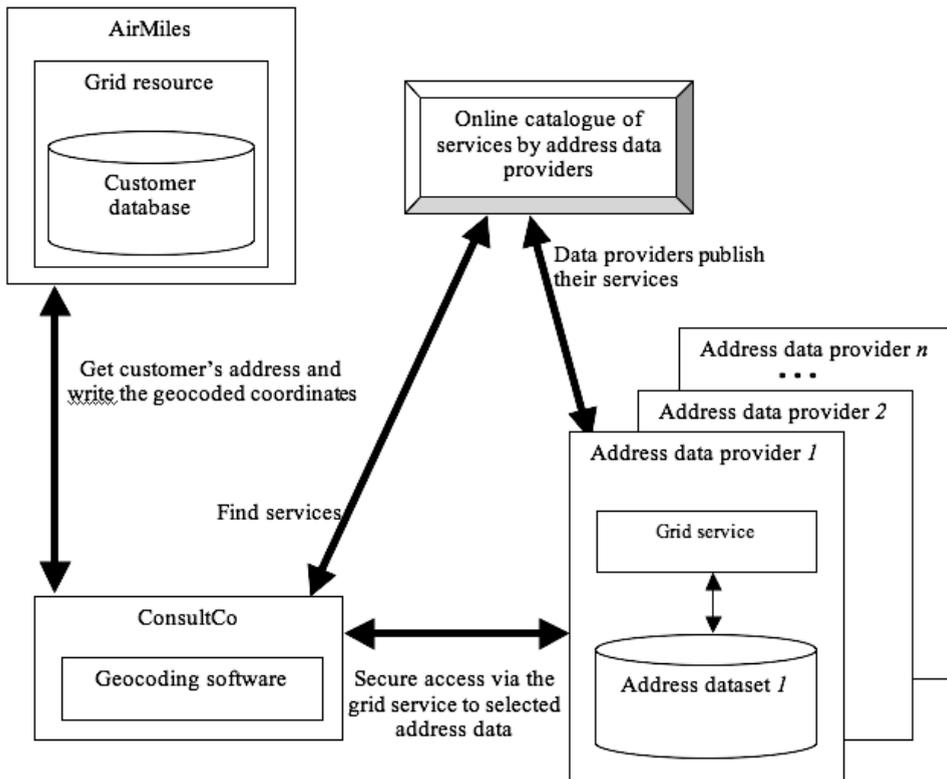
There is no need for ConsultCo employees to be on-site at the AirMiles offices and the logistics are simplified as the employees can continue with the work from the desktops in their respective offices. ConsultCo does not have to purchase the address data for the whole country, nor does it have to consolidate the data from multiple sources, rather it only uses and/or pays for the specific data that is required to geocode the addresses. Thus, the data grid, through data integration, has simplified the logistics and therefore the costs of the project, and more importantly, the costs and network traffic for the address data have been significantly reduced since ConsultCo accesses relevant address data only.

3. THE COMPARTIMOS REFERENCE MODEL

In this section we present the *Compartimos* (Spanish for 'we share') reference model, which gives an abstract representation of the essential components and their relationships that are required to share address data on a data grid in an SDI environment. Compartimos was developed in order to analyse the requirements for realising a scenario such as the one described above. For our research project, the Compartimos reference model is described in terms of the five viewpoints of the ISO Reference Model for Open Distributed Processing (RM-ODP) (ISO/IEC 10746:1998), i.e. the enterprise, information, computational, engineering and technology viewpoints. In this article we focus on the computational viewpoint, which is concerned with the functional decomposition of the system into a set of objects that interact at interfaces - enabling system distribution. Compartimos is based on the OGSA data architecture, which implies that it follows a service-oriented approach and provides services similar to the OGSA data architecture. Compartimos is a domain-specific application of the architecture, emphasising address data in an SDI.

Table 1: Compartimos objects.

Object name	Type	Main purpose
Catalogue	Data	Stores information about services and data
CatalogueService	Service	Provides read and write access to the catalogue
VirtualAddressDataService	Service	Consolidates data
AddressDataAccessService	Service	Provides uniform access to individual address datasets
AddressDataset	Data	The individual address data set
AddressService	Service	A third party address-related service such as routing or mapping
ReplicaService	Service	Replicates data in the address data grid
TransferService	Service	Transfers large volumes of address data

Table 1 lists the Compartimos objects and their purpose, while Figure 2 shows how the Compartimos objects interact with each other in the address data grid. The word 'object' is used here in compliance with RM-ODP where it is used in the broader sense of the word and not with its very specific interpretation in the object-oriented paradigm. Some aspects of the OGSA data architecture, such as policies, storage management and caching, are excluded from Compartimos because they can be used generically for any kind of data and do not have to be tailored specifically for address data in an SDI. In section 4 the purpose and capabilities of each Compartimos object are described and related to the OGSA data architecture, along with the discussion of technology choices. Compartimos' ReplicaService and TransferService are sufficiently generic to be excluded here.

4. TECHNOLOGY CHOICES FOR COMPARTIMOS OBJECTS

Table 2 provides an overview of available technology choices for Compartimos objects. For each Compartimos object, there is also the option of developing it from scratch without using existing technology, which has the same pros and cons of software reuse that have been well documented over the years. Thus this choice is not discussed here.

Figure 2: Object interaction in Compartimos.
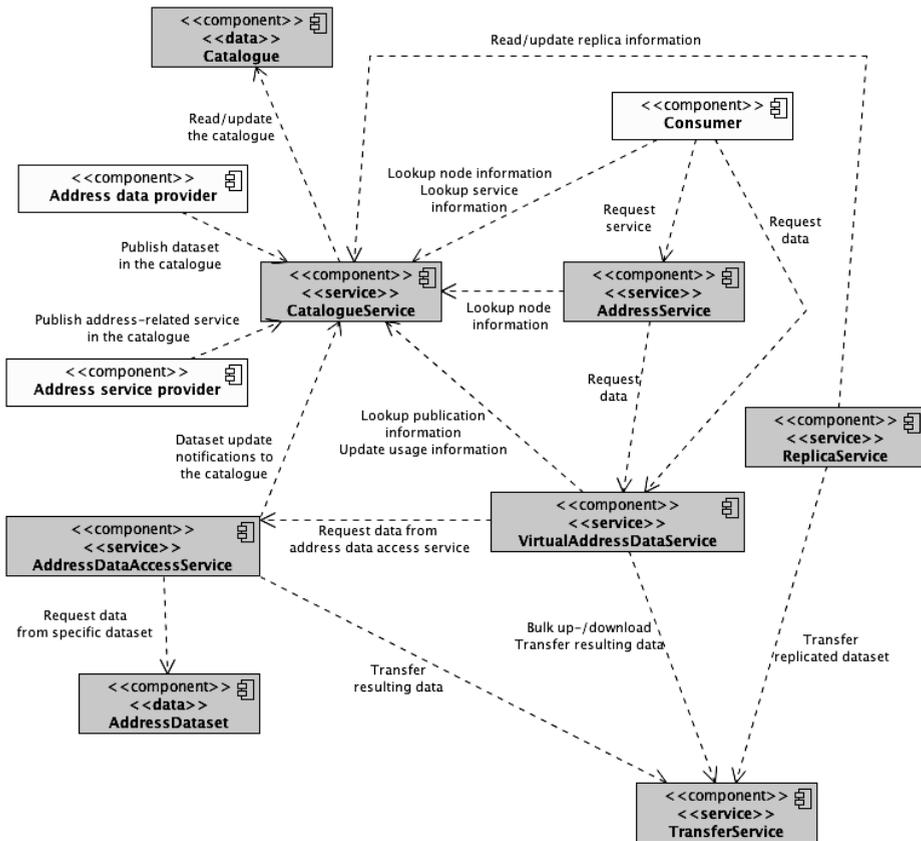
4.1 The Catalogue

The Compartimos catalogue contains the metadata that is required for the operation of the address data grid. Figure 3 shows the elements of the catalogue. The addressing systems describe the types of addresses that are contained in an address dataset, e.g. street and/or intersection address types. A dataset is published on the address data grid by associating it with an AddressDataAccessService. Information about where and how a dataset is replicated is also stored in the catalogue. Address service providers provide address-related services, such as geocoding or mapping, that operate on the single virtual address dataset. The node host provides the resources to host some or all of the catalogue, replica, transfer and virtual address data service, as described in the computational and engineering viewpoint. Any interaction with the catalogue takes place through the CatalogueService.

The size of the catalogue is determined by the size of the catalogue's collections. Based on the number of countries in the world, in an international address data grid these numbers are still relatively small in respect of what relational DBMS, object-oriented DBMS and XML databases are able to cope with, and there is no need to make special provision for huge volumes of data. The data model for the catalogue is

sufficiently simple to allow representation in a relational data model. It is important that the storage mechanism for the catalogue is platform independent so that it can be easily replicated and XML is therefore attractive.

Table 2: Technology choices for Compartimos objects.

Compartimos object	Technology choices
Catalogue	Relational data model vs. other models Relational DBMS vs. other DBMS such as XML or Object DBMSs ISO 19112:2003, *Geographic information – Spatial referencing by geographic identifiers* ISO 19115:2003, *Geographic information – Metadata* ISO 19119:2005, *Geographic information – Services* ISO 19139:2007, *Geographic information – Metadata – XML schema implementation* Dublin Core metadata element set Metadata in the Monitoring and Discovery System (MDS) - Globus Toolkit Metadata in the Replica Location Service (RLS) - Globus Toolkit
CatalogueService	OGC Catalogue Service Monitoring and Discovery System (MDS) - Globus Toolkit Replica Location Service (RLS) - Globus Toolkit
AddressDataAccessService	OGC Web Feature Service (WFS) ISO 19142 (draft), *Geographic information – Web feature service* OGSA-DAI data resources, Globus Toolkit compatible Combination of OGSA-DAI data resources and WFS

VirtualAddressDataService	Intelligence:
	    OGSA-DAI Distributed Query Processing (DQP)
	    Address-specific toolkits, such as AfriGIS Intiendo
	    Any capabilities available as OGC Web Processing Service (WPS)
	Data:
	    OGC Web Feature Service (WFS)
	    ISO 19142 (draft), *Geographic information – Web feature service*
	    OGSA-DAI data resources, Globus Toolkit compatible

AddressDataset	Technology independent, up to the data provider
AddressService	OGC Web Feature Service (WFS) OGC Web Processing Service (WPS)

The Compartimos catalogue includes metadata about address data, as well as the grid configuration. Using existing metadata standards holds the advantage that metadata can be readily imported into the Compartimos catalogue, and tools for capturing metadata according to these standards exist. Metadata about the address datasets can be stored according to existing standards, such as ISO 19115:2003*, Geographic information – Metadata*, with or without the ISO 19139:2007, *Geographic information - Metadata – XML schema implementation*. ISO 19119: 2005, *Geographic information – Services*, includes a data model for service metadata, which is applicable to the address-related services in Compartimos, while ISO 19112:2003, *Geographic information – Spatial referencing by geographic identifiers*, could be used for the addressing systems (Coetzee et al*., 2*008).

Figure 3: The address data catalogue.
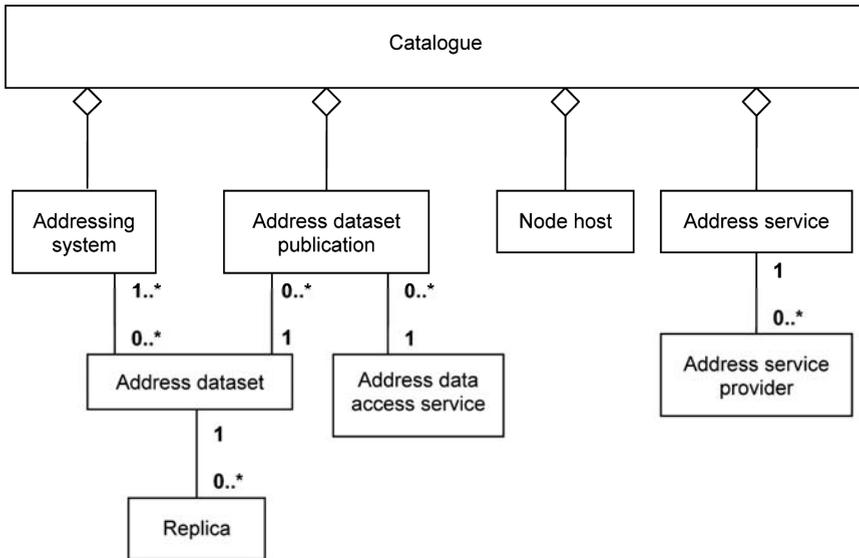
An alternative choice is the Dublin Core Metadata element set (www.dublincore.org), which has been adopted as an ISO standard (ISO 15836:2003), but Dublin Core does not cater for spatial data specifically and is not widely used in the geospatial community. The Globus Toolkit includes a catalogue capability (Singh et al*., 2003*; Zhao et al*., 2004*), but it does not address all the requirements for spatial data, such as the geographic extent of a dataset. The metadata that forms part of the Globus Toolkit's Replication Location Service (RLS) (http://www.globus.org/toolkit/data/rls/) is an option for replica information in the Compartimos catalogue. The metadata that is part of the Globus Toolkit's Monitoring and Discovery System (MDS) (http://www.globus.org/ toolkit/mds/) provides information about the available resources on the Grid and their status and is suitable to store information about the hosts in Compartimos. RLS and MDS were successfully integrated into a geospatial grid, reported by Di et al. (2008).

4.2 The CatalogueService

This service provides read and write access to the metadata in the Compartimos catalogue. Similar to the OGSA data architecture, the Compartimos CatalogueService provides *Publish* (add an entry), *Update* (modify an existing entry), and *Find* (apply query and return matching entries) services. The *Augment* (add additional properties for an entry created by someone else), *AddClassification* (add classification scheme) and *Classify* (classify an entry) services from the OGSA data architecture are not included in Compartimos. Compartimos applies to a very specific kind of data; therefore these services are not required.

The OGC has published a catalogue service implementation specification for the discovery and retrieval of metadata about spatial data and services, which can be implemented in conjunction with the above-mentioned ISO 19115:2003 and ISO 19139: 2007, as well as ISO 19119:2005 for service metadata. Wei et al. (2006) and Di et al. (2008) report on using the OGC catalogue service in their implementation of a geospatial grid for NASA. Zhao et al. (2004) report on a different option in (seemingly) the

same implementation of a geospatial grid for NASA, i.e. augmenting the Globus Toolkit's Metadata Catalogue Service (MCS) with the profile of the OGC Catalogue Service.

From a grid configuration point of view, there are two relevant services in the Globus Toolkit, the Monitoring and Discovery System (MDS) (for hosts in Compartimos) and the Replica Location Service (RLS) (for dataset replicas in Compartimos).

4.3 The AddressDataAccessService

The AddressDataAccessService converts address data from local proprietary format to an interoperable address data model (described in the information viewpoint of Compartimos, which is not included in this article), acting as an interpreter for a specific source of address data and providing a uniform access method to any dataset that is published in the address data grid. The OGSA data architecture proposes three generic data access operations for structured data: *Create*, *ExecuteQuery* and *BulkLoad*. *Create* creates an association between a data service and an underlying data resource, which may be created and populated as a result of this operation.

However, in an SDI environment, the main drive for an address data grid is to publish (as opposed to edit and maintain) address data. The *RegisterDataPublication* of the CatalogueService associates a dataset with an *AddressDataAccessService* (what *Create* does for OGSA). The Compartimos model provides for a one-to-many relationship between a dataset and an access service, thereby increasing scalability and enabling versioning of the interoperable address data model in the Compartimos catalogue.

ExecuteQuery is represented by the *GetAddress* operation of the AddressDataAccessService and *BulkLoad* is represented by the *UploadAddressData operation* in Compartimos, performing more or less the same functionality as their OGSA counterparts, albeit customised for address data.

The OGC WFS, which returns spatial data in vendor independent Geography Mark-up Language (GML) is a natural choice for this service. This implementation specification is currently in the process of being adopted as an ISO standard (ISO 19142 (draft)). However, functionality over and above normal WFS is required for the conversion to and from the Compartimos interoperable address data model. Aloisio et al. (2005a), Di et al. (2008), Goodenough et al. (2007), Wei et al. (2006) and Zhao et al. (2004) report on grid-enabling OGC web services, such as WFS and the OGC Web Map Service (WMS).

An alternative technology choice is OGSA-DAI (Data Access and Integration), which is compatible with the Globus Toolkit. However, OGSA-DAI has been developed for alphanumeric data and would require some extensions to accommodate spatial data. An advantage of OGSA-DAI is that its resources are already usable by other Globus Toolkit services. The choice of OGSA-DAI would influence the technology choice for other services, such as the CatalogueService and the VirtualAddressDataService.

4.4 The VirtualAddressDataService

The VirtualAddressDataService provides the required consolidation functionality to make the distributed heterogeneous address datasets appear to be a single virtual address dataset. This consolidation includes, for example, removing duplicates (due to multiple address data sources) and resolving ambiguities. The OGSA data architecture defines a set of operations of a Data Federation service for the logical integration of

multiple data services or resources to be accessed as if they were a single data service. In a way this corresponds to the VirtualAddressDataService in Compartimos, however, in OGSA *CreateFederation*, *AddSourceToFederation*, *AddAccessMechanism*, and *UpdateFederationAttributes* federate a wide variety of services ranging from input data resources to transformations of data and filters. In Compartimos a dataset (the resource) is automatically included in the federation when it is published in the catalogue and resources are, by definition, limited to address datasets. Therefore, VirtualAddressDataService only has the *GetAddress* and *UploadAddressData* operations, which mirror the AddressDataAccessService operations with the same names.

The VirtualAddressDataService is the centre of intelligence in Compartimos. Due to its diverse capabilities, each within its own field of specialisation, it makes sense to combine different components for an implementation of the VirtualAddressDataService. Below a few examples:

1. the OGSA-DAI Distributed Query Processing (DQP) (http://www.ogsadai.org/about/ogsa-dqp/) could be employed for distributed queries;
2. the address matching functionality provided by independent tools, such as the AfriGIS Intiendo address tool (Rahed et al., *2008*) could be used to remove duplicates and resolve disambiguities; and
3. any processing that is available as an OGC WPS, which is a standardised interface that facilitates the publishing of geospatial processes, and the discovery of and binding to those processes by clients.

Di et al. (2008) implemented their own mediator for geographic data, the Intelligent Grid Service Mediator (iGSM), while Shu *et al.* (2006) propose using OGSA-DAI DQP. Once the address data has been consolidated, similar to the AddressDataAccessService, an implementation of OGC WFS or OGSA-DAI data resources are potential technology choices for the external interface of the VirtualAddressDataService.

4.5 The AddressDataset

The Compartimos AddressDataset object refers to any address dataset that is published on the address data grid. In OGSA data architecture terminology this is the data source or data resource. In Compartimos the data provider determines how address data is stored. The AddressDataAccessService provides access to this proprietary data in the prescribed way (according to the interoperable data model). However, for optimal conversion efficiency it will make sense to store the 'raw' data according to the Compartimos interoperable data model, or as close to it as possible.

4.6 The AddressService

The AddressService refers to any address-related service, such as routing or mapping, that is offered by a third party on top of the single virtual address dataset in the grid. The list of operations of the AddressService is application dependent and defined by the service provider. The AddressService interacts with the VirtualAddressDataService when executing its address-related service. The functionality and interface of this service is determined by its purpose, and therefore not prescribed in Compartimos. For interoperability, it is important that this service uses the same standard and protocol as the other services in Compartimos.

5. RESULTS

Through the development of the Compartimos reference model, we have identified the essential components for sharing address data on a data grid in an SDI environment. We presented the computational viewpoint of Compartimos by describing the purpose of each object and relating it to the OGSA data architecture. We analysed technology choices for individual Compartimos objects and related these to current research. From this analysis it is evident that there is a need for collaboration between grid and geospatial communities to ensure harmonisation between respective standards and tools. The service-oriented approached followed in both OGSA, ISO 19100 and OGC will prove an advantage for collaboration and integration of respective technologies. The results of our analysis contribute towards the mutual understanding between these two communities. The analysis described in this article was a first investigation into the viability of the data grid approach to national address databases in an SDI and has led to further research questions that are discussed below.

The ISO 19100 series of standards together with OGC implementation specifications have been implemented in a number of SDIs (Aalders, 2005). To grid-enable these SDIs, requires grid-enabling these ISO standards and OGC implementation specifications. Aloisio et al. (2005a) and Di et al. (2008) write about such efforts, but more implementations are required to better understand the challenges under different circumstances. Such implementations would also promote the development of tools to streamline the grid-enablement. OGSA-DAI already provides uniform access to different relational databases, similar to OGC web service for heterogeneous geographic information. Future studies should investigate uniform access to spatial data through OGSA-DAI, with or without making use of OGC web services. Also, interesting would be a spatially enabled distributed query processing (DQP) of OGSA-DAI.

We based our initial research on the assumption that address data providers are mostly local authorities in an SDI that can be trusted. In a Web 2.0 world, where the citizens become the sources for data, this assumption does not hold anymore. Citizens, living at an address, are the best available source to verify an address, but the question is whether they can be trusted to provide accurate data. Goodchild (2008) and Craglia et al. (2008) also raise this questions and future work should investigate how such a 'wikification' of address data can be integrated into Compartimos.

Compartimos was developed for address data in an SDI and future research should expand Compartimos for other types of spatial data. Incorporating recent research findings on ontologies for interoperability would be relevant (Brodeur 2004, Shadbolt et al., 2006). A reference model for data grids that caters for all kinds of geographic information could be seen as the first step along the path of standardising geospatial data grids. Also, research is required to better understand the requirements for grid-enabling SDIs in terms of non-technical aspects, such as policies, legislation, agreements, human and economic resources, and organisational aspects.

Finally, this research project started in 2005, before the current hype of 'cloud computing'. However, clouds, such as those by Amazon, IBM, Microsoft and the like, also stand in line as the enabling platform for data sharing in an SDI. Instead of investing servers and bandwidth, local authorities could buy scalable computing power and data storage in a cloud without having to support an IT infrastructure. Thus, the viability of data sharing in an SDI 'in the clouds' should be investigated.

6. CONCLUSION

We presented a scenario that describes for the first time how data grids can solve the problem of sharing address data in an SDI. Our approach is both a novel application for data grids as well as a novel technology in SDI environments, and thus improves the understanding of the requirements and issues related to applying grid technology in SDIs. We presented the computational viewpoint of the Compartimos reference model and analysed technology choices for individual Compartimos objects. From this analysis it is evident that collaboration on standards between the grid and SDI communities is imperative. Our analysis advances the understanding of the requirements for, and the use of, the data grid approach in a specific application domain, namely address data in an SDI. Thus, the article sheds light on a future technological solution that could overcome some of the data sharing impediments that are experienced in SDIs today, and it discusses research that will support the vision of an address data grid for in an SDI.

REFERENCES

Aalders H.J.G.L. (2005). An introduction to spatial metadata standards in the world, in Moellering H., Aalders H.J.G.L. and Crane A. (eds.). *World Spatial Metadata standards*, Oxford, United Kingdom: Elsevier.

Aloisio G., Cafaro M., Conte D., Tiore S., Epicoco I., Marra G.P., Quarta G. (2005a). "A Grid-Enabled Web Map Server, *International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume I*, pp. 298-303.

Aloisio G., Cafaro M., Fiore S., Wuarta G. (2005b). "A grid-based architecture for earth observation data access", *2005 ACM Symposium on Applied Computing*, March 13-17 2005, Santa Fe, New Mexico, USA.

Aydin G., Sayar A., Gadgil H., Aktas M., Fox G.C., Ko S., Bulut H. and Pierce M.E. (2008). Building and applying geographical information system Grids, *Concurrency and Computation: Practice and Experience*, 20(14): 1653-1695.

Brodeur J. (2004). "Interoperability of Geographic Information: A Communication - Based Prototype", *Proceedings of The 8th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI 2004)*, Orlando, Florida, pp 327-332.

Coetzee S. and Bishop J. (2008). Address databases for national SDI: Comparing the novel data grid approach to data harvesting and federated databases, *International Journal of Geographical Information Science*, 26 September 2008, available online ahead of print edition at http://www.tandf.co.uk/journals/tf/13658816.html.

Coetzee S., Cooper A.K., Lind M., McCart Wells M., Yurman S.W., Wells E., Griffiths N. and Nicholson M.J.L. (2008). "Towards an international address standard", *GSDI-10 Conference*, Trinidad and Tobago, 25 – 29 February 2008.

Craglia M., Goodchild M.F., Annoni A., Camara G., Gould M., Kuhn W., Mark D., Masser I., Maguire D., Liang S. and Parsons E. (2008), Next-Generation Digital Earth, a position paper from the Vespucci Initiative for the advancement of Geographic Information Science, *International Journal of Spatial Data Infrastructures Research*, 3: 146-167.

Di L., Chen A., Yang W., Liu Y., Wei Y., Mehrotra P., Hu C., Williams D. (2008). The development of a geospatial data Grid by integrating OGC web services with Globus-based Grid technology, *Concurrency and Computation: Practice and Experience*, 20(14): 1617-1635.

Fahey D. and Finch F. (2006). *GeoDirectory Technical Guide*. An Post GeoDirectory Limited, at: http://www.geodirectory.ie/Downloads.aspx.

Foster I., Kesselman C., Nick J.M. and Tuecke S. (2002)*. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration*, at: http://forge.gridforum.org/sf/go/doc13483?nav=1.

Foster I., Kesselman C. and Tuecke S. (2001). The Anatomy of the Grid – enabling scalable virtual organizations, *International Journal of High Performance Computing Applications*, 15(3): 200-222.

Goodchild M.F. (2008). "Spatial accuracy 2.0", *Proceedings of the Eighth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Liverpool: World Academic Union, pp. 1-7.

Goodenough D.G., Hao C., Liping D., Aimin G., Yaxing W., Dyk A., and Hobart G. (2007). "Grid-enabled OGC environment for EO data and services in support of Canada's forest applications", *IEEE International Geoscience and Remote Sensing Symposium 2007 (IGARSS 2007)*, 23-28 July, pp. 4773-4776.

Hua L., De-ren L., Xin-yan Z. (2005). "Large volume spatial data management based on grid computing", *Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '05)*, 25-29 July.

ISO 15836:2003 (2003). *Information and documentation - The Dublin Core metadata element set*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 19112:2003 (2003)*. Geographic information – Spatial referencing by geographic identifiers*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 19115:2003 (2003). *Geographic information - Metadata*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 19119:2005 (2005). *Geographic information - Services*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 19139:2007 (2007). *Geographic information - Metadata - XML Schema Language*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 19142 (draft) (2008). *Geographic information - Web Feature Service*, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO/IEC 10746:1998 (1998). *Information technology - Open Distributed Processing - Reference Model,* International Organization for Standardization (ISO), Geneva, Switzerland.

Koutroumpas M. and Higgins C. (2008). "A Pipeline Processing Approach to GIS", *11th AGILE International Conference on Geographic Information Science 2008*, Girona. Spain.

Levoleger K. and Corbin C. (*Ed.) (2005). *Survey of European National Addressing as of May 2005*, European Umbrella Organisation for Geographic Information (EUROGI), at: http://eurog.vbnprep.com/POOLED/DOCUMENTS/a101730/EUROGI _Address_Survey_Resp_V3Final.pdf.

Nicholson M (2007). "The address: improving public service delivery", *45th Annual URISA Conference*, 20-23 August 2007, Washington DC, USA.

OGC OGF (2007). Memorandum of Understanding, *Directions Magazine*, at: http://apb.directionsmag.com/index.php?url=archives/3567-OGC-OGF-OU.html& serendipity[cview]=linear.

Open Geospatial Consortium (2007). *OpenGIS Catalogue Service Implementation Specification 2.0.2*, at: http://www.opengeospatial.org/standards/cat.

Open Geospatial Consortium (2005). *OpenGIS Web Feature Service Implementation Specification 1.1.0*, at: http://www.opengeospatial.org/standards/wfs.

Open Geospatial Consortium (2003). *OpenGIS Web Map Service Implementation Specification 1.3.0*, at: http://www.opengeospatial.org/standards/ wms.

Open Geospatial Consortium (2007). *OpenGIS Web Processing Service Implementation Specification 1.0.0*, at: http://www.opengeospatial.org/standards/wps.

Open Grid Forum (OGF) (2007a). *OGSA Data Architecture*, Editors: D. Berry, A. Luniewski, and M. Antonioletti, at: http://www.ogf.org/gf/docs/?final.

Open Grid Forum (OGF) (2007b). *Open Grid Services Architecture Glossary of Terms Version 1.6*, Editor: J. Treadwell, at: http://www.ogf.org/gf/docs/ ?final.

Open Grid Forum (OGF) (2006). *The Open Grid Services Architecture, Version 1.5*, Editors: I. Foster, H. Kishimoto and A. Savva, at: http://www.ogf.org/gf/docs/ ?final.

Paull D. (2003). *A Geocoded National Address File for Australia: The G-NAF What, Why, Who and When*. PSMA Australia Limited, at: http://www.psma.com.au/ resources/the-g-naf-what-why-who-and-when.

Rahed A.A., Coetzee S. and Rademeyer M. (2008). "A data model for efficient address data representation - Lessons learnt from the Intiendo address matching tool", *Proceedings of the academic track of the 2008 Free and Open Source Software for Geospatial (FOSS4G) Conference, incorporating the GISSA 2008 Conference*, 29 September - 3 October 2008, Cape Town, South Africa.

Shadbolt N., Berners-Lee T., Hall W. (2006). The Semantic Web Revisited, "*IEEE Intelligent Systems*", May/June 2006, pp. 96-101.

Shu Y., Zhang J.F., Zhou X. (2006). "A Grid-Enabled Architecture for Geospatial Data Sharing", *IEEE Asia-Pacific Conference on Services Computing* (APSCC '06), December 2006, pp. 369–375.

Singh G., Bharathi, Chervenak A., Deelman E., Kesselman C., Manohar M., Patil S. and Pearlman L. (2003). "A Metadata Catalog Service for Data Intensive Applications", *Proceedings of the 2003 ACM/IEEE conference on Supercomputing (SC '03)*, November 2003.

Wei X., Yue P., Dadi U., Min M., Hu C., Di L. (2006). Effective Acquisition of Geospatial Data Products in a Collaborative Grid Environment, "*IEEE International Conference on Services Computing*", SCC '06, Sept. 2006, pp. 455-462.

Williamson I., Grant D. and Rajabifard A. (2005). "Land administration and spatial data infrastructures", *Proceedings of 8th Global Spatial Data Infrastructure Conference (GSDI-8)*, 16-21 April 2005, Cairo, Egypt.

Xue Y., Wan W., Li Y., Guang J., Bai L., Wang Y., and Ai J. (2008). Quantitative retrieval of geophysical parameters using satellite data, *IEEE Computer*, 41(4): 33-39.

Zhao P., Chen A., Liu Y., Di L., Yang Q., Li P. (2004). "Grid Metadata Catalog Service-Based OGC Web Registry Service", *GIS'04*, November 12–13, 2004, Washington D. C., USA.

SDI and Metadata Entry and Updating Tools

Abbas Rajabifard, Mohsen Kalantari and Andrew Binns

Centre for Spatial Data Infrastructures and Land Administration
Department of Geomatics, The University of Melbourne
{abbas.r, saeidks, a.binns}@unimelb.edu.au

Abstract
Metadata is a vital tool for management of spatial data and plays a key role in any spatial data infrastructure (SDI) initiative. It provides users of spatial data with information about the purpose, quality, actuality and accuracy and many more of spatial datasets. Metadata performs crucial functions that make spatial data interoperable. However, current metadata models and standards are complex and very difficult to handle. Also, metadata for spatial datasets is often missing or incomplete and is acquired in heterogeneous ways.
Typically, it is acquired after the spatial data itself, through lengthy and complex efforts. Metadata is usually created and stored separately to the actual data set it relates to. Separation of storage creates two independent data sets that must be managed and updated - spatial data and metadata. These are often redundant and inconsistent. Thus, the reliability of spatial data and the extent it can be used are often unclear. To respond to this issue, this article discusses the importance of having an integrated system for both spatial data and metadata in which that metadata and spatial data can be integrated within the one spatial dataset, so that when spatial data is updated, metadata related to that data is also automatically updated. The article highlights the significance of spatial data and metadata integration through developing a set of criteria for metadata application development and the result of applying the criteria against a selection of metadata entry tools (METs).

Keywords: Spatial Data infrastructure (SDI), Metadata integration, Metadata Entry Tools, Metadata Update.

1. INTRODUCTION

SDI is an enabling platform that facilitates access to spatial data and sharing spatial resources and tools among different practitioners. The creation of an enabling platform for the delivery of the spatial data and tools will allow users from diverse backgrounds to work together with current technologies to meet the dynamic market place (Rajabifard et al., 2005). Within the enabling platform, metadata plays a key role to facilitate accessing up-to-date and high quality spatial data and services (Williamson et al., 2003).

Metadata is data about data and is a vital component of spatial data. Users of spatial data need to know who created it, who maintains it, its scale and accuracy, and more. It not only provides users of spatial data with information about the purpose, quality, actuality and accuracy of spatial data sets, but also performs vital functions that make spatial data interoperable, that is, capable of being shared between systems. Metadata enables both professional and non-professional spatial users to find the most appropriate, applicable and accessible datasets for use.

According to international definition (ISO/TC211 2001), metadata comprises "… a schema required for describing geographic information and services. Information about

the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data …". Metadata are according to ISO/TC211: "applicable to the cataloguing of datasets, clearinghouse activities, and the full description of datasets."

Consistent description of the content and use of spatial data requires standards to define which attributes are needed and the structure of a metadata schema or model. This means that metadata can be used both for human interpretation of data sets, and computer processing and utilisation in search engines.

Different countries and jurisdictions are building many extensive and expensive spatial data systems in which access to up-to-date metadata is vital to delivering high quality spatial data services to their vast areas (Crompvoets et al., 2004). Meanwhile, by increasing distribution of spatial data over the Internet, the demand for spatial metadata to describe spatial data is growing in the networked environment. However, current metadata models are often complex and very difficult to handle. The creation and maintenance of spatial metadata is also seen as an expensive overhead by the spatial data industry (Philips et al., 1998; Najar et al., 2007).

Meanwhile, an integrated model for metadata and spatial datasets will benefit the spatial data industry in general, as well as all industries that increasingly utilise spatial data in their day-to-day tasks. This will enable metadata to be maintained dynamically in a way that addresses the more real-time requirements of people and organisations that use spatial data.

This article aims at discussing the significance of an integrated approach for handling spatial metadata by combining spatial data and metadata in a seamless approach. The methodology used in this article is based on assessing a number of metadata applications in order to reveal the importance of integrated approach. The article is based an ongoing research by authors on the automation of spatial metadata update process.

The article first develops a number of important criteria in coding, developing, installing and exploiting the metadata entry tool. It then reviews and assesses a number of existing metadata entry tools. Based on the result of the assessment, the article presents a discussion on the importance of having an integrated system for both spatial data and metadata.

2. METADATA COLLECTION AND METADATA ENTRY TOOLS

Typically, metadata is collected after the spatial data itself, through lengthy, complex efforts. Metadata for spatial data sets is often missing or incomplete and is acquired in heterogeneous ways. Metadata is usually collected and stored separately to the actual data set it relates to, and is often managed by people with a limited knowledge of its value. Separation of storage creates two independent data sets that must be managed and updated - spatial data and metadata. These are often redundant and inconsistent. Thus, the reliability of spatial data and the extent it can be used are often unclear.

The spatial industry has already identified the major factors about metadata collection and developed a number of applications to manage it. Crucial questions when developing a metadata entry tools are: How can the process of entering metadata be automated for the users? What functionalities should a metadata entry tool provide?, and How can the metadata collection process be facilitated?

A critical problem for metadata collection applications is flexibility. A metadata application must be sufficiently flexible to cope with changes to metadata standards over time and to allow users to extend a standard to cope with local requirements (Waugh, 1998). Also, a key component of supporting flexible metadata applications is user friendliness which can facilitate metadata entry and update.

Overall, the challenge of developing a right metadata entry tool (MET) lies in the structured arrangement of a substantial number of different disciplines and the examination of a large number of factors and issues that are discussed below.

In order to develop a MET, this article categorised the criteria into four main groups including technical requirement, compliance with international standards, user friendly interface and finally availability of necessary functions for handling metadata records.

2.1 Technical requirements

Consideration of technical criteria includes ensuring proper technology development with easy deployment and an efficient database technology to support accessing and maintaining metadata. Technical criteria also should consider the outlay of a MET with a low cost and low risk.

2.1.1 Development technology

There are generally two options for the development of a MET: (1) standalone and (2) web based. As an entry tool a MET is not necessarily required to be a web based application. However, for integration in online search engines, spatial clearinghouses, web base spatial libraries and web mapping systems, a web based development technology will have a better position comparing standalone technologies. Also using a web based technology the tool can be available any time and any where for different range of users.

Meanwhile, the Web places some specific constraints on the development of METs such as the ability to deal with a variety of protocols and formats (e.g. graphics) and programming tasks; performance in terms of speed and size of communication; safety; platform independence; protection of intellectual property; and the basic ability to deal with other Web tools and languages. The web based approach can be helpful for integration of spatial data and metadata which are distributed over the network.

2.1.2 Database Connection Technology

Database Connection Technology provides the connection between a MET and a spatial metadata database. The connection means a link having a formal and published definition for communication in order to record, edit and retrieve metadata. This definition identifies the interface that MET must use to issue query and receive database content through the link.

In this regard, when choosing a MET, the first consideration is the type of databases that the metadata is stored in. Based on the type of database technology, a proper connection technology can be chosen. For instance, Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC) are important technologies as they are available on many disparate platforms and they provide common interfaces to several different database products (Shekhar and Chawla, 2003). More importantly for the

integration of metadata and spatial data a comprehensive seamless data should be developed.

2.1.3 Robust code

For further development, when preparing an open source MET, having robust code is essential. Robust coding is a style of programming that prevents abnormal termination or unexpected actions. A robust coded software is easy to follow, well commented, well tested, well-named, has good error messages and can be easily maintained and, if necessary, modified. However this criterion is not applicable when choosing proprietary software with metadata entry facilities.

2.1.4 Easy deployment

When installing a MET, the general deployment process consists of several interrelated activities and transitions between them. With this in mind that every software system is unique, a complete and easy deployment process for a MET should at least include release, installation, activation, deactivation, adaptation and un-installation.

2.1.5 Open source or freeware software

A vendor independent open source modular coding, and, to a lesser extent, freeware, can enable ease of adapting a MET and of future development. For preference, the language should be of an industry standard to match available skills that can be purchased cost effectively from the market place. Similar to the robust code criterion, this factor is not applicable for proprietary software.

2.1.6 Standards

The MET must support international metadata standards that support spatial metadata such as Dublin Core Metadata Standard or ISO 19115:2003 Geographic information – Metadata elements whether core, conditional, mandatory or optional. Besides, the tool should support the implementation of the metadata standards such as ISO 19139:2006, Geographic information – Metadata – XML schema implementation (Moellering et al., 2005).

2.1.7 User friendly interface

Ease of use in a MET includes ensuring consideration of providing an intuitive, simple and familiar user interface to perform the necessary functions and applications. The familiar interfaces would help to hide some very complex operations and provide good navigation logic.

The navigation logic should enable novice, 'low-end' users to easily find their way around. The interface should enable novice low-end users to easily create and edit metadata records. This includes consideration of operational and navigational design, graphical and visual design, help information and assistance, the process of entering, editing and retrieving metadata records, and finally technical issues such as response and navigation speed.

2.1.8 Functionalities

A MET must allow metadata records to be created, edited, copied, compiled, searched, saved and deleted in accordance with the standard Metadata Profile, including not only the mandatory elements but also the remainder of the standard's elements. The basic functionality of the MET should cater for the needs of authors, users and managers of metadata.

The MET should import and export different standards' compliant metadata without the loss of content. To enable ease of human readability and presentation the MET should also produce valid text files such as HTML or RTF from the content of metadata. The application should also print metadata records using a print friendly layout.

The tool should provide support for reuse and linkage of contact details. Normally contact details should be entered once, and re-used for all subsequent edits. This enables duplication of existing records for use as a first draft of a new metadata record.

The tool should also be linked to the spatial application to be able to update the metadata when any change occurs to the spatial data. This significantly can reduce redundancies in the metadata database. More importantly the reliability of metadata linked to spatial data will increase. An integrated approach for handling metadata and spatial data together will require an integrated data model and integrated application.

The ability to search for records using spatial, free text, keyword and other search terms through a simple or advanced search will increase the usability of the tool. The integrated help facility should also be available through the tool with table of contents, search facility and links to related web documents or websites, and a context help linking each of the elements to a summary and the detailed text of relevant sections of the user guidelines.

Finally in a MET, the metadata administrator should be able to generate reports, statistical data based on specific metadata (elements), agency inputs, exports, and searches. Table 1 summarises the criteria developed for the assessment of metadata entry tools.

3. METADATA ENTRY TOOL ASSESSMENT

A three stage methodology of assessment has been developed for evaluating a selection of metadata tools against the developed criteria. Figure 1 illustrates process flow and stages of the methodology.

Stage 1
In this stage, a collection of related documents about the tools have been reviewed and explored. This stage helped with understanding of background and objectives of the tools development. The criteria developed above were finetuned in accordance with the overall purpose of the tools.

Stage 2
In parallel to stage 1, the selected metadata entry tools were installed and the deployment process of each tool was investigated. Based on the results of the first stage, a questionnaire was developed and designed to be used by a number of five clients testing the user friendliness of each tool. The clients were selected from different background related to spatial science.

Table 1: Selection criteria for Metadata Entry Tools.

Category	Criteria
Technical	Web based development technology
	Database technology and access
	Robust Code
	Easy Deployment
	Open source or Freeware software
Standards	Standard 1
	Standard 2
	Standard n
User Interface	Similarity to common software
	Intuitive navigation logic
	Intuitive interface enabling easy creation and editing of metadata records
Functionalities	Contact reuse
	Record duplication
	Automatic and integrated update
	Search tools
	Edit tools
	Import and export tools
	Format Conversion
	Print Friendly layout
	Help facility

Stage 3

Having developed the assessment criteria, the stage 3 was dedicated for evaluating the selected tools. Functionality available, technical requirements and the tools user friendliness were explored.

This section introduces the three metadata entry tools that have been studied and evaluated against the developed criteria. Each of these tools is explored in detail in this section, including background information essential to effectively assessing each of the tools in relation to the overall objectives of the article.

3.1 IDEC MetaD

The Infraestructura de Dades Espacials de Catalunya (IDEC) project is a common initiative of the Cartographic Institute of Catalonia (ICC), the Department of Territorial policy and public works, the Secretariat for Telecommunications and Information Society, and the Department of Universities, Research and Information Society of the Generalitat de Catalunya, within the framework of the III Research Plan 2001-2004 and the Strategic Plan for the Information Society (Catalonia on the Network). Its aim is to create a Spatial Data Infrastructure for Catalunya (MetaD, 2007).

Within IDEC initiatives, MetaD allows for the creation, edition and consultation of metadata stored in a data base. This includes the creation of new metadata records, as well as the maintenance of records already created. The tool incorporates various functionalities including an XML viewer. XML Viewer is a tool that allows the display of XML

Figure 1: Metadata Entry Tools evaluation methodology.
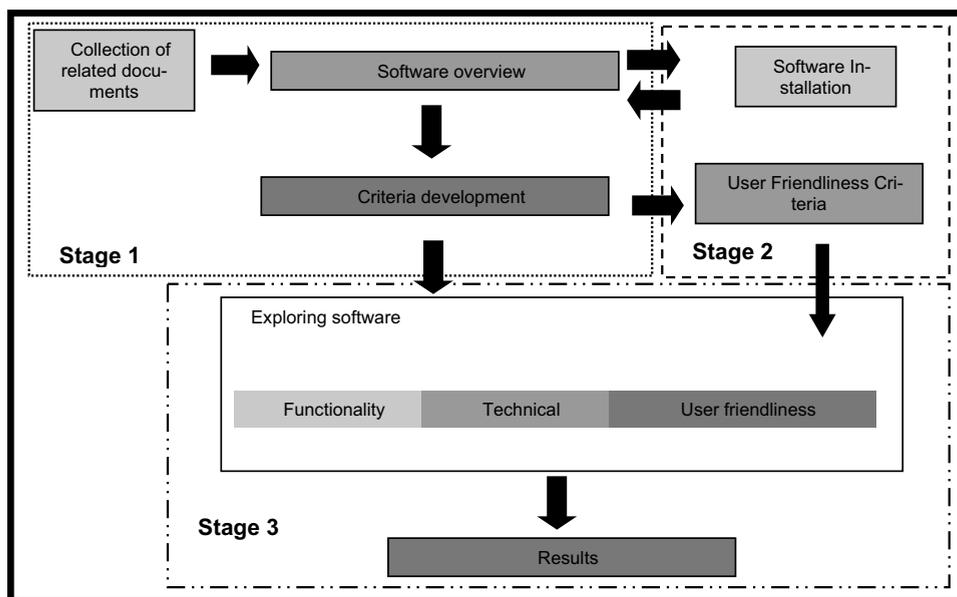
files generated by other programs. The application also facilitates the incorporation and maintenance of thesauri of subjects and key words, with the purpose of facilitating the edition of metadata. This includes the ability to incorporate a users own list of key words. The software also allows the user to incorporate and exchange metadata with other organisations or departments stored within other data bases.

MetaD has been developed with visual basic (VB) programming language, which limits the ability of this tool to be developed for web based applications. However, MetaD has the ability to be connected to a database in the network.
MetaD uses Microsoft Access as the database technology. Access is used by small businesses create ad hoc customised desktop systems for handling the creation and manipulation of data. Some professional application developers use Access for rapid application development, especially for the creation of prototypes and standalone applications that serve as tools for on-the-road salesmen. Access does not scale well if data access is via a network, so applications that are used by more than a handful of people tend to rely on Client-Server based solutions.

The structure and terminology of this tool follows the Standard ISO 19115. As this standard is very generic, the IDEC has adapted it to the particularities of Spain, giving as result the elements that appear in the application.

In summary, the MetaD application uses a graphical style interface, in order to simplify the complexity of the metadata standard for users of the application. The application also has the ability to convert coordinates into other formats which are required for specific standards.

The objectives of the application are not just the ability to create metadata records for geographic information, but also to be able to easily maintain or change metadata re-

cords as required. The application allows users to change former metadata records created. Once created, metadata can be exported to catalogues as XML files, either as single record sheets, or the entire file. Table 2 summarises the tool against the criteria developed.

Table 2: MET assessment result for MetaD.

Category	Criteria	
Technical	Web based development technology	No
	Database technology and access	Yes
	Robust Code	Yes
	Easy Deployment	Yes
	Open source or Freeware software	Yes
Standards	ISO 19115 support	Yes
	ISO 19139 compliant XML metadata	Yes
User Interface	Similarity to common software	Yes
	Intuitive navigation logic	Yes
	Intuitive interface enabling easy creation and editing of metadata	Yes
Functionalities	Contact reuse	Yes
	Record duplication	Yes
	Automatic and Integrated Update	No
	Search tools	Yes
	Edit tools	Yes
	Import and export tools	Yes
	Format Conversion	Yes
	Print Friendly layout	No
	Help facility	Yes
	Multi level access	Yes

3.2 CatMdEdit

CatMDEdit is a metadata editor tool that facilitates the documentation of resources, with special focus on the description of geographic information resources. CatMDEdit has been developed by the TeIDE consortium. TeIDE is a Spanish consortium constituted by the R&D groups of the University of Zaragoza, the Universitat Jaume I, and the Polytechnic University of Madrid.

The CatMDEdit library is freeware; it can be redistributed and/or modified under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either the current version of the License, or any later version.

The tool has been implemented in Java and is multi-platform (Windows, Unix). As it has been developed in Java and the storage of metadata records is managed directly through the file system, the application can be deployed in any platform with the minimum requirement of having Java installed.

CatMdEdit enables users to create consistent metadata describing spatial data resources. The functionality includes basic services for creating and editing metadata, as

well as more enhanced tools to improve the quality of metadata, including the thesaurus management tool and a metadata automatic generation tool. The thesaurus tool facilitates mapping between a selected vocabulary and a large collection of datasets.

The automatic metadata generation tool is able to derive metadata from data sources by means of interconnection with commercial GIS tools or proprietary software. Examples of derived metadata are information about spatial reference systems, number and type of geographic features, extension covered by a dataset, or information about the entities and attributes of alphanumerical related data (Ballari et al., 2006).

The basic functionality works with any relational database management system, such as MS Access, Oracle, mySQL, which is responsible for the storage of the metadata entries using a SQL-92 metadata database model. The more advanced functionality is oriented to more advanced metadata contributors, as well as to catalogue administrators in charge of management and improvement of metadata controlled under the geographic catalogue. In this case, the system works with Oracle 8i because of its special capacities for the use of spatial objects, text management and thesaurus (Ballari et al., 2006).

CatMDEdit metadata edition is in conformance with the "ISO 19115. Geographic Information-Metadata" standard. Four interfaces are provided for the edition of metadata records:

1. A detailed interface following the ISO 19115 comprehensive profile.

2. A reduced interface following the *Ncleo Espanol de Metadatos* (NEM). NEM, a subset of ISO 19115, is a recommendation under development which has been defined by the Spanish National Geographical High Board (*Consejo Superior Geografico*). This subset includes all the elements defined for the ISO19115 Core metadata profile ('Core metadata for geographic datasets').

3. An interface following the SDIGER - INSPIRE metadata profile, which has been developed under the framework of the SDIGER project. This profile is based on the international standard ISO 19115 that was customised to meet the requirements set up in the proposal for a Directive of the European Parliament and of the Council establishing an infrastructure for spatial information in the European Community (INSPIRE).

4. An interface following the SDIGER - WFD metadata profile, which has been developed under the framework of the SDIGER project. This profile is based on the international standard ISO 19115 customised to follow the guidelines for metadata to implement the GIS Elements of the Water Framework Directive.

CatMDEdit metadata edition is also in conformance with the SDIGER - Dublin Core Metadata Application Profile for geographical data mining, which has been developed under the framework of the SDIGER project. This profile is based on the Dublin Core Spatial Application Profile developed by the European Standardisation Committee to improve the discovery of geographic information.

CatMDEdit permits the reuse of contact information (e.g., name, address, telephone …) of organisations and individuals, which must be filled in several metadata elements. The contact information about a responsible party is inserted only once and used whenever it is required. It also permits creation of an identical copy of the selected

element in the metadata record. This tool allows making all common edit operations on the record selection window or on any other metadata edition sub-window. Apart from edit, save, cancel and refresh operations, it also adds operations to treat rows in windows that contain multi-valued table-represented elements. These tables appear both in the record selection window and in the metadata edition windows that contain several occurrences of a metadata element.

CatMDEdit supports exchange of metadata records according to different standards and formats. It also permits import and export of ISO 19115 metadata in XML format in compliance with ISO19139 technical specification. Additionally, there is interoperability with other metadata standards apart from ISO19115. The application allows input and output XML files in conformance with the standards CSDGM (Content Standard for Digital Geospatial Metadata, defined by U.S. FGDC), Qualified Dublin Core, SDIGER - Dublin Core Metadata Application Profile for geographical data mining, or MIGRA (Spanish standard for geographic information exchange). Table 3 summarises the tool against the criteria developed.

Table 3: MET assessment result for CatMDEdit.

Category	Criteria	
Technical	Web based development technology	Yes
	Database technology and access	Yes
	Robust Code	Yes
	Easy Deployment	Yes
	Open source or Freeware software	Yes
Standards	ISO 19115 support	Yes
	ISO 19139 compliant XML metadata	Yes
	SDIGER- INSPIRE and WFD	Yes
	SDIGER- Dublin Core	Yes
User Interface	Similarity to common software	Yes
	Intuitive navigation logic	Yes
	Intuitive interface enabling easy creation and editing of metadata records	Yes
Functionalities	Contact reuse	Yes
	Automatic Update	No
	Record duplication	Yes
	Search tools	Yes
	Edit tools	Yes
	Import and export tools	Yes
	Format Conversion	Yes
	Print Friendly layout	No
	Help facility	Yes
	Multi level access	No

3.3 Geonetwork

GeoNetwork open source is a standard based, Free and Open Source catalogue application to manage spatially reference resources through the web, designed to enable

access to geo-referenced databases, cartographic products and related metadata from a variety of sources (Geonetwork Community, 2007).

The development of GeoNetwork has been undertaken by the Food and Agricultural Organisation (FAO) on the United Nations. It is beginning to attract attention with its adoption by a number of international programs, countries and regional SDI initiatives, including in the USA, France, Czech Republic and Hungary.

GeoNetwork opensource is a standardised and decentralised spatial data management environment, designed to enable access to geo-referenced databases, cartographic products and related metadata from a variety of sources, enhancing the spatial data exchange and sharing between organisations and their audience, using the capacities of the internet.

GeoNetwork comes with an internal DBMS server, the McKoi SQL database. For more than one connection to the same database, Mckoi SQL Database is a multi-threaded multi-user server. Mckoi SQL Database is an SQL Database written entirely in Java.

GeoNetwork opensource has been developed to connect spatial data communities and their data using a modern architecture, which is at the same time powerful and low cost, based on the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols (from ISO/TC211 and OGC). It supports metadata creation based on the ISO19115, FGDC and Dublin Core standards.

For contact reuse and record duplication, a template can be fully customised online and can be pre-filled with repetitive content (contact information for example). Geonetwork benefits from a search index capable of handling large amounts of metadata. The indexing is build using Jakarta Lucene. The full sets of query parameters used to search the GeoNetwork catalog are also available for harvesting jobs. It provides a uniform way of searching through the metadata.

Geonetwork provides a method storing and indexing of metadata in its original format. It also provides editing the different metadata standards online in default, advanced and XML mode. It is also possible access to the full set of ISO19115 and FGDC metadata elements through the generic online editor Geonetwork also has a Metadata Template system. This system allows to quickly creating new metadata entries. A template can be fully customised online and can be pre-filled with repetitive content (contact information for example). Templates can also be searched in the same way normal metadata is searched. But only editors and administrators have access to templates. Further, more templates can be created for specific user groups.

Geonetwork permits import of XML formatted metadata and possible conversion of the input file through XSL transformation. It also supports batch import of XML formatted metadata and possible conversion of the input files through XSL transformation. Table 4 summarises the assessment result for Geonetwork.

3.4 Assessment result

During the assessment, the aim was to observe clients using the products in an as realistic a situation as possible, to discover the effectiveness of the assessment methodology. The development of metadata applications not only should focus on technical capabilities of the tool, but also it should concentrate on usability and functionality of

Table 4: Assessment result for Geonetwork.

Category	Criteria	
Technical	Web based development technology	Yes
	Database technology and access	Yes
	Robust Code	Yes
	Easy Deployment	Yes
	Open source or Freeware software	Yes
Standards	ISO 19115 support	Yes
	ISO 19139 compliant XML metadata	Yes
	Dublin Core	Yes
User Interface	Similarity to common software	Yes
	Intuitive navigation logic	Yes
	Intuitive interface enabling easy creation and editing of metadata records	Yes
Functionalities	Contact reuse	Yes
	Automatic Update	No
	Record duplication	Yes
	Search tools	Yes
	Edit tools	Yes
	Import and export tools	Yes
	Format Conversion	Yes
	Print Friendly layout	No
	Help facility	Yes
	Multi level access	Yes

the functions. This is often caused by pressure to develop systems based on technicians' expectations instead of overall non-professionals' needs.

For instance, overall observation of five clients with MetaD user interface showed the tool was simply understood with them. The clients did not have too many difficulties with understanding instructions, manipulating parts, or interpreting functionalities. Another example was the user interface testing with CatMEdit, which did not uncover difficulties with learning operating and navigating in software for five clients and entering and editing metadata was a simple process, but it still seemed that CatMEdit needed improvement.

With a large community supporting development of Geonetwork, clients can easily find information and assistance if needed. Geonetwork opensource enjoyed a user friendly interface to search and manage Metadata. A comprehensive metadata editor also supported its popularity among the five clients. The web based catalogue application and the integrated InterMap interactive map client application made it very attractive for clients. The tool is easy to learn and operate.

As illustrated in Tables 2, 3 and 4, none of the tools has an integrated approach to handle spatial data and metadata together in a seamless database. Integration of

metadata and spatial data is new and enables the spatial data to carry their own metadata description with them. The next section discusses the significance of an integrated method for managing spatial data and its metadata together.

4. METADATA INTEGRATION-SIGNIFICANCE AND INNOVATION

The research in metadata integration should focus on utilise metadata standards and developments in order to combine metadata and spatial data within an integrated package, so that the process of updating or creating spatial data and metadata – where feasible – becomes one process rather than two.

This approach distinguishes between already existing spatial data models, which have to be extended. If common metadata-spatial data sets exist, the concept of views allows metadata and spatial data to be extracted according to various standards and published in an OGC compliant registry. This aligns with the OGC Open Services Framework which is based on a publish-find-bind architecture. This creates flexibility and interoperability. Tools developed to both integrate spatial data and metadata and to automate the process of updating metadata would be used widely within the spatial data sector.

Some elements of metadata obviously cannot be automatically updated. These would not be stored in an integrated fashion with the spatial data. Only those metadata elements that can be automatically updated would be integrated with the spatial data. This will save producers of data both time and money associated with the updating of metadata records, and will also aid data users who require up to date metadata to be delivered with data for their use.

Research undertaken within ETH Zurich University in conjunction has examined the possibilities of integrating metadata and spatial data and creating an automated process (Najar et al., 2007). This initial research lays the ground work for the development of appropriate metadata management tools, applications and models which will directly aid the development of the integrated approach for managing and automatic updating spatial metadata.

5. CONCLUSION

Many national and international issues concerned with land management, environmental sustainability, water and disaster management can be addressed by having the ability to find and access high quality spatial data within SDIs. The ability to find and access the appropriate information relies on having up-to-date metadata. However, current metadata models, application and standards are complex and very difficult to handle, often with missing or incomplete metadata. It is also viewed as an overhead and extra cost by organisations.

The key criteria for supporting flexible metadata applications are those of technical requirement, compliance with international standards, user friendly interface and availability of necessary functions for handling metadata records. Ensuring proper development technology together with efficient database technology to support access and maintain metadata records have been identified as critical factors within the technical criteria. Meanwhile, robust coded open source METs will be more efficient as the metadata standards evolve. Consequently, a metadata entry tool must be sufficiently flexible to address the changes to metadata standards over time and to allow users to create and extend a standard to satisfy organisational and local needs. Finally, the

MET designers should focus greatly on creating designs that satisfy the user requirements in terms of functionality and usability. A tool primary function should be more than handling the different standards, rather making maintenance of the metadata records easy for the user.

More importantly, the ability to automatically generate metadata relating to spatial data, and make it available through SDI will have important benefits all practitioners including spatial data producers, vendors, distributor and user. One of the easiest ways in which to investigate and search for spatial data is through integrated metadata directories. If users can find data and services, then they will be utilised, growing the spatial data market. This highlights the importance of the integrated tools to businesses and agencies that produce spatial data. Data producers can cut down on the amount of time and money spent producing metadata, while at the same time increasing the ability for customers to find and hence use their data and services. This will enable metadata to be maintained dynamically in a way that addresses the more real-time requirements of people and organisations that use spatial data.

ACKNOWLEDGEMENTS

REFERENCES

Ballari, D., Maganto, A., Nogueras, J., Pascual, A., Bernabé,M. (2006). "Experiences in the Use of an ISO19115 Profile within the Framework of the Spanish SDI", *GSDI-9 Conference Proceedings*, 6-10 November 2006, Santiago, Chile.

Crompvoets J., Bregt A., Rajabifard A., Williamson I. (2004). Assessing the worldwide developments of national spatial data clearinghouses*, International Journal of Geographical Information Science*, 18(7): 665-689.

Geonetwork Community (2007). *Geonetwork Opensource Complete Manual,* at: http://geonetwork-opensource.org/documentation/manual/geonetwork-manual/Manual.pdf.

ISO/TC211 International Standards Organisation ISO (2001). Geographic information – Metadata, ISO 19115.

MetaD (2007). *User manual of the application MetaD for the creation, edition and exportation of geographical metadata,* at: http://www.geoportal-idec.net/geoportal/eng/inici.jsp?pag=metad&home=s.

Moellering, H., Aalders H.J., Crane A. (2005). *World Spatial Metadata Standards*, Elsevier.

Najar, C., Rajabifard, A., Williamson, I., and Giger, C. (2007). "A framework for comparing spatial data infrastructures: an Australian-Swiss case study", in H.J. Onsrud (ed.). *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, Redlands, California, USA: ESRI Press, pp. 201-213, at: http://gsdidocs.org/gsdiconf/GSDI-9/papers/TS24.1paper.pdf.

Phillips, A. and Williamson, I. P. and Ezigbalike, D. (1998). "The importance of meta-data engines in spatial data infrastructures", *Proceedings of AURISA '98*, Perth, Australia.

Rajabifard, A., Binss A., Williamson I. (2005). "Development of a Virtual Australia Utilizing an SDI Enabled Platform", *Proceedings of FIG Working Week/GSDI-8*, Cairo, Egypt.

Shekhar S. and Chawla S. (2003). *Spatial Databases*: *A Tour*, Prentice Hall.

Waugh A. (1998). Specifying metadata standards for metadata tool configuration, *Computer Networks and ISDN Systems*, 30(1-7): 23-32.

Williamson, I., A. Rajabifard and M.-E.F. Feeney (Eds.) (2003). *Developing Spatial Data Infrastructures: from Concept to Reality*, London, UK: Taylor & Francis.

A Prototype Metadata Tool for Land Use Change and Impact Models – a Case Study in Regional Victoria, Australia

Stephen Williams, Christopher Pettit, David Hunter and Don Cherry

Department Primary industries, Victoria, Australia
{steve.williams, christopher.pettit, david.hunter, don.cherry}@dpi.vic.gov.au

Abstract

The use of models to infer or predict changes and impacts in natural resources and environmental systems is a fundamental research activity around the world. A recent audit of such modelling activities in eastern Australia uncovered a plethora of models in use and a number of instances where models were implemented across various groups and agencies. Often the active parties were unaware of each others research. The preparation of data and development of model parameters to support deployment of a model can take considerable effort and this can often be leveraged by subsequent research. Additionally, previous modelling when accessible may reduce expenses and inform by lessons of experience the selection of models and approaches to their future implementation. Addressing these research needs is the subject of this article. A prototype tool for storing and managing model metadata has been developed that extends the utility of the more traditional model register allowing storage of details associated with each instance of a model run. A non-standard approach has been taken to enable efficient registration of the spatial context for model runs. The overall approach taken has implications for the development of Spatial Data Infrastructures (SDI), model automation and e-science.

Keywords: Metadata, spatial models, e-science, natural resource management.

1. INTRODUCTION

Since the advent of the computer in the 1940s there has been considerable research into the development and application of spatial models for better understanding landscape process, function and futures. In fact a Google[TM] search on the term 'landscape model' resulted in 811,000 hits, and although some of these hits are extraneous, most exemplify the proliferation of modelling endeavours. With such large numbers of models developed and applied at a range of scales from local to global it should be possible to search where such models have been applied and when, what datasets are required to run a particular model, and who are the custodians and experts associated with such a model. Much research on developing a spatial data infrastructure (SDI) has addressed such issues for datasets. However, little research has been done with respect to models and model outputs. This article describes such a prototype SDI interface developed, known as the Model Information Knowledge Environment (MIKE).

MIKE began as a pencil sketch outlining a desired flow diagram showing how a client's query might lead through to data and model discovery. The early concept of MIKE focused on addressing land management questions that could be informed by a spatial modelling tool applied within the context of a landscape. A common global agricultural goal shared by primary industries sector in Victoria focuses on the need for productive and sustainable landscapes. A better understanding of landscape health and ecosystem services in relation to potential agricultural industries can be acquired through the application and development of a growing number of spatial modelling tools. Such models can be used to assess and inform understanding of the incremental and cumu-

lative effects of activities on the landscape. Spatially explicit models which rely on such common parameters as magnitude, frequency and extent are used extensively to simulate potential change, highlight patterns and identify critical impacts such as habitat fragmentation, dry land salinity and soil erosion within agricultural areas.

Models can help in understanding the impacts of changes in climate, highlighting at-risk portions of the natural environment and immediate threats to agriculture. For example, the current dry period (below average rainfall over the last 12 years across south eastern Australia) places significant pressure on limited water resources. The application of water balance models in specific catchments is recognized as crucial in protecting those resources (Boughton, 2005; Ranatunga et al., 2008). Furthermore the impact on the environment is felt when wildfires race through communities and forests alike due to the prolonged drying effect on the land. Review studies highlight the need to develop search and discovery systems to locate each model instance, the data used and its spatial/temporal extents.

The MIKE prototype has been populated with a number of land use change and impact models as reported by (Nichol et al., 2005). Many of these models have been applied in Victoria to better understand: adaptive management of native vegetation, rural land use change, groundwater dependencies and socio-economic conditions. These research methodologies and results are only accessible via a recent published volume on spatial models for natural resource management and planning (Pettit et al., 2008). This typically is how models and model outputs are shared within the modelling community. The outputs from a number of land use change and impact model also make their way into Government reports which are accessible via planners and decision-makers. This raises all kinds of issues around discoverability, reuse of models and model outputs. Thus, the fundamental question our research endeavours to address is: how can the plethora of models for understanding land use change and impact be made more accessible to end users? We believe the key lies in model metadata, a term which refers to 'data about models'. This article concludes by outlining on-going research in designing a methodology for evaluating the MIKE prototype and identifying future research priorities.

2. METADATA A CRITICAL COMPONENT TO SDI

Spatial Data Infrastructures are platforms that facilitate a wide variety of users to access data in an easy and secure manner, assist stakeholders to cooperate and enable interaction with spatial technologies in more cost effective ways (Rajabifard 2002). SDIs can be characterized by their sphere or scale of influence. Rajabifard (2002) describes 5 levels of SDI hierarchy based on local, state, national, regional and global scales. In this context the MIKE aligns more strongly in respective order at the state, local and national scales. The use of metadata is a critical component within an SDI. While the application of a variety of standards provides commonality underpinning a reliable SDI it is the metadata content within the system that delivers the 'contextual intelligence' required to support the diversity of data and applications utilizing the infrastructure. There are a number of research initiatives both nationally in Australia such as the Dataset Acquisition Accessibility and Annotation e-Research Technologies (DART) (see Website 1) and Australian Research Repositories Online to the World (ARCHER) (see Website 2) projects and the international Infrastructure for Spatial Information in Europe (INSPIRE, 2003) that are investigating development of SDI's and their inherent metadata components. Additionally projects such as the Science Collaboration Environment for New Zealand Grid (SCENZ-GRID, 2008) are beginning to combine spatial

data and grid computing infrastructures with associated metadata to deliver on-line spatial data processing and modelling.

Figure 1 organizes metadata into several simple functional classes and assists in illustrating the conceptual relationships between spatial data, spatial data metadata and the different forms of model metadata.

Figure 1: Functional classes of metadata.
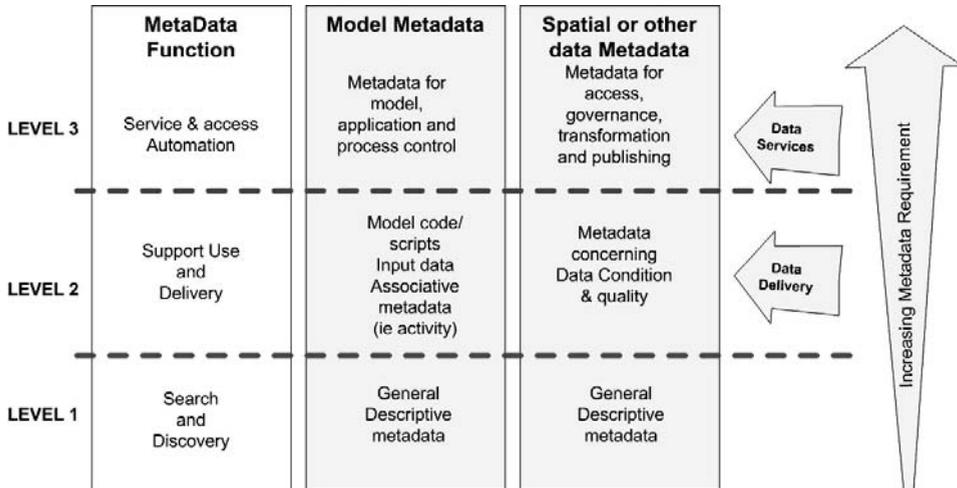
This approach to describing metadata is slightly unorthodox because it extends the common paradigm that uses metadata largely for purposes of search and discovery. The term metadata, purportedly first used in 1969 (Howe, 1996) is defined as "structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage an information resource" (NISO, 2004). A more commonly used definition is that metadata is "data about data". Here we use 'model metadata' as a term or descriptor meaning 'data about models'. This is not to be confused with the terms meta-modelling and meta-models as these describe a contemporary approach to model selection, construction and assembly (Keller and Dungen, 1999; Bridewell et al., 2006).

The NISO (2004) metadata report states "Metadata is key to ensuring that resources will survive and continue to be accessible into the future". In Agriculture and Natural Resource Management (NRM) this is undoubtedly true for metadata about data and information resources as these are often primary sources of knowledge describing the state and nature of environmental systems and are important for comparative analysis regardless of their age. However, this is probably less so for model metadata as models represent the interpolative, inferential or processing systems that are used in research. These methodologies are generally subject to continual improvement and consequently it could be argued that model metadata has a greater potential to age and is most useful when it is more current. Although there are metadata standards for data and information (ISO 19115) similar standards are only beginning to emerge for model metadata.

The functional levels shown in Figure 1 are closely related to three main types of metadata: descriptive metadata, structural metadata and administrative metadata (NISO 2004). At level 1 the metadata is primarily used as a resource for discovery and identi-

fication both of data and models. The usage of spatial data metadata for this purpose is a fundamental service within SDI and is in wide use (van der Wel, 2005; Stout et al., 2007; Nogueras-Iso et al., 2005). At this point there are many model registers supporting model search and discovery (i.e., NASA Global Change Master Directory (Website 3), Freebase - Environmental Modelling Collection (Website 4), Catchment Modelling toolkit (Website 5)).

At level 2 a deeper description of both information and models is required to assist their utilisation. For data and information in particular this includes metadata describing the genesis, quality and condition of the data resource. In the fields of Agriculture and NRM this metadata is of high importance as data often exists as a collection of data instances that have been produced by a number of projects and initiatives. These are often spread over time and may be subject to an evolving standard or improving collection methods and associated technologies. Unless there has been strong adherence to a data collection standard the specific instances in the collection may differ in subtle ways that can have a profound influence on their utility. At this level these differences are captured in the metadata which is often of a more technical nature. At present this type of metadata does not get much attention in metadata schemas (NIOS, 2004) and can be costly to collect. An exception is the emerging Numerical Model Metadata standard associated with climate modelling (Steenman-Clark et al., 2004). Stout et al. (2007) provide an example where the US Federal Geographic Data Committee (FGDC) metadata standard is being extended to support some of this type of metadata. At level 2 the use paradigm for model metadata is changed. Rather than generally describing the model itself emphasis is placed on storing contextual information about the modelling activity and how to implement. To support effective model use not only are technical descriptions of the inputs and description of how to execute the model important but understanding where and who implemented the model is extremely useful. This can benefit to prevent duplication in modelling effort. In contrast to level 1 there are not many tools that effectively address and use model metadata at this level. Most approaches to date are associated with theme based modelling groups and a good example is the 'Earth System Curator' and its prototype database used by climate modellers. In this context the record of model activity is confined to the community of interest. This prototype has the stated objective to "store metadata related to model runs and datasets" but is still under development. The authors have not yet found published model metadata systems recording details of model instances spatially (other than the MIKE system described herein). Practical support for this function requires user-friendly and efficient approaches to the spatial registration of these instances. At level 2 there are opportunities to link model and data metadata repositories (for those models with stable or fixed data inputs) and support queries informing understanding of spatial data availability for modelling. These approaches require a common referencing scheme for datasets in both registers. This is a key component within the case study described in the following section.

At the third level the emphasis for metadata shifts to record details that can support associated processes both for models and data. In the data area examples include metadata to support publication and transformation of data such as in web based mapping or even workflows associated with data such as managing collection by third parties (Stout et al., 2007). In relation to models this metadata serves as a base to underpin applications for automated model processing or pre-processing of data for models. Current development in this area is largely confined to process models in remotely sensed data (Jianto et al., 2003) or in climate modelling. The development of metadata in this area is a crucial part of an SDI to support automated publication of data and models. Where the metadata is extended to store processing instructions, algorithms or

model elements it becomes possible to utilise these in scientific services or workflows. This is one of the key elements to e-science. An example where these concepts are being applied is on the SCENZ-GRID web-site (see Website 6).

3. MODEL METADATA

The form and detail of model metadata is by its nature inextricably linked and influenced by the models themselves and their application context and deployment environment. Models are infinitely diverse. In designing a framework to store, manage and support the use of model metadata it is important that this diversity is accommodated. Equally the scale or size of the design needs to support the required metadata but also ensure unnecessary detail that will confound implementation and use is minimized (Gangsheng et al., 2008). Underpinning this approach is the premise that model metadata needs to serve a purpose and it is precisely for this reason that we have adopted the functionally based view for metadata shown in Figure 1.

Models can be grouped in a number of ways including classing by algorithm type and architecture (Jørgensen, 2008), by subject or application area (NLWA, 2004; NASA GCMD, 2009), by how they fit into a scientific research framework (Steinitz, 1990; Nichol, 2006) or simply by size and complexity. Irrespective of the model characteristics used to establish classes it is common for these classes to require different types of model metadata. As an example, Jørgensen (2008) proposes a scheme containing eleven types of model including one type called 'spatial models'. If spatial model instances are to be re-used then the precise spatial boundary associated with an instance can be a very useful element of model metadata. In contrast the same metadata element will have little relevance or a different contextual meaning for a market or sector based economic model. Classifying models by how they are currently deployed provides a practical means to assist in gaining a better understanding of some of the major differences in model metadata requirement. This can be done simply by grouping models into firstly those models (and modelling assemblies) that essentially operate in a stand-alone manner, secondly those that operate within the context of a modelling environment (i.e., ISEE systems Stella®, ESRI™ ModelBuilder, Kepler) and finally those that operate as a service (such as a web service). Examination of recent reviews on natural resource and landscape impact models in Australia (NLWA, 2004; Nichol, 2006) indicates that to date almost all models in use lie in the first group with some in the second and very few in the third. As many of these models require significant amounts of data, effort and expertise there will be demand from new users for model metadata describing previous instances of model implementation. In particular they will want referral information (to contact those involved for knowledge transfer) and contextual information about the modelling instance to support an assessment of it's potential usefulness. By implication the greatest demand for model metadata in natural resource management in the immediate future is likely to be at level 1 and 2 (see Figure 1) and oriented towards finding suitable models and assisting in their local deployment and use. This situation is unlikely to change rapidly unless there is significant effort and investment made to re-architect or replace existing models with service oriented alternatives. For this reason the diversity in models and model metadata requirement is likely to persist and continue to pose challenges for research into model metadata development, management and associated standards, systems and applications. The advancements in conceptualization of model interaction and the kinds of model metadata needed to support interoperability (Nilsson et al., 2006) represent steps towards a more organized, standardized and desirable future. However, as these approaches are more aligned to a service or web based modelling paradigm this is at odds with current reality. Correspondingly, partly in order to gain acceptance, the major focus of the case

study described in the next section is on model metadata pertinent to finding and using models.

4. CASE STUDY

4.1 Overview of the study and prototype tool

The overall goal of the case study is to support researchers in understanding the availability, use and application of models and landscape analysis tools by providing pertinent and current information through a spatial web environment (metadata tool). The key questions that the tool was designed to address include:

– What models are in use and what are they, what do they do?
– Who is using them and where?
    a) What data do these models require, and
    b) Is it available?

Some of the key functions that the tool was designed to support include:

– Search and query of model and data metadata;
– Registration of models;
– Registration of modelling instances or activity, and
– Display and visualisation of metadata.

The case study area comprises the State of Victoria, Australia. Victoria is the smallest mainland state in Australia with an area comprising approximately 228,000 square kilometres and a population of just over five and a quarter million people. It has significant mineral and coal deposits and is a major producer of food and fibre. Natural resource management challenges include food and resource security associated with land degradation, salinity, climate change bio-security and changing land use. In recent times water shortage has become a major issue at an unprecedented level across the state. A large number of models are being applied to assist in understanding opportunities and responses to these and other issues. However, the number of models examined for trial in the MIKE was constrained to ensure the work remained manageable. The key functional elements to the MIKE are outlined in Figure 2. The greyed out sections indicate areas for future development that were not the focus of development at this time. The storage of both model and data metadata is via a simple 4 tier hierarchical data model that can readily be converted to XML. Where possible the ISO 19115 and ANZLIC 2 geospatial data standards were applied; especially in respect to naming and code lists. The metadata content stored includes data associated with models (including versions), model instances, model features and elements associated with model features (see Table 1). The data repository is a Microsoft SQL 2005 database and an ESRI product suite, ArcGIS, ArcIMS and Spatial Data Engine (SDE) was used to the spatially enable functions within the tool. A web front end is provided and supported.

4.2 Facilitation of spatial referencing

The spatial registration of modelling projects is essential to enable collaborators to understand who is using what models and where. To enable users of the prototype system to do this easily and efficiently the spatial boundaries were determined for a sub-sample of known modelling projects (i.e. Catchment Analysis Toolkit, SLEUTH, MODFLOW, BC2C) identified in the study by Nichol et al. (2006). The results indicated that

Figure 2: Overview of the system components.
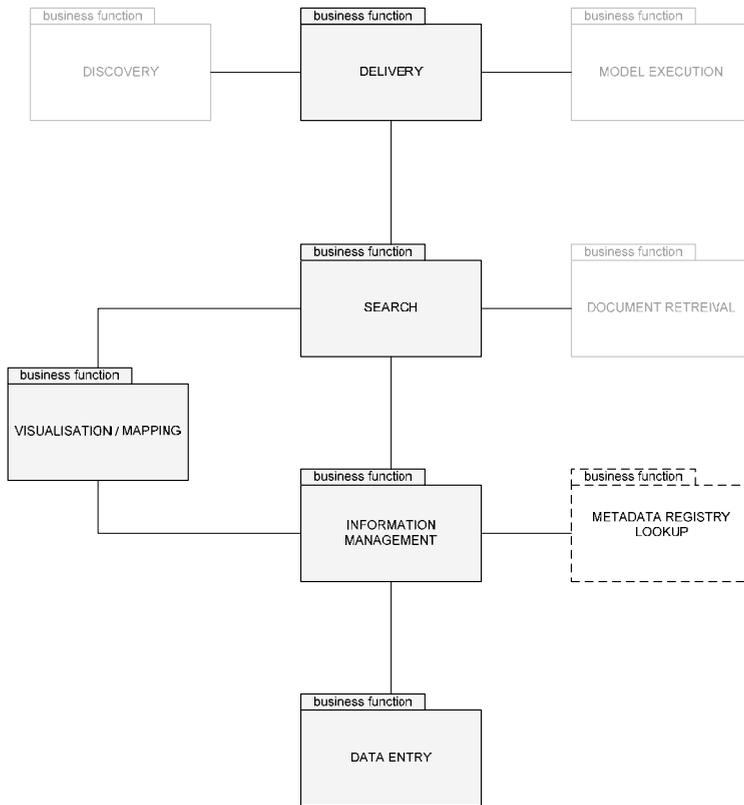
a combination of administrative (local government areas, parishes, catchment management regions and irrigation regions) and physiographic boundaries (drainage basins) held the spatial boundary definition required to register almost all natural resource modelling projects. The major exceptions to this rule in Victoria, Australia being fine scale modelling in small sub-catchments or studies associated with specific biodiversity remnants and reserves. With respect to the former there is currently no source of accepted or formalised boundaries for sub-catchments within the state of Victoria. As the departmental focus is Agriculture (not biodiversity) for the purpose of the case study it was deemed that the combination and intersection of all the identified boundary datasets (excepting those for sub-catchment and biodiversity remnants) would provide an adequate index to spatially reference modelling projects. This produced an index with 2881 polygonal elements. Each element retained all the attributing from all the contributing parent boundary datasets. The resultant index provides a spatial matrix enabling registration of modelling projects by any of the contributing spatial boundary layers.

4.3 Interface design

The interface is constructed to provide both spatially enabled views and the more classic subject or theme based model discovery services associated with a model register. Seven tabs are provided at the top of the home page and persist within most of the tools windows. Apart from the home tab these represent themes or groupings of func-

tionality. Themes are 'map view' (spatial interrogation and display of models and data), 'model query' (model centric query), 'model discovery' (model search and selection based on type, application or other model feature of interest), 'data discovery' (aspatial data search), 'administration' (includes metadata entry and management and tool administration and management functions) and 'support' (provides contact information, manuals and links to other sources of relevant information). At present the system largely stores metadata at level 1 and 2 (see Figure 1 and Table 1) and further research into the design will be required to support automated modelling or processing applications (although the storage of some relevant metadata is currently possible).

Table 1: Some of model feature descriptors used in the MIKE.

Model feature metadata	
Model Feature description	**Model Feature description**
A classification (or type) of the model	Model Physical Domain
A classification of model function based on Steintiz Framework	The level of model integration with a GIS
A key contact for the model includes full details such as address and phone	The model computational basis
A limitation of model	The model development stage
A model called by this model but not fully integrated	The principle application of the model
A model input	The spatial scale of the model
A model integrated within this model	The temporal characteristics of the model
A model output	The version details of the model
A program language that the model is written in	Abstract describing the model
A web page related to the model	The background or history to the model
An author of the model	The key model objectives
An operating system that supports the model	Description of the model packages or component programs
Keywords associated with the model	Description of the modelling procedure
Model access type	Description of the model purpose
Model application type	

In the main the default interface design standards associated with the component technologies and tools have been adopted and only varied as necessity demanded.

The registration mechanism for model projects or instances is shown in Figure 5. Note that these instances at this point are not equivalent to the concept of model runs supported by the Earth System Curator (DeLuca, 2008). Within that and similar systems a model run is characterised by the execution of a model with a specific set of parameters at a date and time. Due to the complexity and dispersion (geographic and institutional) of NRM modelling our use of the term model instance is equivalent to the application of a model to a specific landscape at a specific time (usually called a 'modelling project'). Consequently a model instance in this context can represent a large number of modelling runs. To record a model instance in the tool requires:

1. selection of a model;
2. provision of a contact name for the instance;
3. selection of spatial area type (i.e., river basin);
4. nomination of the spatial member of this type (i.e., Loddon river);
5. entry of a host project title (not required to be pre-registered), and
6. the option to enter additional notes to support or describe the instance.

Figure 3: Mapping interface showing a model instance (shaded area).
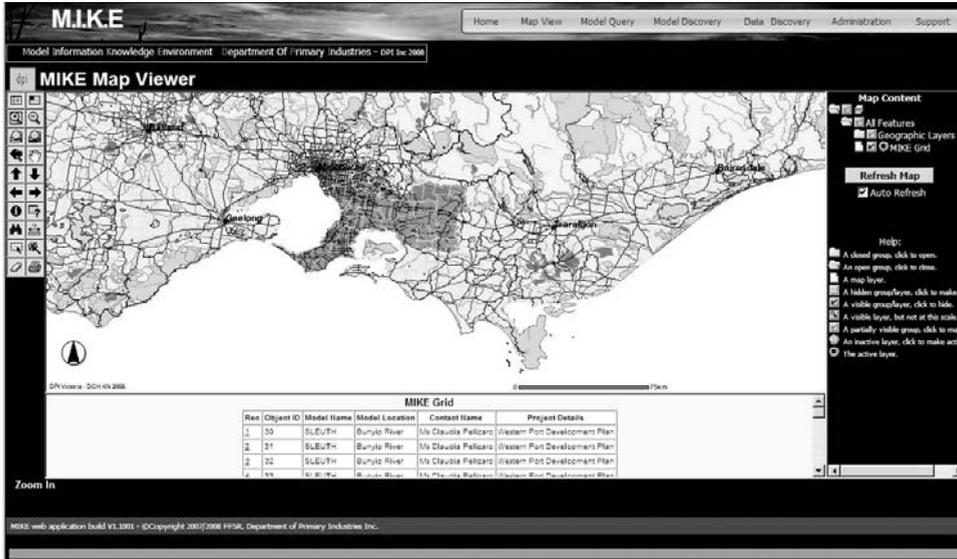
Figure 4: Close up of a model instance.
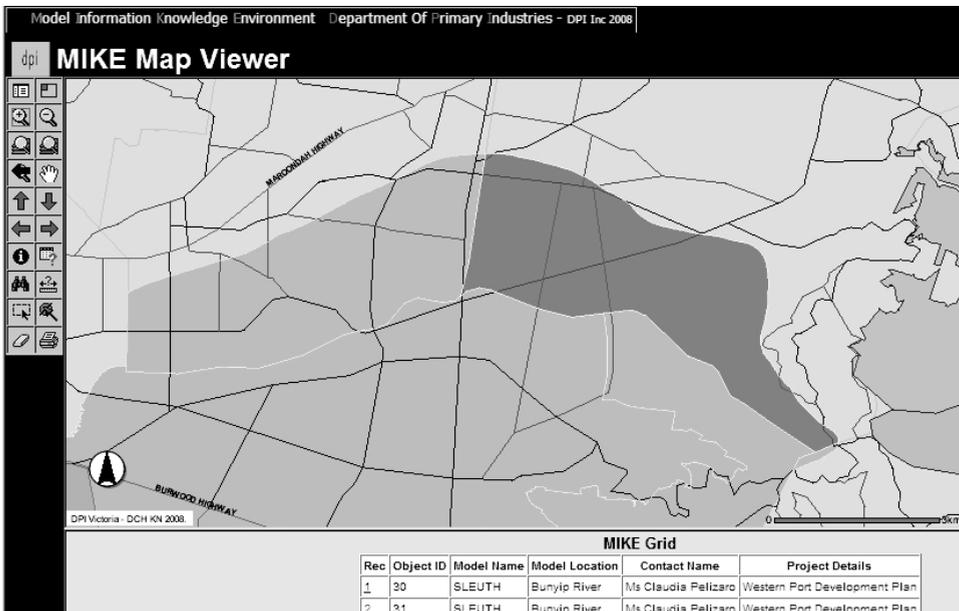
Figure 5: Registration screen for a model instance for the SLEUTH land use change model.

Figure 6: Administrative functions within the prototype.

Typically we have found this process is completed by users within 60 seconds. It is essential the process of registering a modelling project is streamlined and efficient in order that this does not become a barrier to tool uptake. Additional information such as date of the instance registration is generated invisibly to the user. In the future when

the tool is deployed more widely users will need individual registration and login before they can add model instances; in this case the instance will then inherit their personal details automatically. This is not currently the case. Figure 4 shows more detail of a model instance.

The administrative elements contained within the system are shown in Figure 6. Several role based tiers of security are currently supported and this controls access to specific functions. Additionally the system supports some elements of model governance (although this requires further development) allowing model developers to record and communicate endorsement levels and/or development stages for registered models.

5. EVALUATION

Following the initial discussions to develop the MIKE system the prototype was built. To date the evaluation of the product has been limited to one workshop and a number of one-to-one interviews. This process has revealed a range of areas for further development and also highlighted a number of issues. For instance one significant requirement is to link the model metadata to a data register to identify gaps in existing data so that this can inform model choice and understanding of the resources required for application of the model. However, not all models have fixed data inputs and consequently this some-what confounds and complicates design of the interface. Another issue is that some models contain component models (models within models). At this point although MIKE allows storage of parent child relations between models this is currently a simple implementation and needs improvement.

A simple analysis of the strengths and weaknesses of the prototype based on feedback received to date indicates the following. In terms of ease of use and intuitiveness, parts of the system are easy to use especially the component supporting registration of model instances. In particular the implementation allowing the rapid spatial registration and query of an instance is somewhat unique. The system supports storage of metadata for a model, its versions and the instances of application. In concept this is similar to the three level hierarchy described in Gangsheng et al. (2008). By volume, the bulk of the model metadata currently in the system describes the base models. Unlike model instance information this is currently cumbersome and time consuming to enter in the prototype. This is a key weakness and further research and improvement to the design of this component of the interface is required. Additionally, refinements to the design of the underlying data structures to support this will also be required. Systems like NASA GCMD (2009) provide good insight and guidance for this work. At the back end the metadata model currently appears (based on user feedback) to store the bulk of the requisite information to support model discovery and use, but provides little support for the storage of metadata to support on-line model services and automation. This intentional gap and weakness will be addressed in the future and will benefit from rapid advances occurring in metadata design for interoperability. Another major weakness and threat to the prototype is that current spatial data Metadata repositories and services contain deficiencies in the design, implementation and level of content in respect to storage and delivery of metadata about data quality. This generally confounds the ability to meaningfully link model metadata to existing data registration services and understand data gaps and deficiencies in support of the use of models.

Although aspects to support model governance have been included in the design of the prototype we have found that the present interest in these features has not been widespread. We suspect this is due to cultural factors more than system design and implementation. Further evaluation is required.

As the research progresses more feedback will be sought from key stakeholders and additional evaluation workshops will be programmed. The focus will not be on what has been done so far, but to garner ideas and concepts to assist in progressing the development of the tool.

6. CONCLUSIONS AND SUMMARY

The MIKE metadata tool shows promise in assisting the use of Natural Resource Management (NRM) models within Victoria, Australia and in providing details of modelling activities throughout the state. Users have found it easy and efficient as a tool to register their modelling projects. The entry of the model metadata content describing the models is proving more cumbersome and future ways to streamline or improve this activity need to be sought. Although some other groups have demonstrated success in automating storage of model runs and model run metadata (Steenman-Clark, 2004) this is usually in the context of more tightly focused modelling at few locations or more isolated laboratory based modelling and processing systems. In our case the modelling activity is widely dispersed geographically and is undertaken using a multitude of continually evolving models. This makes automated approaches difficult to achieve at this time. With the growing number of digital datasets and models being developed and deployed it is critical that metadata not just around fundamental datasets, but also around models such as land use change and impact models be properly developed. Across the board the development and maturation of standards for models and model metadata is required, similar to those that have been developed and continue to evolve for data and data metadata. This will be a crucial step in the evolution of an e-science framework for supporting knowledge management and collaboration.

Further areas of research will include improving the model metadata schema and collaboration to assist further standardisation of the model metadata content (for improved search and query). Exploration of the potential to more automatically extract and capture model metadata and store parameters and data associated with a model instance is also a next step. However progressing this is a real challenge due to model dispersion, complexity and variety. More immediate next tasks will be to further improve the entry and display of the model metadata. In particular there will be a focus on ways to geographically visualise multiple modelling efforts and temporal views of modelling activity within landscapes.

ACKNOWLEDGEMENTS

REFERENCES

Boughton, W. (2005). Catchment water balance modeling in Australia 1960-2004, *Agriculture and Water management*, 71: 91-116.

Bridewell, W., Sanchez, J., Langley, P. and Billman, D. (2006). An interactive environment for the modeling and discovery of scientific knowledge, *International Journal of Human-Computer Studies,* 64(11): 1099-1114.

Chan, T., Beverly C., Ebert, S., Garnett, N., Lewis, A., Pettit, C., Thankappan, M. and Williams S. (2008). "Platform for Environment Modelling Support: A Gris Cell data Infrastructure for Modellers", in Pettit, C., Cartright, W., Bishop, I., Lowell, K., Pullar, D., Duncan, D. (eds.). *Landscape Analysis and Visualisation: Earth and Environmental Science*. SpringerLink, pp. 97-117.

DeLuca, C., (2008), *Earth System Curator*, at: http://www.earthsystemcurator.org/.

Gangsheng, X., Jianwen, X., Hailong, Y. (2008). "Study on framework of GIS-based model metadata", *Proceedings of Geoinformatics 2008 and Joint Conference on GIS and Built Environment.*

Howe, D. (1996). *Free on-line Dictionary of Computing (FOLDOC)*, at: http://wombat.doc.ic.ac.uk/.

INSPIRE (2003). Report on the feedback of the Internet consultation on a forthcoming EU initiative establishing a framework for the creation of an Infrastructure for Spatial Information in Europe, Inspire.

Jianto, B., Tianhe C., Xintong L., Yanrong C., Hongqiao W., Jianbang H. (2003). Research on Application of the Method-Metadata in Resources & Environment Information System, at: http://www.gisdevelopment.net/policy/international/interna015.htm.

Jørgensen, S. (2008). Overview of the model types available for development of ecological models, *Ecological Modelling,* 215(1-3): 3-9.

Keller, R. and Dungen, J. (1999). Meta-modeling: a knowledge-based approach to facilitating process model construction and re-use, *Ecological Modelling,* 119(2-3): 89-116.

Lithong, S., Yuxia H., Xiaowen L., Guoliang T. (2001). Spatial data and model management based on XML, *Geo-science and Remote Sensing Symposium 2001,* 4(2001): 1951-1952.

NASA GCMD (2009). *NASA Global Change Master Directory*, at: http://gcmd.nasa.gov/.

Nichol, G. (Ed.) (2006). *Landscape Analysis Models and Frameworks – a review*, Department of Primary Industries, Victoria, Australia.

Nichol, G. E., R. J. MacEwan, C. Pettit, J. Dorrough, H. Hossain, H. Suter, D. Cherry, C. Beverly, X. Cheng, V. Sposito, J. McNeill, A. Melland and A. Shanks (2005). "A Review of Models Applicable to 'Our Rural Landscape'". *International Congress of Modelling and Simulation Advances and Applications for Management and Decision Making (MODSIM05)*, Melbourne, Australia.

Nilsson, M., Johnston, P., Ambjörn, N., Powell, A. (2006). Towards an Interoperability Framework for Metadata Standards, at: http://www.scribd.com/doc/43235/Towards-an-Interoperability-Framework-for-Metadata-Standards

NISO (2004). *Understanding Metadata.* National Information Standards Organization, USA.

NLWA (2004). *Natural Resource Models in the Rangelands – a review undertaken for the National land and Water Resources Audit.* National Land and Water Resources Audit, Australia.

Nogueras-Iso, J., Zarazaga-Soria, F. and Muro-Medrano, P. (2005). *Geographic Information Metadata for Spatial data Infrastructures- resources, interoperability and information retrieval*. Springer.

Pettit, C., W. Cartwright, I. D. Bishop, K. Lowell, D. Pullar and D. Duncan (2008). *Landscape Analysis and Visualisation*. Berlin: Springer-Verlag.

Rajabifard A. (2002). Diffusion of Regional Spatial data Infrastructures: with particular reference to Asia and the Pacific, *PhD thesis,* Department of Geomatics, Faculty of Engineering, University of Melbourne, Australia.

Ranatunga K., Nation E., and Barratt D. (2008). Review of soil water models and their application in Australia, *Environmental Modelling & Software,* 23(9): 1182-1206.

SCENZ – GRID (2008). Science Collaboration Environment for New Zealand – GRID, at: http://scenz-grid.bestgrid.org:8080/.

Steenman-Clark, L. (2004). Numerical Model Metadata Goals, at: http://ncas-cms.nerc.ac.uk/NMM/component/option,com_wrapper/Itemid,72/.

Steenman-Clark, L. and Bouton, K. (2004). Numerical Model Metadata – a draft standard.

Steenman-Clark, L. Numerical Model Metadata Schema, at: http://ncas-cms.nerc.ac.uk/svn/NMM/NMM_Model/trunk/v.011/schema_def/NMMModel.xsd.

Steinitz C. (1990). A framework for theory applicable to the education of landscape design architects (and other environmental design professionals, *Landscape Journal,* 9(2): 136-143.

Stout, D., Kodis, M., Wilen, B. And Dahl, T. (2007). *Wetlands Layer – National Spatial Data Infrastructure: A phased approach to completion and modernization.* US Fish and Wildlife Service.

van der Wel, F. (2005). Spatial data infrastructure for meteorological and climate data, *Meteorological Applications,* 12: 7-8.

Web references

Website 1: Dataset Acquisition Accessibility and Annotation e-Research Technologies (DART): http://dart.edu.au/.

Website 2: Australian Research Repositories Online to the World (ARCHER): http://arrow.monash.edu.au/vital/access/manager/Index.

Website 3: NASA Global Change Master Directory: http://gcmd.nasa.gov/.

Website 4: Freebase - Environmental Modelling Collection: http://www.freebase.com/view/user/robertm/environmental_modelling.

Website 5: Catchment Modelling toolkit: http://www.toolkit.net.au/cgi-bin/WebObjects/toolkit.

Website 6: SCENZ-GRID: http://scenz-grid.bestgrid.org:8080/.

Implementation of Recent Metadata Directives and Guidelines in Public Administration: the Experience of Sardinia Region

Luisa Manigas, Michele Beneventi, Luca Corvetto, Rita Vinelli and Marco Melis

Region of Sardinia (Italy), Local Authority for Municipalities, Finance and City Planning
Department for Cartography and Information Systems
{lmanigas, lcorvetto, rvinelli, melisma}@regione.sardegna.it
mbeneventi@sardegnait.it

Abstract

Since 2006, the Region of Sardinia (Italy) has been developing its Geographic Information System and the related Spatial Data Infrastructure, known as *Sistema Informativo Territoriale Regionale - Infrastruttura dei Dati Territorial* (SITR-IDT). The aim of SITR-IDT is creating and managing a spatial database and the technologic infrastructure, services and application needed for data access. This is in agreement with the principles of geographic data sharing and accessibility expressed by the INSPIRE Directive; for this purpose, metadata play a fundamental role. In SITR-IDT a high importance has been given to metadata definition, acquisition and management. Proper tools for metadata management, according to INSPIRE Implementing Rules (IR), are being developed. Both INSPIRE IR and the requirements coming from an appropriate daily spatial database administration have been considered in metadata organisation. In this article, the metadata organisation adopted in SITR-IDT for data of different typologies is illustrated. It presents the evolution of metadata organisation according to European Union Directives and technical guidelines. It also explains the switch from a quite rigid and constraining metadata schema to a more flexible and standards-compliant one. Different questions such as the metadata manager and the organisation of metadata for non-geographic data related to geographic data are discussed. The article presents a tool, the metadata manager, that should help to create, collect, and manage metadata at the appropriate levels of a Spatial Data Infrastructure.

Keywords: Geographic Information System (GIS), SITR-IDT, Sardinia, metadata, INSPIRE, ISO19115, metadata manager, non-geographic data.

1. INTRODUCTION

The most important target of SITR-IDT is to create a spatial database containing all geographic data produced by the Regional institutional agencies in Sardinia, Italy. The data quality should be granted by official validation and certification. Metadata play a fundamental role in arriving at this target.

Metadata are the data 'passport'; they describe data features and therefore represent the data identity. When structured according to international standards, metadata allow data consultation, access and use by international stakeholders. In the definition of the metadata organisation initially chosen in SITR-IDT, the 'IntesaGIS' project played a fundamental role. This project, started in 1996 and nowadays promoted by the CNIPA (*Centro Nazionale per l'Informatica nella Pubblica Amministrazione*), aimed at defining and sharing operational guidelines for geographic data creation among the public administrations of Italian Regions. The main goal of this project was to make the implementation of spatial databases standardised at a national level (Intesa GIS, 2004a; Intesa GIS, 2004b).

Over the last few years, important in-depth analyses concerning geographic metadata were carried out, and several technical guidelines were delivered. Internationally, ISO19115 (ISO, 2003) and ISO19119 (ISO, 2005) standards, INSPIRE Implementing Rules (IR) (Drafting Team Metadata, 2007) and the European Commission Regulation 1205/2008 have been published. In Italy, operational guidelines have been delivered by the Italian national authority for geographic data (CNIPA, *Centro Nazionale per l'Informatica nella Pubblica Amministrazione*).

In 2006, CNIPA has delivered two drafts of technical guidelines (CNIPA, 2006a; CNIPA, 2006b) for creating and updating a National Register of Spatial Data (*Repertorio Nazionale dei Dati Territoriali*, RNDT). The RNDT should collect metadata of all geographic data produced by the Italian public administrations, in order to create a national catalogue. The drafts explained the metadata organisation that Public Administrations should adopt to feed the RNDT. They contained the mandatory plus some optional fields of the EN ISO19115:2005 core set of metadata. Therefore, in SITR-IDT all metadata were structured according to the ISO19115 standard.

In October 2007, the INSPIRE Draft IRs for Metadata, Version 3, were published. With their implementation in December 2008, CNIPA delivered the draft Decree of the President of the Council of Ministers 'Regulation governing the definition of the content of the National Register of Spatial Data, as well as the modalities of establishment and update of the latter'. This draft descended almost completely from the 2006 drafts. The only change was that this final version fully implemented the INSPIRE metadata IR, giving further restrictions concerning multitude and obligation and adding some further fields.

First, this article explains the geographic data organisation in IntesaGIS. Then the metadata organisation in SITR-IDT and the requirements of regional authorities for metadata are provided and a tool to manage metadata proposed. This metadata manager should help to create, collect, and manage metadata at the appropriate levels of a Spatial Data Infrastructure, including the regional level.

2. INTESAGIS GEOGRAPHIC DATA ORGANISATION

'IntesaGIS' laid down a hierarchical organisation of geographic data; according to this schema, data were classified in a pyramidal structure constituted by *Layers*, *Themes* and *Classes*. In particular, *Classes* coincided with the most atomic data level (usually corresponding to the shapefile), while *Themes* were aggregations of *Classes,* and *Layers* were aggregations of *Themes*. Initially, in SITR-IDT metadata organisation followed strictly the hierarchical data organisation of 'IntesaGIS'. The individuation of Series, Dataset and Tiles according to the schema reported in Figure 1 descended almost naturally.

However, it must be clear that the 'IntesaGIS' guidelines concerned geographic *data* structure, while no rules concerning *metadata* organisation were given.

3. CHANGES IN SITR-IDT METADATA ORGANISATION

Some changes are going to be done in the near future in SITR-IDT metadata organisation, in order to fulfil the following requirements:

a) possibility of organising metadata out of the strictly hierarchical schema of 'IntesaGIS';

b) definition of a metadata organisation suitable with the requirements coming from the daily database management, and

c) compliance with CNIPA guidelines.

These questions will be better explained in subsection 3.2 (points a and b), and subsection 3.3 (point c).

Figure 1: Correspondence between data structure and metadata hierarchy in SITR-IDT, according to 'IntesaGIS' data organisation.
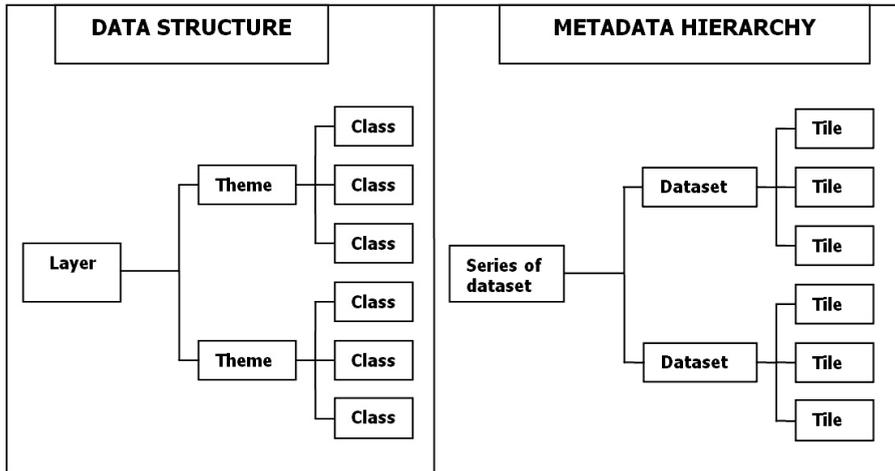
3.1 Current metadata organisation in SITR-IDT

In SITR-IDT, metadata of several cartographic products were successfully organised according to the hierarchical structure of 'IntesaGIS'. For instance, metadata of Sardinian Regional Technical Map were organised with the following criteria:

– the whole Regional Technical Map (as a cartographic product) constituted a *Series*;
– the single sections of the cartographic product, produced in different periods or covering different areas, coincided with *Datasets*, and
– further divisions of the sections into geographic areas where catalogued as *tiles.*

This metadata schema, following the concept of 'dataset and series' laid down by ISO19115, avoids the repetition of several metadata, which, due to the data typology, occurred several times in the whole cartographic product. For instance, different production sections of the Regional Technical Map (each of them constituting a Dataset) shared the same metadata concerning scope and use of the map, lineage, authors, points of contact, and similar information. Therefore, joining these datasets in a series avoided the repetition of those metadata shared by all of them.

Following the same structure, metadata of the Sardinian Regional Topographic Database were effectively organised. This spatial database, which contained the data of the Regional Technical Map organised in an RDBMS, was built according to the 'IntesaGIS' data specifications. For this product, the hierarchical metadata schema was easily identifiable and fitted very well with the data model, which was the following (see Figure 2):

- the Regional Topographic Database was considered as the whole cartographic product;
- the whole spatial database was divided into several *Layers* (such as Orography, Viability, and Administrative boundaries);
- for each *Layer*, several subsets, called *Themes*, were individuated; for instance, the *Layer* of Viability was divided into several *Themes*, each of them containing the data of streets, squares, bridges and similar objects;
- the smallest components of the cartographic product were the shapefiles which constituted each *Theme*; each shapefile was identified as a *Class*, i.e. the most atomic entity identifiable in the database.

According to this data structure, metadata were organised following this schema:

- each *Layer* constituted a Series;
- each *Theme* coincided with a Dataset;
- each *Class* was catalogued as a Tile;
- the whole cartographic product was considered as a series of series, constituted by all the *Layers*.

Figure 2: Data and metadata structures for the Regional Topographic Database.
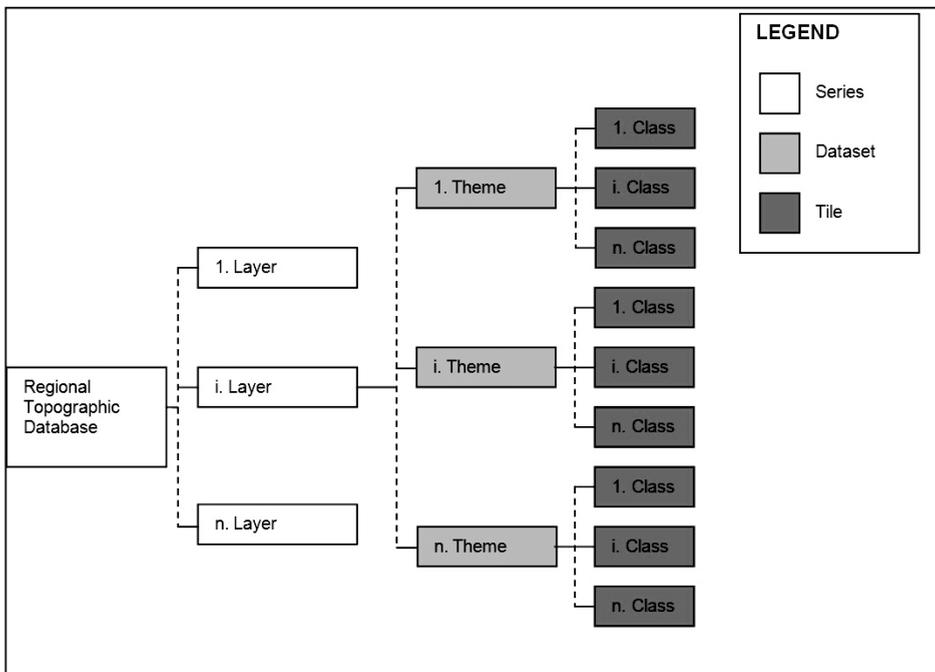
The metadata organisation previously described showed to be effective for the categories of data quoted before, particularly for the Regional Technical Map and the Regional Topographic Database. However, different typologies of data, not organised in a hierarchical structure, can require a different metadata organisation, as explained in the next subsection.

3.2 Definition of a metadata schema suitable with the requirements of SITR-IDT management

If cataloguing only general metadata is the most proper choice for a register at national level, it is not suitable for cataloguing geographic data at regional level. In creating and managing a regional spatial database, data of very variegate typologies can occur, requiring specific metadata for their appropriate and complete description. Managing these metadata can be even more complex than what is required by standards.

In SITR-IDT, several non strictly geographic data are catalogued. For instance statistical, demographic, sanitary or environmental data, being geo-referenced, are part of the SITR-IDT spatial database. In comparison with strictly geographic data, different metadata are required to describe the alphanumeric information contained in these data. Also the detail level of information to be described can be different: for instance, specific metadata for non-geographic attributes, which have different lineages, are needed. In this case, a simple 'series-dataset-tile' schema is not sufficient; a higher detail is required.

For example, consider the metadata organisation adopted for cataloguing a set of statistical data concerning local job areas. The data on the local job areas consists of a shapefile, made both by the geographic data and by the attributes containing the alphanumerical information. The shapefile of all the job areas on the regional extent is catalogued as a dataset, while each single area is catalogued as a tile. Since each shapefile attribute is created by different authors, in different periods and with different lineages, each single attribute of the shapefile needs its own metadata. In this case, cataloguing metadata as series, dataset or tile is not sufficient. The value 'attribute', present in the 'MD_ScopeCode' list of the ISO19115 standard would be required to catalogue metadata at the most appropriate level.

If structuring metadata at a different hierarchical level than series, dataset or tile can be intuitive for non-geographic data, it can be necessary also when managing simple geographic data.

In SITR-IDT, this requirement occurred for instance when the Sardinia Regional Geological Map was catalogued. This geographic product was constituted by 40 distribution units, each of them covering a different area of the regional territory. Each distribution unit was constituted by four different shapefile and a *.mdb* file with the map legend. The four shapefile contained in each distribution unit were:

– shapefile of the geological surveys authors;
– shapefile of the geological surveys reviewers;
– shapefile of the polygonal geometries containing the geological information, and
– shapefile of the linear geometries containing the landslides information.

Each shapefile with the geological and landslides geometries was further divided into geographic sections, according to the local official cartography division. The shapefile of the authors and of the revisers contained the alphanumerical information in its attributes. Each attribute differed from the others for the creation period and for the lineage. With this foreword, metadata for the cartographic product of the Sardinia Regional Geological Map were organised in this way:

– the shapefile of authors and reviewers, for each single distribution unit, constituted two separate datasets;

- the attributes of each of these two shapefile should have been catalogued with the 'attribute' value of the ISO19115 'MD_ScopeCode' list. Since, in SITR-IDT it is not possible to chose the value of 'attribute' as hierarchical level, the attributes were catalogued as 'tiles'. Tiles were thus constituted by non-geometrical subsets of geometrical datasets, individuated on the basis of the differences of production periods and of lineage;
- each of the two shapefile containing the geology and landslides geometries was catalogued as a dataset;
- each of the geographic sections that constituted the shapefile was considered as a single tile;
- each distribution unit was considered as a series, and
- the whole Regional Geological Map was considered as a whole cartographic product, constituted by the mosaic of the forty distribution units (so covering the whole regional extent). It was considered as a series itself, made by the joint of all the forty series, as a sort of 'series of series'.

In this way every components of the cartographic product were described by their own metadata. Sardinia Regional Geological Map data and metadata structures are schematised in Figure 3.

Figure 3: Data and metadata structures of the Sardinia Regional Geological Map.
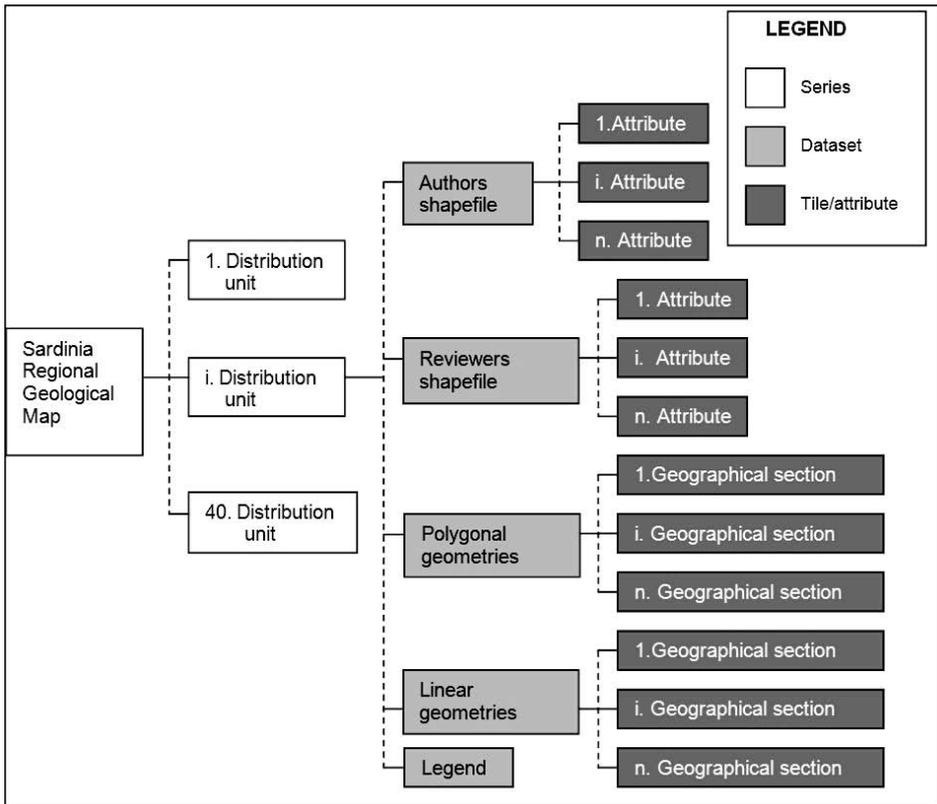
This is just an example of how complex a complete cataloguing of geographic data can be. Many other data present in SITR-IDT required as complex metadata structure. As above mentioned, some data require that metadata are classified into attributes, features or the other values of the ISO19115 'MD_ScopeCode' list. At the moment, this is not possible in SITR-IDT, and therefore, in the next months, important changes in the metadata database conceptual model are going to be made.

3.3 Compliance with CNIPA (and INSPIRE) operational guidelines

The aim of the National Register of Spatial Data (RNDT) is to give a complete and general overview of all geographic data existing at a national level. In order to ensure its effective feeding, management and consultation, a geographic data register at a national level should contain a limited number of metadata for each single datum. A too detailed description of every data set would imply a great complexity in managing the whole national register. For this reason, the ISO19115 metadata set is not the most proper choice for the RNDT. Instead RNDT was structured according to the INSPIRE metadata requirements, plus some further fields.

Metadata in the RNDT should be organised in datasets and series of datasets, and, if needed, in subsets of datasets, i.e. in tiles. No further hierarchical levels should be individuated, because a higher detail level is out of the scope of a data catalogue at a National level.

Since CNIPA metadata structure is organised according to the INSPIRE IR, adherence to the CNIPA metadata schema ensures the compliance with INSPIRE. SITR-IDT metadata organisation does not need many important changes to reach this compliance. Since its beginning, for defining metadata fields SITR-IDT has followed CNIPA technical operational guidelines. Metadata elements are constituted by the ISO19115 core, plus some further fields of the same standard. In order to be fully compliant with the most recent version of CNIPA technical guidelines, these minor changes will be made in SITR-IDT:

– change for some existing fields that they do not have to be filled out (obligatory status), and
– introduction of some new fields, mandatory for the most recent CNIPA operational guidelines.

With these small modifications, SITR-IDT will be able to produce the *.xml* files necessary for feeding the RNDT, which will become mandatory in 2010.

4. METADATA OF NON-GEOGRAPHIC DATA

Textual documents, legends or other non-geographic data correlated with geographic data require appropriate metadata. INSPIRE IR and CNIPA operational guidelines do not give any precise indications about these data, which are, however, very significant and constitute a very important part of the geographic data.

However, ISO19115 standard recognises this kind of data, affirming that its criteria can be extended to many other typologies of geographic data such as maps, charts, textual documents and other non-geographic data (ISO19115, 2003). In SITR-IDT metadata of non-geographic data are described according to the Dublin Core set.

5. METADATA MANAGER

For metadata creation and management, a fundamental role is played by the tool adopted. In SITR-IDT this tool, which consists of dedicated software, is still being developed. It is a client desktop working on the local SITR-IDT intranet and connected via Open Database Connectivity (ODBC) with the Oracle metadata database.

Software was created at the beginning of the SITR-IDT project to access and manage the metadata database. Its conceptual model was structured according to the 'IntesaGIS' hierarchical data structure. SITR-IDT's metadata manager is tightly linked to this hierarchical metadata organisation, so that in creating metadata, its organisation in series, datasets and tiles is compulsory. Therefore, it is not possible to create a single dataset without the related series. This is clearly out of the ISO19115 conceptual model.

A new metadata manager is therefore under analysis. The new metadata manager should allow a flexible metadata organisation, and should work on a different metadata database conceptual model, according to INSPIRE and ISO19115. It should allow to catalogue metadata with all the values of the ISO19115 'MD_ScopeCode' list.

Currently, the SITR-IDT metadata manager can be accessed (both in consultation and in editing modality) only from the local SITR-IDT intranet. However, it might be helpful to extend this access in client-server mode to other Sardinian Public Administrations producing geographic data and provide these to the SITR-IDT database. It is nowadays well recognised that the most effective way to produce good quality metadata is to create it contemporarily to data production. This would be possible with a tool which allows any geographic data author, external to the SITR-IDT intranet, to access metadata manager and to create the metadata in parallel to data creation. After the data author has completed the metadata entry, the SITR-IDT metadata administrator should consult and validate the metadata. The implementation of this scenario implies to create different profiles for all the users of the metadata manager. The set up of different profiles for metadata consultation, editing, validation, publication, creation or elimination would then be necessary.

The possibility of creating metadata by several users implies that precise guidelines must be made available for all metadata creators. SITR-IDT will write down and publish among all the Regional institutional producers of geographic data, appropriate operational guidelines illustrating practical case-studies covering the largest range of data that are likely to be created at a regional level.

The new metadata manager should be able to export the *.xml* metadata files of different metadata sets, customisable depending on specific requirements (e.g. metadata set of the ISO19115 'Core', of the RNDT, or others).

The capability of connecting to different spatial databases (Oracle, PostgreSQL, HSQLDB) is being evaluated. Also compilation wizards could be implemented. This would allow the pre-compilation of some fields as the geographic bounding box for a shapefile, or as the distribution format for different typologies of file. A functionality for the versioning management of a metadata set should be also implemented. Finally the metadata manager should be so flexible to allow the implementation of all possible variations of national or international guidelines.

6. CONCLUSIONS

The Public Administration of Sardinia Region is making a great effort in terms of human and economic resources to create its Spatial Data Infrastructure (SITR-IDT). In the management of SITR-IDT, great importance is given to metadata. They are object of study concerning their creation, collection and management but also for more general questions such as the individuation of the most suitable database structure and of their appropriate organisation.

A good basis consisting of the individuation of the metadata structure compliant to ISO19115 standard and to INSPIRE IR has already been created in SITR-IDT. Feeding metadata to the National Register of Spatial Data will be possible without significant changes to the metadata structure adopted in SITR-IDT.

However, many further steps are still to be performed in SITR-IDT to develop the metadata database and to create the tools that allow for managing metadata of different typologies of geographic data in the most appropriate way, cataloguing the information at the most appropriate level of detail and allowing metadata management by different geographic data producers.

REFERENCES

CNIPA, Segreteria tecnica del Comitato tecnico nazionale per il coordinamento informatico dei dati territoriali (2006a). Repertorio Nazionale dei Dati Territoriali - Linee guida per l'applicazione dello Standard ISO 19115 Geographic Information - Metadata. Versione 0.2.

CNIPA, Segreteria tecnica del Comitato tecnico nazionale per il coordinamento informatico dei dati territoriali (2006b). Repertorio Nazionale dei Dati Territoriali - Linee guida per l'applicazione dello Standard ISO 19115 Geographic Information - Metadata. Versione 0.3.

Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata, at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R 1205:EN:NOT.

Draft Decree of the President of the Council of Ministers "Regulation governing the definition of the content of the National Register of Spatial Data, as well as the modalities of establishment and update of the latter", at: http://www.cnipa.gov.it/ site/it-it/Attivit%C3%A0/Sistemi_Informativi_Territoriali/Specifiche_tecniche/Regol amento_Repertorio/.

Drafting Team Metadata (2007). Draft Implementing Rules for Metadata (Version 3), at: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/INSPIRE_Metadata_ ImplementingRule_v3_20071026.pdf.

Intesa GIS (2004a). Specifiche per la realizzazione dei database topografici di interesse generale - Specifiche di contenuto: gli strati, i temi, le classi.

Intesa GIS (2004b). Specifiche per la realizzazione dei database topografici di interesse generale - Specifiche di contenuto: documento di riferimento.

ISO (2003). ISO 19115:2003(E); Geographic Information – Metadata.

ISO (2005). ISO 19119:2005(E); Geographic Information – Services.

An Integrated Framework for the Implementation and Continuous Improvement of Spatial Data Infrastructures

Ara Toomanian[1] and Ali Mansourian[2]

[1]Lund University, Sweden, Ara.Toomanian@gis.lu.se
[2]K.N.Toosi University of Technology, Iran, Mansourian@kntu.ac.ir

Abstract
Development and usage of proper frameworks for implementation, evaluation and continuous improvement of spatial data infrastructures (SDIs) is currently an important research topic. A wide range of methods are being researched. In this respect, methods and techniques on performance measurement and evaluation techniques from business management literature are not yet considered. Some techniques and methodologies from business management literature could be developed based on Six Sigma, ABC (Activity Based Costing), BSC (Balanced Scorecard) and TQM (Total Quality Management). This article describes these techniques and then provides an integrated framework, based on these business management techniques, for implementation and continuous improvement of SDIs.

Key Words: spatial data infrastructure (SDI), Implementation, Continuous Improvement, Six Sigma, ABC, BSC, TQM.

1. INTRODUCTION

In recent years, many countries implement and develop NSDI (Masser, 2005a). Also, scientists suggest operational platforms for the SDI implementation such as SDI business model (Wagner, 2005), SDI partnership (Warnest et al., 2002) and spatially enabling governments (Masser et al., 2007). Considering the fact that SDI implementation is a matter of technical, technological, social, institutional, political issues and also financial challenges (Nedovic-Budic et al., 2004; Masser, 2005b; Mansourian et al., 2006; Onsrud, 2007), different aspects and perspectives must be brought into attention for the progress of SDI implementation. Moreover, considering various dimensions of an SDI as spatial data production issue, data accessibility, data sharing, updating, standardisation and institutional matters, the need for a structured and integrated implementation framework is inevitable.

The next significant and essential requirement for an SDI implementation is the performance measurement and the continuous improvement due to the complexity and long term procedure of SDI implementation. In an SDI, it is important to have feedback from different dimensions and perspectives and to improve the weak points in order to have an effective and operational SDI. Such improvements may help to decrease additional costs and will lead to high quality spatial data products. Furthermore, all SDI users as well as the whole society will be satisfied with standard, accessible spatial data products and delivery within a high performance SDI.

A variety of research is conducted in accordance with SDI evaluation and performance measurement (Georgiadou et al., 2005; Georgiadou et al., 2006; Kok and van Loenen, 2005; McDougall, 2006; Van Loenen, 2006; Najar et al., 2006; Giff and Lunn, 2008; Fernández and Crompvoets, 2008; Lance et al., 2006; Grus et al., 2007). However, few attention is paid to business management literatures which provide proper techniques for performance measurement and evaluation.

In the business management literature, there are a variety of techniques which are used for continuous improvement of industries and/or organisational activities. Six Sigma, Activity Based Costing (ABC), Balanced Scorecard (BSC) and Total Quality Management (TQM) are some of these techniques that are also the targets of the article. Each of the mentioned techniques covers a dimension of SDI implementation. This article aims to address utilisation of these techniques as an integrated framework for implementation and continuous improvement of SDIs. Such integration will cover different aspects of SDI implementation and evaluation requirements. With this in mind, first, the techniques are reviewed briefly and then their feasible applicability for SDI implementation and evaluation is described.

2. CONCEPTUAL FRAMEWORK FOR SDI MEASUREMENT AND IMPLEMENTATION

In this section, we introduce a number of measurement methods used in the business management literature and describe their original purpose, then denote an integrated framework as an SDI implementation and evaluation procedure.

Six Sigma is a problem solving and continuous improvement method based on statistical methods where all the employees within an organisation have different roles within the entire technique. Six Sigma framework and guidelines can be used as a basic framework for SDI implementation.

Activity Based Costing (ABC) is a useful method to find the real price of the products according to the organisational costs and overheads. It also tries to assign costs to each activity and removes unnecessary and unprofitable tasks in an organisational process. ABC can be useful for estimating SDI costs as well as cost reduction and spatial data valuation.

Balanced Scoreboard (BSC) is a performance evaluation method used for evaluating and monitoring the strategic plans and the objectives. It can be used as an evaluation and monitoring method for SDIs and also for measuring the progress of SDI implementation according to different perspectives.

Total Quality Management (TQM) is a method to monitor the process quality. It deals with the entire product procedure and tries to keep the work process in a high standard. This method can be utilised as a proper technique for both quality control of SDI work process and spatial data.

Table 1 represents the usage domains and a description for various techniques discussed above. In the following sections, we will describe each method separately.

Table 1: A general overview of different business management methods.

Method	Key Premises
Six Sigma	Problem Solving and continuous improvement
ABC	Financial management and evaluation
BSC	Performance evaluation
TQM	Quality enhancement

2.1 Six Sigma

Six Sigma is one of the most effective problem solving methodologies for improving business and organisational performance. It was first originated and introduced by Motorola Company in 1987 and targeted an aggressive goal of 3.4 parts per million defects (Barney, 2002; Folaron, 2003). The background of Six Sigma method is a statistical approach where two main items are discussed:

− the roll up of characteristic behaviours, and
− the natural increase of variation for each characteristic in the long term.

Here, the sigma scale is a universal measure of how well a critical characteristic performs compared to its requirements. It works in such a way that if sigma score increase, the characteristic will be more capable (Gygi et al., 2005). Six Sigma is using a scientific, structured method for business improvement that could be used for any aspect of organisation, process or person.

Six Sigma is defined as ''high-performance, data-driven approach to analyse the root causes of business problems and solving them'' (Blakeslee, 1999). Other persons described Six Sigma as a disciplined and statistically based approach for improving product and process quality (Hahn et al., 2000). Also, Six Sigma refers to a business process that allows organisations to improve drastically their bottom line by designing and monitoring everyday business activities in ways that minimise waste and resources while increasing customer satisfaction (Harry and Schroeder, 2000). To achieve these aims, Six Sigma involves all employees in the organisational activities, according to their skills, and also obtains their feedback for problem solving and continuous improvement of the processes. Solving complex and strategic problems is conducted through experts and professionals and moderate tasks are carried out via medium level of skill and average trained employees. The regular transactions are conducted by other staffs.

The Six Sigma methodology has two project strategies, DMAIC (Define, Measure, Analyse, Improve, and Control) and DMADV (Define, Measure, Analyse, Design, and Verify), which are describe bellow. These strategies are a set of standardised and systematic methods that each project has to use in order to have a continuous improvement.

2.1.1 DMAIC

DMAIC is a problem solving and continuous improvement strategy for any kind of organisational strategy. It includes the following steps (Gygi et al., 2005):

− Define: writes the problem statement context and project objective setting;
− Measure: understands the process and improves the baseline performance and capability of the process or system;
− Analyse: uses data and tools to understand the cause and effect relationships in the process or system*;*
− Improve: determines and develops the modifications that lead to a validated improvement in the process or system and tries to implement solutions to achieve the objective statement, and
− Control: establishes plans and procedures and implements processes control methods to ensure the improvements are sustained.

To use this strategy, effective contribution of skilled and trained staffs, at different management levels, is essential. In other words, all employees have *fundamental* role for the DMAIC implementation. In addition, completing one step is a prerequisite for moving to the next step. After passing all steps successfully, a Six Sigma project is completed.

This strategy can be utilised for in an early stage of SDI implementation. In such situations, there are a number of initial tasks to start the SDI implementation procedure. Strategic plans, action plans, general objectives are some of the primary documents which have to be completed in the *define* step. Afterwards, within the implementation procedure of SDI, data production and delivery processes, collaboration among organisation for data exchange and also maintenance and standardisation of spatial data are measured and evaluated in the *measure* step. To perform this, integration of the Six Sigma measurement methods and the SDI evaluation indicators is suggested. *Analyse* deals with analysing the results of the measurements and identifying those barriers that impede SDI implementation and those positive points that facilitate the implementation. *Improve* enhances the procedures of SDI implementation regarding to the information derived from previous stage. Finally, *control* aims to check whether improvements of the previous steps caused the SDI implement in a proper way or not. As SDI implementation is a long term process, this methodology might be used many times as loop within the period of implementation.

2.1.2 DMADV

There are many similarities between DMADV and DMAIC. The major difference is in the last two letters which refer to *Design* and *Verify. Design* refers to either a new process or a corrective step to the existing one, eliminating the error origination that meets the target specification. *Verify* means verification by simulation of the performance of developed design and its ability to meet the target needs (Gygi, 2005). In DMADV, the processes change and redesign according to the customer's needs. Such change is needed in order to fit to the on demand requests instead of the improvement and control steps which more focus on readjusting and controlling by one way or other.

Although there are many overlaps in this strategy with the previous one, nations and societies which have already started an SDI implementation procedure and would like to extend or adjust it can use the DMADV strategy. In this strategy, the re-design of the SDI may extend or restructure the previous framework and then in the validate stage, it will be evaluated and monitored according to the new process and situation.

2.2 Activity Based Costing (ABC)

Financial aspects and cost are main features for SDI development. Even though the SDI budgets mainly stem from the government resources, these subjects are essential for the SDI managers to succeed in the spatial data market.

In the traditional way of management and accounting methods in the 1930s, corporate rules had a basic role to force companies for providing financial accounts. Although the application of strict rules was a proper way for financial accounts, management accounts were proposed as a decision-making tool in business atmosphere and therefore required more flexibility (Letza and Gadd, 1994). In such a method, production overhead was absorbed to the product cost to valuate the stock. Moreover, labour costs were used as a convenient overhead recovery base, although the ratio of the total labour cost was not proportional.

However, the traditional methods often fail to incorporate the final cost today. The reason is that the technological costs and other overheads have increased rapidly, due to the expansion of global competition, and the increase of interactions via communication media, development of IT and access to inexpensive information systems. Therefore, new accounting methods such as Activity Based Costing (ABC) have been introduced.

ABC was first introduced in the late 1980s by Johnson and Kaplan (1987). Scientists expanded the first initial idea and developed a method for cost drivers to calculate activity costs for each product and service. They argued that such method supplies accurate cost data needed to make proper strategic decisions for product mix, sourcing, pricing, process improvement, and evaluation of business process performance (Cooper and Kaplan, 1992; Swenson, 1995).

ABC is a costing model which determines the activities in an organisation and assigns the cost of each activity resource to products and services separately regarding to the actual usage by each. It also generates the real cost of products and services by removing unprofitable activities and eliminate lowering prices of overpriced ones. Here, an activity is defined as a discrete task that a company makes in a product or service, and uses cost drivers to assign activity costs to products, services or customers related to these activities (Cooper, 1988; Ittner et al., 2002). In this method, products use activities and the activities use resources.

ABC has two main stages to assign overhead costs to products and services (Hilton, 2005). First, based on the definition, the main activities are determined and overhead costs are assigned to the activity cost pools according to the amount of resources used by activities. The activities are often derived from information gathered from interviews, questionnaires, and time cards (Cooper and Kaplan, 1991). The second stage contains cost allocation from each activity cost pool to each product line concerning to the amount of the cost driver utilised by the product line (Bjornenak and Mitchell, 2002). In other words, at the first step, organisational resources are grouped in the different pools such as salaries, license fees, operational costs and depreciation. Then, different institutional missions and tasks are grouped into homogeneous activities such as data preparation, research and development (R&D), data delivery (Ooi and Soh, 2003). In this way, each activity will use a percentage of a single or multiple cost pools. For example, the data preparation activity will use 10% of the rental cost, 20% of the salary and 40% of the operational costs.

As ABC reveals the links between performing particular activities and the demands those activities make on the organisation's resources, it provides managers with a clear picture of how products and services both generate revenues and consume resources. The profitability picture that emerges from the ABC analysis helps managers focus their attention and energy on improving activities that will have the biggest impact on the result.

An important part of SDI implementation are the financial and economical issues. A proper financial funding model may lead the SDI coordinators to a successful and operational SDI. Furthermore, having a clear idea about the SDI cost and the way of cost reduction will also increase the efficiency of SDIs. With this in mind, using the ABC method, main activities of SDIs are determined and according to the transparent implementation tasks, unprofitable and parallel activities will be eliminated. Also, in each step, the financial resources can be predicted with respect to different contributors

whether the financial support is from the spatial data market or authorities. Moreover, for any task and process within a clear financial and economic perspective, evaluation and monitoring can be easily performed by the SDI coordinators.

2.3 Balanced Scorecard (BSC)

The success of the next generation of Spatial Data Infrastructures (SDIs) will, in part, depend on the ability of SDI coordinators to comprehend, analyse and report on the performance of their initiatives (Giff and Lunn, 2008). Therefore, it is necessary for SDI coordinators to use proper models and measuring techniques to assess and monitor the progress of SDIs.

BSC, as a technique from business management literature for strategic performance management, was introduced by Kaplan and Norton (1992) as a set of different measures that allow for a holistic, integrated view of business performance. It was a complementary solution for the traditional financial parameters to measure the performance in organisations. In other words, BSC is a performance measurement framework that provides an integrated look at the business performance of an organisation by a set of measures including both financial and non-financial metrics (Kaplan and Norton, 1992; Kaplan and Norton, 1996). Also, BSC refers to a multi-dimensional framework that uses measurement to improve an organisation's strategy.

There are some basic elements in the BSC structure which leads the strategy measurement in a proper way. A *perspective* is an element into which the strategy is decomposed to drive implementation. In most BSC structures, there are four perspectives: *financial*, *internal process, customer*, and *learning and growth*. As Norton and Kaplan (2000) mentioned, "Balanced Scorecards tell you the knowledge, skills and systems that your employees will need (learning and growth) to innovate and build the right strategic capabilities and efficiencies (internal processes) that deliver specific value to the market (customer) which will eventually lead to higher shareholder value (financial)". It is possible to add other perspectives or sometimes replace the mentioned perspectives according to the specific strategies. The perspective can be defined as an interpretation of the strategy in different dimensions.

The second main element of the BSC design is called *objective*. An objective is a statement of strategic intent, describing how a strategy will be made operational in an organisation. In other words, objectives are the main elements of the strategic plan and the entire strategy can be broke down into many objectives. In the BSC design, normally a limited number of objectives exist relating to one of the perspectives, which is normally described in one or two sentences.

The next basic element in a BSC design is the *cause and effect linkage*. In the BSC structures, objectives are related and depend on each other through cause and effect relationships. The cause and effect linkages are like *if – then* statements where the objectives in each perspective are linked with the graphical connectors according to the rules derived from different dimensions.

Another element of the BSC is the *measure* term, which is a performance metric one can calculate the progress of an objective. A measure must be quantifiable. In a BSC design there are reasonable numbers of measures explicitly linked to an objective. In addition, the measure concept is typically represented via mathematical formulas.

The fifth element for BSC design is called *target*. A target is a quantifiable goal for the each measure. A combination of targets on the BSC design is the general goal of an organisation. They help the organisation monitor the progress toward strategic goals, and give proper feedbacks if necessary.

Strategic initiative is the last element of a BSC design. They are action programs that drive strategic performance and the activities which will lead the organisation to achieve the strategic results. All ongoing initiatives in an organisation should be associated with the strategy in the BSC.

BSC design can be used as an evaluation and monitoring framework for SDIs. By defining performance indicators as well as desired targets, for each objective, SDI coordinator and managers can measure a current situation, compare it with the target and then evaluate the progress of an SDI. Considering the four main perspectives in the BSC structure, BSC provides a general framework for evaluating SDIs from users' and data producers view point. It also helps to evaluate internal processes, financial affairs and even capacity building at the individual level. So BSC can be regarded as a general framework for an SDI evaluation.

2.4 Total Quality Management (TQM)

SDI implementation requires intra-organisational activities which imply that there are various hierarchical management decision making steps in different levels. Having a proper tool for increasing the quality of the entire procedure leads the SDI to succeed in not only high quality data production and management, but also in facilitating data sharing and access. Therefore, applying a quality management approach for the development of SDI is essential.

TQM consists of three main concepts. *Total* refers to the organisation (e.g., SDI organisation) and includes the whole supply chain and product life cycle. *Quality* means a high degree of excellence in products and also the comparison indicators with the existing standards. *Management* is the process of planning, organising, leading, coordinating, controlling and staffing (Fayol, 1966). TQM is a collection of principles, techniques, processes, methodologies, tools and best practices that over the time have been proven effective in order to increase the internal and external customer satisfaction with a minimum amount of resources.

Sashkin and Kiser (1993) defined TQM as an intense and long-term commitment to quality implementing such a commitment requires the use of tools and techniques. The commitment is more important than the way of utilising the method. TQM is a method to change the organisational values and beliefs in order to let everyone know the most basic aim which is the quality for the customer. Also the ways of working together are determined by what will support and sustain this basic aim (Sashkin and Kiser, 1992). On the other hand, they argued such a system as a shift in the way of thinking and the culture of an organisation rather than using a specific software, technique or specific tool (Sashkin and Kiser, 1993). TQM tools include quality training, process improvement, benchmark management, Statistical Process Control (SPC), Quality Control circle (QCC) and quality information computerisation (Huarng and Chen, 2002).

There are many scientists working to improve the TQM method. Edwards Deming (1986, 1993) introduced fourteen management principles as requirements to remain competitive in providing products and services. These include management commitment and leadership, statistical process control, removing barriers to employee partici-

pation and control of their own quality, and continuous improvement of processes. Juran (1989) emphasised planning and product design, quality audits, and orienting quality management toward both suppliers and customers. Crosby (1984) focused on such organisational factors as cultural change, training, and leadership, and the ongoing calculation of quality costs. Important extensions to the TQM framework have included the development of customer-based specifications in the design of a product or process (Taguchi and Clausing, 1990), and benchmarking or the measuring of products/services and processes against those of organisations recognised as leaders (Camp, 1989).

TQM can be used as a general instrument for quality control of the SDI implementation procedures. To utilise such a technique in SDI, the fourteen step approach of Deming can be used in the SDI implementation procedure.

3. DISCUSSION

To investigate the applicability of the mentioned techniques for improving the development and maintenance of SDIs, this section investigates the pros and cons of each technique and their affects on SDI. Table 2 summarises the strength and weak points of each technique.

SDI is a collaborative effort: various organisations and institutions are involved in the development and implementation of SDI. Thus, team work and joint activities have a major role in arriving at the objectives of an SDI. One of the strengths of Six Sigma relates to team building and facilitating team working (see Table 2). This technique can be used for creating the collaborative environment, which is required for the development of SDI. In addition, to develop an SDI, different procedures (spatial data production and updating during daily businesses, inter-and intra organisational data sharing, managing databases and web services) have to be diffused within the organisations. Integrating human elements (culture change, user focus, spatial data-related responsibilities) with process elements (process management, measurement system analysis) can facilitate such diffusion. As highlighted in Table 2, 'integration of human and process elements' is another strength point of Six Sigma, which makes it a suitable technique in the work with implementing SDI.

With respect to the weak points of Six Sigma, 'the need for high quality data for the evaluation' can be considered as the weakness of the most evaluation and improvement methods. Also, since the priority of major activities for implementing an SDI is generally clear, so 'the prioritisation of projects' (Table 2) is not too critical for using Six Sigma for SDI implementation and continuous improvement.

Financial management of SDI is a complex task. Due to diversity of activities required for implementing an SDI, calculating the costs associated for each activity as well as relevant overheads calls for adopting proper financial frameworks. The framework should also provide the possibility of monitoring SDI funding for each activity, based on the mentioned estimations. ABC with the advantage of 'clarification and calculation of the real cost for the products, services, processes and distribution channels' and 'supporting performance measurement' (Table 2) can satisfy such an SDI's requirement. ABC is also easy to understand and well integrated with Six Sigma.

Table 2: Strengths and weaknesses of discussed methods according to the SDI.

Method	Strengths	Weaknesses
Six Sigma	– Team Building and Facilitation – Integration of the human and process elements	– Requiring quality data available for the measurement – Prioritisation of projects is critical
ABC	– Easy to understand – Accurate measurement of costs – Well integration with Six Sigma and other continuous improvement tools – Supports performance measurement and scorecard – Enables costing processes	– Time consuming for data collection – ABC implementation cost
BSC	– The ability to link: o Financial and non-financial indicators o Internal and external aspects o Performance drivers and outcomes – Organising disparate data, and providing benchmarks for management discussion and operations. – Highlighting inevitable trade-offs	– Higher weight of financial measure
TQM	– Encourages effective participation	– Requires much time and effort

Two weak points of ABC, mentioned in Table 2, might not be critical from an SDI perspective as:

– 'time consuming for data collection' is the limitation of the most monitoring and evaluation approaches, not specifically for ABC, and
– ABC implementation cost will be a small percentage of the financial resources required for the SDI implementation.

Due to complex and multi-dimensional nature of the SDI development, its evaluation and monitoring should be based on a multi-view framework linking financial and non-financial indicators, internal and external aspects, and performance drivers and outcomes. BSC not only has the advantage of linking the mentioned factors, but also can highlight inevitable trade-off among them. Therefore, BSC can be a proper framework for the implementation and evaluation of SDIs.

Regarding the weakness of BSC, from an organisational perspective, a financial measure has much greater organisational weight than its new non-financial sibling. However, in SDI, besides financial benefits of spatial data sharing, non-financial benefits of spatial data usage in decision making and planning is also of high value. Furthermore, social benefit gained from SDI has more weight than any financial indicator. Governments spend much money for SDI development to promote the society and better life for citizens, so the financial perspective is important, but not the most significant dimension of SDI implementation.

Finally, TQM encourages effective participation by involving people in the decision making process for development of SDI and improving the quality of their work environment provides them with a sense of value and purpose. Similar to the other methods, TQM is also a long-term procedure and implementation of TQM takes too much time and effort.

With respect to this description, the mentioned techniques can be used for different aspects of SDI implementation, monitoring and improvement.

4. CONCLUSIONS

This article proposes instruments and frameworks from the business management field for the implementation and evaluation of SDIs. We first reviewed different strategic and continuous improvement methods including Six Sigma, ABC, BSC and TQM. Then the applicability of each technique for the implementation of SDI was investigated. The primary investigation shows that each of these techniques can be used in some aspects of SDI implementation. In a nut shell, an integrated general framework for the SDI implementation consists of the Six Sigma as a core methodology. For implementing an SDI, the DMAIC (Define–Measure–Analyse–Improve–Control) approach can be used; ABC (Activity Based Costing) for economic management of SDI; BSC (Balanced Scorecards) for monitoring the progress of SDI and TQM (Total Quality Management) for the quality management of the entire procedure of SDI implementation.

Table 3 illustrates the summary of the usages and value of each method for SDI implementation and continuous improvement.

Table 3: Summary of the usage and value for SDI from discussed methods.

Methods	Value for SDI (Where it can be used)	Usage Domain
Six Sigma	A general framework for the SDI Implementation	Core methodology
ABC	Economic management and evaluation of SDI	Define, Measure, Improve
BSC	Monitoring the progress of SDI	Measure, Analysis, Control
TQM	Quality management of the whole procedure of SDI implementation	Measure, Analysis, Improve, Control

It is worth to be noted that the discussed techniques are originally used for business management and continuous improvement within an organisation. However, also for the implementation of SDI with its collaborative and intra-organisation nature, applying these techniques may be beneficial and worthwhile to be considered by the SDI community.

ACKNOWLEDGMENTS

REFERENCES

Barney, M. (2002). Motorola's second generation. *Six Sigma Forum Magazine,* 1(3): 13-16.

Bjornenak, T. and F. Mitchell (2002). The development of activity-based costing journal literature, 1987-2000. *European Accounting Review*, 11(3): 481-508.

Blakeslee, J. (1999). Implementing the Six Sigma solution, *Quality Progress,* 32(7): 77-85.

Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance*. Milwaukee, WI: Quality Press.

Cooper, R. and R. Kaplan (1992). From ABC to ABM, *Management Accounting*, 74: 54-57.

Cooper, R. and R. Kaplan (1991). Profit priorities from activity based costing, *Harvard Business Review,* 69(3): 130-135.

Cooper, R. (1988). The rise of activity-based costing – part two: when do I need an activity-based costing system? *Journal of Cost Management* (fall) 2(3): 41-48.

Crosby, P. B. (1984). *Quality without tears: The art of hassle-free management*. Milwaukee, WI: Quality Press.

Deming, W. E. (1993). *The new economics for industry, government, education*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Deming, W. E. (1986). *Out of the Crisis*. Cambridge, Massachusetts Institute of Technology: Center for Advanced Engineering Study.

Fayol, H. (1966). Administration industrielle et générale - prévoyance organisation-commandement, coordination – contrôle, Paris: Dunod.

Fernández, T.D. and J. Crompvoets (2008). "Evaluating Spatial Data Infrastructures in the Caribbean for sustainable development", *Proceedings of the 10th GSDI conference*, February 25-29, 2008, St. Augustine, Trinidad.

Folaron, J. (2003). The evolution of Six Sigma. *Six Sigma Forum Magazine,* 2(4): 38-44.

Georgiadou, Y., Rodriquez-Pabón, O. and Lance, K.T. (2006). Spatial data infrastructure SDI and e-governance: a quest for appropriate evaluation approaches, *URISA journal*, 18(2): 43-55.

Georgiadou, Y., Puri, S.K. and S. Sahay (2005). Towards a potential research agenda to guide the implementation of spatial data infrastructures - A case study from India, International Journal of Geographical Information Science, 19(10): 1113-1130.

Giff, G. and R. Lunn (2008). "Designing Performance Indicators for Local SDI Assessment: A City of Fredericton Case Study", *Proceedings of the 10th GSDI conference*, February 25-29, 2008, St. Augustine, Trinidad.

Grus, L., Crompvoets, J.W.H.C. and Bregt, A.K. (2007). Multi-view SDI assessment framework, *International Journal of Spatial Data Infrastructures Research*, 2: 33-53.

Gygi, C., DeCarlo, N., Williams, B., Stephen and R. Covey (2005). Six Sigma for Dummies: For Dummies.

Hahn, G.J., Doganaksoy, N. and R. Hoerl (2000). The evolution of Six Sigma, *Quality Engineering*, 12(3): 317-26.

Harry, M.J. and R. Schroeder (2000). *Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations*, New York: Doubleday.

Hilton, R.W. (2005). *Managerial Accounting: Creating Value in Dynamic Business Environment*, New York: McGraw-Hill Irwin.

Huarng, F. and Y.T. Chen (2002). Relationships of TQM philosophy, methods and performance: A survey in Taiwan, *Industrial Management+Data systems*, 102(3/4*):* 226-234.

Ittner, C. D., Lanen, W. N. and D. F. Larcker (2002). The association between activity-based costing and manufacturing performance, *Journal of Accounting Research*, 40: 711-726.

Johnson, H. and R.S. Kaplan (1987). *Relevance Lost: the Rise and fall of Management Accounting,* Boston, MA: Harvard Business School Press.

Juran, J. M. (1989). *Juran on leadership for quality: An executive handbook*. Wilson, CT: Juran Institute.

Kaplan, R. S. and D. P. Norton (2000). Having trouble with your strategy? Then map it, *Harvard Business Review,* 78(5): 167-176.

Kaplan, R.S. and D.P. Norton (1996). Using the balanced scorecard as a strategic management system, *Harvard Business Review*, 74(1): 75-85.

Kaplan, R.S. and D.P. Norton (1993). Putting the balanced scorecard to work, *Harvard Business Review*, 71(5): 43-2.

Kaplan, R.S. and D.P. Norton (1992). The balanced scorecard: measures that drive performance. *Harvard Business Review*, January–February: 171-9.

Kok, B. and B. van Loenen (2005). How to assess the success of National Spatial Data Infrastructures? *Computers, Environment and Urban Systems*, 29: 699-717.

Lance, K.T., Georgiadou, Y. and Bregt, A. (2006). Understanding how and why practitioners evaluate SDI performance, *International Journal of Spatial Data Infrastructures Research*, 1: 65-104.

Letza, S. and K. Gadd (1994). Should activity-based costing be considered as the costing method of choice for total quality organizations? *The TQM Magazine*, 6(5): 57-63.

Mansourian, A., Rajabifard, A., Valadan Zoje, M.J. and I. Williamson (2006). Using SDI and web-based systems to facilitate disaster management, *Journal of Computers and Geosciences*, 32: 303-315.

Masser, I., Rajabifard, A. and I. Williamson (2007). Spatially enabling governments through SDI implementation, *International Journal of Geographical Information Science*, 22(1): 5-20.

Masser, I. (2005a). "Some priorities for SDI related Research", *FIG Working Week 2005 and GSDI-8*, Cairo, Egypt April 16-21, 2005.

Masser, I. (2005b). "The Future of Spatial Data Infrastructures", *ISPRS Workshop on Service and Application of Spatial Data Infrastructure*, XXXVI (4/W6),14-16 Oct 2005, Hangzhou, China: pp. 7-15.

McDougall, K. (2006). A Local-State Government Spatial Data Sharing Partnership Model to Facilitate SDI Development, The University of Melbourne.

Najar, C., Rajabifard, A., Williamson, I. and C. Giger (2006). "A Framework for Comparing Spatial Data Infrastructures: An Australian-Swiss Case Study", *Proceedings of GSDI 9 - International Conference for Global Spatial Data Infrastructure*, 3-11 Nov 2006, Santiago, Chile.

Nedovic-Budic, Z., Feeney, M.E.F., Rajabifard, A. and I. Williamson (2004). Are SDIs serving the needs of local planning? Case study of Victoria, Australia and Illinois, USA, *Computers, Environment and Urban Systems*, 28: 329-351.

Onsrud, H. (ed.) (2007). *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ESRI Press: Redlands, California.

Ooi, G and C.Soh (2003). Developing an activity-based costing approach for system development and implementation, *SIGMIS Database*, 34(3): 54-71.

Sashkin, M. and K.J. Kiser (1993). *Putting total quality management to work*. San Francisco: Berrett-Koehler.

Sashkin, M. and K.J. Kiser (1992). *Total quality management assessment inventory*, Seabrook, MD: Ducochon Press.

Swenson, D. (1995). The benefits of activity-based cost management to the manufacturing industry, *Journal of Management Accounting Research*, 7: 167-180.

Taguchi, G. and D. Clausing (1990). Robust quality, *Harvard Business Review*, 68(1): 65-75.

Van Loenen, B. (2006). Developing geographic information infrastructures, the role of information policies. PhD thesis, Delft University of Technology, Delft, the Netherlands.

Wagner, R.M. (2005). "Need for a Framework SDI Business Model", *Joint EuroSDR/ EuroGeographics Workshop*, BKG Frankfurt, Germany, 23-25 February.

Warnest, M. and Rajabifard, A. and I. P. Williamson (2002). "Snapshot of SDI Development in Australia: Models, Partnerships and Lead Agencies Advancing Implementation", *Proceedings of Joint AURISA and Institution of Surveyors Conference*, Adelaide, South Australia.

The Value Chain Approach to Evaluate the Economic Impact of Geographic Information: Towards a New Visual Tool

Elisabetta Genovese[1], Stéphane Roche[1], Claude Caron[2]

[1]Département des Sciences Géomatiques, Université Laval
{elisabetta.genovese, stephane.roche}@scg.ulaval.ca
[2]Groupe Géobusiness, Faculté d'administration, Université de Sherbrooke
claude.caron@usherbrooke.ca

Abstract
Geographic Information (GI) is becoming more important everyday at all levels of society. GI has a central role in supporting economies, improving business effectiveness in the private sector, enabling more efficient governments, and increasing citizens' quality of life. Assessing the value of digital information products, services and infrastructures is particularly complex due to the specific characteristics of GI as a not- standard economic good (Krek and Frank, 2000) and the nature of the GI market itself (Krek, 2006). One promising assessment approach is the value chain: value is created step-by-step along the chain. Thus, pricing in a value chain serves to determine the way in which the value created for the end user is distributed among the contributors. In theory, the value chain is one of the most suitable approaches to assess GI. However, it is also one of the most complex one due to the number of variables connected to how GI is produced and used. Therefore, it is often impossible to determine a single and constant value to specific GI (Longhorn and Blakemore, 2008) and a concrete example of application of a formal economic analysis based on the value chain concept still does not exist (Genovese et al., 2008). The EcoGeo project, in its first phase, has developed a prototype computer tool named Socioscope, which provides cartography of the links existing between various public and private contributors (Plante, 2006). In EcoGeo's second phase, Socioscope will be upgraded and the value chain of the geomatic sector in Quebec will be defined. The final goal of the project is an economic evaluation for a test-area inside the value chain: the ability to measure the GI economic value will provide key decision support for both institutional and private sectors.

Keywords: geographic information, geomatics, value chain, socio-economic assessment.

1. INTRODUCTION

Geographic information (or geomatics, the term which is used in Canada and France to express the concept) occupies a growing place in modern societies. Common examples include supporting both routine and key decisions in areas such as public health, public safety, emergency services, environmental issues, land management, forestry, agriculture, urban and rural planning, and retail analysis, among others (Samborski, 2007).

The impact of the Internet on the pricing of information and communications has been substantial and GI has gradually followed a democratisation process (Gauthier, 1999; Noucher and Archias, 2007). Information that was expensive and reserved for specialists (Longhorn and Blakemore, 2008) is now accessible to all users. For example, the advent of free web mapping services has allowed the wider public to have free access to GI and easy-to-use web-mapping technology. Likewise, Global Positioning System

(GPS) devices have changed the nature of GI production for both geomatic experts and common users (Bergeron et al., 2005; Caron et al., 2006).

GI is considered to be extremely valuable and its collection, processing and management are expensive. Conversely, it is inexpensive to disseminate (Longhorn and Blakemore, 2008). Everyone is a user of GI, and the same information can be used by all segments of society - citizens, businesses, and public bodies - usually for different intentions. Therefore, there is a wide debate on how society assigns different values to GI.

Despite the massive use of GI and the substantial body of literature on return on investment for general GI technology projects, a scientific framework providing the appropriate criteria (necessary to determine and compare the value and the benefit associated with GI projects) has yet to be defined (Genovese et al., 2008). The high number of variables and attributes that has to be taken into consideration when evaluating GI may explain this (Longhorn and Blakemore, 2008).

Thus it is necessary to define the basis on which or the conventions with which it will be possible to evaluate the effectiveness of investment in the implementation of GI projects. Recent studies (Crompvoets, 2006; Samborski, 2007; Grus et al., 2007; ACIL Tasman, 2008; Genovese et al., 2008; Longhorn and Blakemore, 2008; Crompvoets et al., 2008) have developed theoretical frameworks to evaluate GI.

On the practical level, the "lack of knowledge" about GI clearly hinders decision-makers and policy-makers, both in public and private sectors. Without accurate quantitative studies, it is difficult to identify the GI value chain and to evaluate the benefits of committing inadequate budgetary funds to investments in GI products or infrastructures (Genovese et al., 2008).

2. ASSESSING GEOGRAPHIC INFORMATION

Increasing volumes of GI resources were produced in the 1990s by public, private and non-government agencies (Onsrud, 1998). Over the past decade, many countries have responded to the need of standardising data quality and documentation formats by investing in national, regional or supranational Spatial Data Infrastructures (SDIs). The Canadian Geospatial Data Infrastructure (CGDI) and the United States' National Spatial Data Infrastructure (NSDI) are two examples of national SDIs, while Europe's INSPIRE and Australia and New Zealand's ANZLIC are examples of supranational SDIs (Genovese et al., 2008).

These investments in SDIs have improved access to GI and services, reducing data production costs through less duplication, providing policies, tools and mechanisms to promote data sharing at regional, national, and international levels (Crompvoets, 2006). Consequently, at all community levels, it is increasingly important to assess the GI return on investments to justify the resources spent on those infrastructures (Grus et al., 2007).

Until recently, the impact of these expenditures had yet to be evaluated systematically (Crompvoets, 2006). The INSPIRE directive (2007/2/EC), which establishes an Infrastructure for Spatial Information in the European Community, explicitly calls for regular monitoring and reporting to determine the extent to which the initiative is successful and to establish the impacts the infrastructure has on social and economic systems (Directive 2007/2/EC).

The definition of assessment strategies is challenging due to several characteristics and attributes distinctive of GI products, services and infrastructures. Several chief studies come from the Unites States. The Geospatial Information and Technology Association's suggest a formal methodology for the preparation of business cases for shared data and services for GI technologies within and across multiple agencies (Samborski, 2007). The U.S. Geological Survey (USGS) conducted a cost-benefit analysis of the National Map (Halsing et al., 2004) and the National Aeronautic and Spatial Administration (NASA) used cost-benefit methodologies to quantify the value of geospatial interoperability standards, as well as to determine when and for whom the benefits increase (Booz, 2005).

However, most of the studies focusing on cost-benefit analysis are performed before a purchase or a project to determine whether the expense is justifiable on organisational and financial grounds (Didier, 1990; Rodriguez, 2005) and a negative benefit cost ratio for GI has never been reported (Longhorn and Blakemore, 2008). Analyses of this type are usually focused on set-up costs and short-term efficiency benefits, which are relatively easy to assess compared to longer-term social, political and economic benefits (Craglia and Nowak, 2006). Indeed, since adopting GI technologies can have transformative effects on organisations, tangible benefits can take several years to materialise (Samborski, 2007).

Determining the value of investments in SDIs and in the wider GI sector has a further peculiarity: some benefits are intangible and could include customer and citizen goodwill, decision making quality, employee morale, quality of life, and environmental health, among others (Didier, 1990; ACIL Tasman, 2008). These intangible benefits are not easily estimated. Although they do not affect the monetary analysis, they can be equally or even more important than the tangible benefits (Genovese et al., 2008). Therefore, when quantifying the effectiveness of GI solutions, it is significant to consider both the affect these technologies have on society, as well as society's influence on the development of these technologies (Goodchild, 1995; Tulloch et al., 1998; Roche and Caron, 2004; Chrisman, 2005).

3. VALUE CHAIN CONCEPT

3.1. The value of GI

There are several ways to define value, some are monetarily quantifiable while others are not, but most depend upon the reason for which GI was initially collected or required. The commercial value of data or services, also known as exchange value, is just one of them. In fact, different kinds of value can be assigned to the same information depending on whether it is used by or for different people, at different moments, in different formats, or used for purposes different from the one for which it was originally collected (Longhorn and Blakemore, 2008).

As a result, information has several types of value and therefore different measures will not apply equally to all information in all conditions. In the private sector, the monetary value of a GI product must be adequate enough to recover the costs of data collection, processing, dissemination, and management with a suitable return on investments; if not, the product will not survive on the market (Krek, 2004).

Understanding the value of a good is essential to defining the issue of pricing in both the public and private sectors because the price of a good is intimately tied to the potential buyer's perception of the value. In order to set a price for GI, the producers are

required to know which properties the potential buyers consider significant and the extent to which these properties are valued (Krek, 2002).

The commercial sector of GI has often identifiable monetary value for its producers and vendors, but difficulties arise when considering the public sector, for which GI also has a direct and indirect social value. The public sector value is problematic to quantify due to the countless uses and social objectives, which can be aimed at the same GI data or services (Longhorn and Blakemore, 2008). Moreover, GI is supposed to be valuable not only for the data owner or user, but also for society as a whole.

3.2. The value chain

The multi-stage processes of modifying GI from its original form to create new derived products are particularly important when assessing the value GI has to an economy. Porter (1985) defined the concept of the value chain for classic production, expressing that activities within the organisations add value to the services and products that they create and distribute. This set of value-adding activities is defined the value chain. The concept of value chain is also used in the context of supply chain management to describe the added-value flow, which provides the revenue stream for each stage of the supply chain (Cox, 1999).

In the context of GI, the value chain relates to the set of value-adding operations undertaken by one or more producers, to transform GI (datasets or analogue maps) to the final product (Krek and Frank, 2000). Assessing the GI value chain entails many variables, the foremost being that GI is not a standard economic good as it is often not possible or reasonable to restrict any person's use of the information (non-excludible). Thus, the same dataset may be used repeatedly and new products can easily be created by forming different combinations of datasets. Furthermore, as a public good, GI is defined as "non-rivalrous" by Krek and Frank (2000): one person's use of that information does not limit the amount of the good available for consumption by others (Samuelson, 1954; Krek and Frank, 2000). These characteristics of GI data were emphasized by the spread of the Internet and the resulting democratisation of GI. To manage this, organizations frequently face enforcement costs, including cost of protecting rights, policing and enforcing agreements (Krek, 2003).

Value is created step-by-step along the chain, even if most of the costs are incurred during the initial data collection (Krek and Frank, 2000). This initial cost represents a high percentage of the total cost of producing a dataset, a usual characteristic of information goods (Shapiro and Varian, 1999), which is due to a high cost of labour while capturing or measuring the data from the data sources, as well as the cost of data transformation, analysis, and modification. The high cost of data collection seems to justify high prices for data at the first stages of the value chain.

Moreover, exchanging GI involves transaction cost. Transaction costs consist of cost of searching for the information about the possible data sellers or producers and cost of contact the possible providers. They include measurement cost (the cost of measuring the valuable attributes of that which is being exchanged) and enforcement cost (the cost of protecting and enforcing the property rights) (North, 1997; Krek, 2003). As transaction costs are not often transparent, they are difficult to measure. Moreover, they were completely ignored by neoclassical economic models, which are thus not easily applicable to GI (Krek, 2003).

There may be several activities performed at each stage of the value chain, and more than one type of value chain may apply. Having defined the value chain for a product or service, the organisation can assign costs to the activities along the chain. The value chain concept has been extended to the information sector by a range of proposals for an information value chain. As reported from Longhorn and Blakemore (2008), many authors have proposed different information value chains stemming from different points of view (Kaplinsky and Morris, 2001; Phillips, 2001; Spataro and Crow, 2002; Oelschlager, 2004).

A value chain dedicated to GI has still not been defined, since a high number of variables and factors are related to the production and dissemination of GI: context, attributes, timeliness, quality, accuracy, provenance, history (when data was collected, validated, and updated), among others (Longhorn and Blakemore, 2008).

4. THE ECOGEO PROJECT

4.1. Geomatics in Quebec: a brief overview

The use of geomatic technologies in Quebec's public and private sectors has expanded substantially over the past twenty years, in terms of the on-line application of GI technologies to new sectors such as insurance, banking, public health, transportation, and emergency services.

The government of Quebec is the most extensive user and producer of GI in Quebec and has consistently supported, directly and indirectly, policies and practices of innovation on the global scene of geomatics (website 1). To support the development of geomatics within government Ministries and Organisations (M/O), Quebec established an interdepartmental coordination infrastructure called Geomatic Plan of the Government of Quebec (PGGQ) in 1988. The PGGQ has contributed to the development of geomatic applications within the M/O by publishing four reports. The first was completed in 2000-2001.

In 2005, the Council of Ministers requested that the Ministry of Natural Resources and Fauna undertake a study to develop a model to guide the cooperative M/O production and distribution of free on-line GI by 2010. This initiative follows existing initiatives such as Ontario's Land Information Ontario (website 2) and the federal GeoGratis portal (website 3) provided by Natural Resources Canada (NRCan).

In 2007, the Ministry for the Economic Development of Quebec (MDEIE) recognised the geomatic sector as one of the poles of excellence of Quebec, within the framework of the governmental program ACCORD, which is aimed at financially supporting the sectors of economic activities of excellence in the greater metropolitan regions of Quebec and Montreal. Therefore, the geomatic sector of Quebec constitutes a relevant test sector to develop a new approach for the evaluation of GI.

4.2. Origins and goals of the EcoGeo project

In 2000, the Treasure Council of the Government of Quebec asked the General Direction of Geographic Information (DGIG) of the Ministry of Natural Resources (MRNF) to evaluate the economic benefits of a provincial map for the entirety of Quebec society. In this context, the EcoGeo project started in 2004. Its main goal was establishing an economic evaluation model of the geomatic sector in Quebec.

The first phase of the EcoGeo Project (EcoGeo I, in 2006) provided a visual represen-
tation of the overall flows of GI between the main private and public stakeholders of the
geomatic sector in Quebec. Firstly, the EcoGeo I team identified stakeholders involved
on the Quebec market. Secondly, it proposed a new visual tool (called Socioscope) to
represent and analyse GI flows (see section 5).

Based on this previous work, the EcoGeo phase II has two main goals. The first one is
to improve the first version of Socioscope, taking into consideration the entire life-cycle
of GI products in Quebec, and including data collection, production and processing.
This will also permit to define the specific value chain for the GI sector in Quebec. The
second challenging objective is to develop an evaluation model on a test subsector of
the value chain, in order to assess the economic impacts of GI on that selected area.

5. THE SOCIOSCOPE PROTOTYPE

5.1. GI flow modelling

The GI flow modelling of Socioscope is based on three steps: (1) a literature review
and analysis; (2) a Delphi survey focused on GI flows; and, (3) a focus group of experts
on the same issue. The primary private and public stakeholders have been involved in
the process from the onset.

This analysis has highlighted the following details:

– GI flows could have various forms according to the organisations involved (sale,
  loan, gift, donation, exchange, sharing and other);
– a significant number of GI flows are based on non-monetary transactions, and some
  of them are managed on an informal basis;
– for a few sub-sectors such a GI flows are not known at all;
– these GI flows are quite complex and their overall analysis and representation would
  require a complete survey involving all the stakeholders.

Figure 1: Global representation of GI flows within the geomatic sector in Quebec.
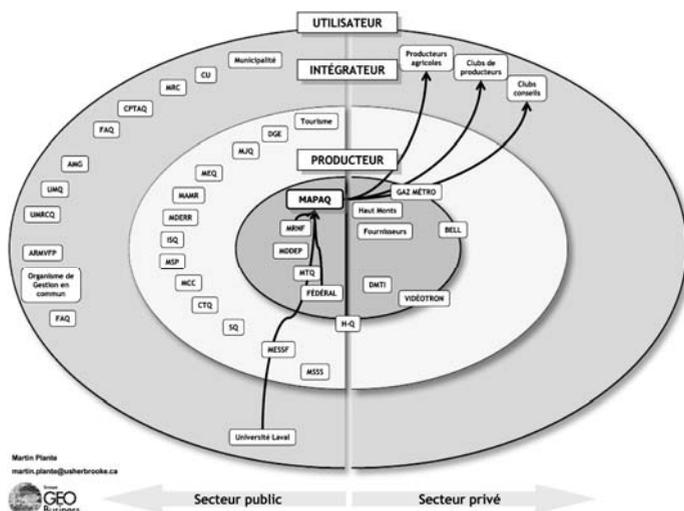
The complexity of work, which was required to produce such a representation, illustrated the eventual need of a systematic and computerised approach in order to provide an exhaustive and complete portrait of the sector. Therefore, a computerised dynamic model of GI flows was implemented.

With the goal of rendering the understanding and manipulation of the model easier, different versions were developed before implementing the latest version of the Socioscope prototype (Figure 1).

5.2. Concepts and principles of Socioscope

The modelling process of GI flows within the geomatic sector in Quebec is complex and the representation that has been provided could not be understood without an accurate analysis. The Socioscope prototype was developed, firstly, in the context of a master's thesis (Marion, 2005). It aimed to present the formal and informal links between people inside a single organisation, based on the concept of social network analysis.

Figure 2: Main interface of the prototype.
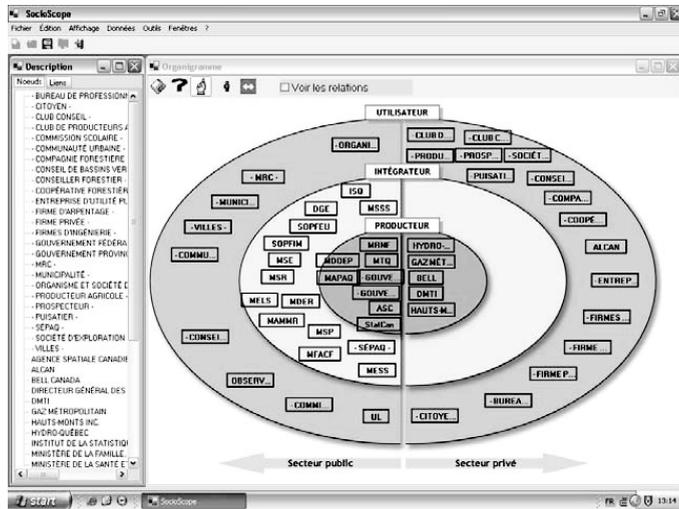
This first prototype was adapted for EcoGeo I purposes and applied to several organisations. A specific interface was developed and a database was built allowing the storage of relevant information about GI flows. The structure of the database also allowed to easily update information during the EcoGeo Phase II.

The main user interface of the prototype includes the organisations that were identified. If these organisations play more than one role in the geomatic sector, they were grouped by their primary role - data producer, solution integrator, and user (listed in Figure 1 and 2 with the corresponding French names: producteur, intégrateur, and utilisateur).

Organisations are located on the visual interface depending on their statute (private sector on the right, public sector on the left). The drop-down menu on the left extremity of the interface gives users direct access to the list of organisations that are graphically represented in the main windows.

Adding a new organisation in the procedure is straightforward. In order to perform that task, a contextual window allows the user to add relevant information about that organisation (Figure 3).

Figure 3: How to add a new organisation.
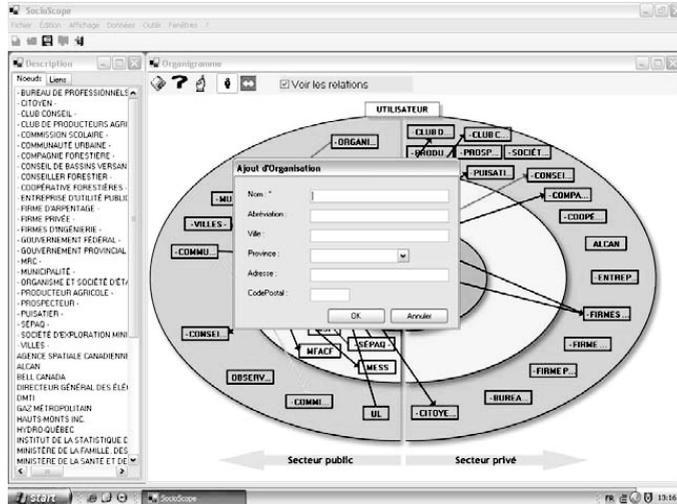
The main window also represents relationships between different organisations (Figure 4). These links could be single direction or double direction. Each relationship is characterised by a few attributes that qualify the GI flow.

Figure 4: Graphical representation of GI flows.
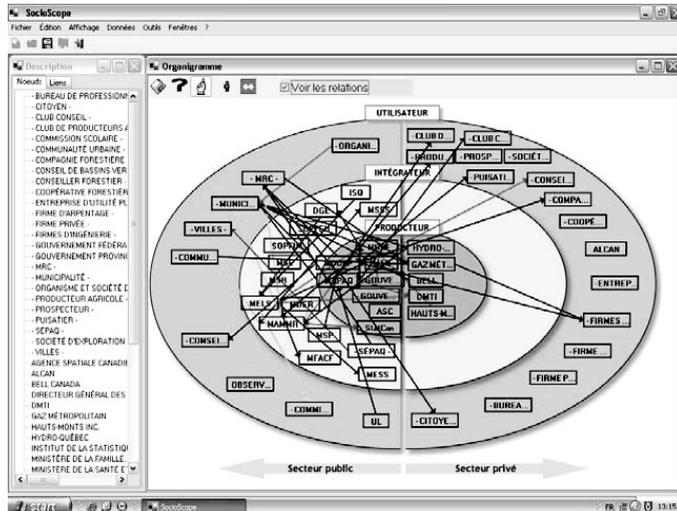
When a new flow between two organisations is added, the user may describe the GI flow characterisation - sale, loan, gift, donation, exchange, sharing and other - the direction of the flow to integrators or users, and other relevant details which can be added to the analysis.

Figure 5: Secondary interface: organisation and GI flows.
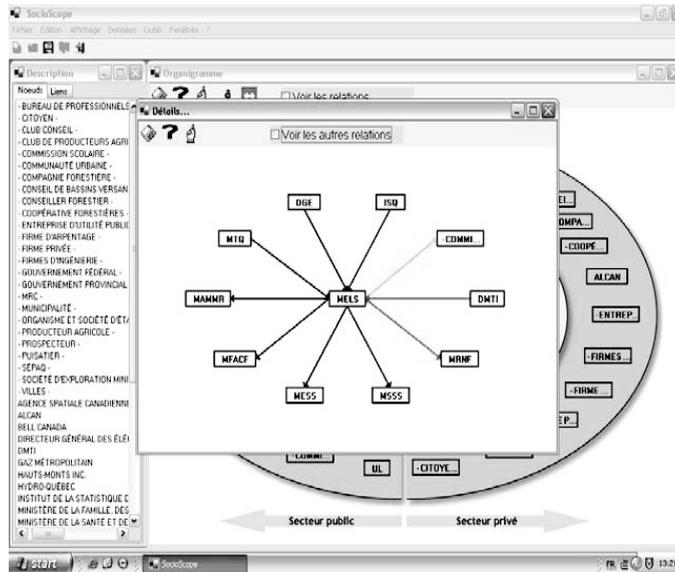
Clicking on a specific organisation provides access to a new pop-up window where other organisations having a relationship to the first are linked in terms of GI flows (Figure 5). Inside this new window, users may also navigate from one organisation to another, and also drill-down an organisation in which sub-organisations are available.

5.3. Applying the value chain concept to GI: perspectives

At the moment, the prototype can be used to create cartographical links, which exist between various stakeholders of the geomatic sector in Quebec, representing the movement of GI among the organisations. Socioscope could be implemented in a more complex system that would allow us to trace the steps within which the value is generated, from data acquisition and production, to sales, improvements, and distribution.

With the information gained by adding new attributes, defining the steps of the life cycle of GI products, determining where and how the products are exchanged, as well as when the value is created, it would be possible to support strategic decisions on the market. For example, investing in a geodetic network in a region would normally lead to a better GI capturing process. But what is the real value at the end of installing a geodetic point? Or we can imagine that someone, for his private economic activity, needs GI that can be acquired from an existing geographic dataset. Who on the market offers a dataset that has economic value for the potential buyer?

We affirm that Socioscope is a highly appropriate tool to enable easy access to the knowledge of GI flows among various organisations, and that it possesses remarkable potential to the future goals of EcoGeo II.

6. CONCLUSIONS AND FUTURE WORK

The next step of the project EcoGeo will be to define a value chain specific enough for the Quebec GI sector, encompassing an extensive range of issues including production, maintenance, distribution, and consumption of GI. The new focus is on developing and improving the Socioscope functions in order to reach the goal of implementing the value chain inside the prototype. In particular, the specific components necessary to fitting the value chain concept requirements need to be added.

A number of variables, which have been for the most part identified in literature (Krek 2003; Longhorn and Blakemore, 2008), will be inserted inside the chain. These attributes are connected to the way that GI is produced and used (value of the location attribute, time dependency, quality and others more), to the costs of the GI product (as the transaction costs), and to the price definition.

In section 3 we said that value is gradually created along the chain, so pricing in a value chain helps to determine how the value created for the final user is distributed among the contributors. It is fundamental to know the stakeholders involved in the value process (data producer and data owner, final users or the entire society) and understand their roles in defining the value of GI, and Socioscope helps us in this task.

Value is generated both for producers and customers of GI, thus it is essential to establish a common agreement on the value of GI products to assess the price and exchange value (Krek and Frank, 2000). The economic value of GI derives from its use in a decision making process and is measured by the improvement of the process.

The prototype differentiates the economic reality of the GI market for public and private sectors. This distinction is of primary importance, considering the differences in perception of the value chain by public sector GI owners (government agencies) who collect and use the GI for purposes relating to governments of society, involving a quantity of socio-economical uses and objectives for which the value is more difficult to evaluate (Didier, 1990; Craglia and Nowak, 2006; Longhorn and Blakemore, 2008; ACIL Tasman, 2008; Genovese et al., 2008). Therefore, while calculating value in the private sector means considering only the monetary/exchange value, the socio-economic one must also be considered when defining value in the public sector.

With the definition of a value chain for the Quebec GI sector, with modalities that are defined by EcoGeo II, it will be possible to follow, with some approximation, the generation of added value on a specific network of GI flows, starting from the original producer and ending with the final consumer. However, just the professional users (the data owners and vendors) are considered in this first phase of Socioscope. The users attributable to the democratisation process (final consumers or the society as whole) will be subsequently considered and their contribution in the value chain added in a further phase.

A formal economic analysis of the value of GI still does not exist in the literature (see Longhorn and Blakemore, 2008; Genovese et al., 2008). This is because of the extreme complexity of the GI sector and the number of variables connected to GI production and dissemination. The final goal of the EcoGeo project is the definition of a model of economical evaluation to be applied on a chosen subsector of the GI market in Quebec, determining the value added for the activities selected on the local value chain. The ability to assess economically the GI value will provide key support in strategic decisions and business efficiency, helping private companies to obtain a positive return

on investments. More importantly, it will improve citizens' quality of life and deliver more efficient government in the institutional sector.

REFERENCES

ACIL Tasman (2008). *The value of spatial information: the impact of modern spatial information technologies on the Australian economy*, Report prepared for the CRC for Spatial Information and ANZLIC, the Spatial Information Council, Australia.

Bergeron F., L. Gingras, P. Hadaya and C. Caron (2005). "A Framework for Evaluating Strategic Location-based Applications in Businesses", *HICSS 38 - Proceedings*, Hawaii.

Booz Allen Hamilton (2005). *Geospatial Interoperability Return on Investment Study*, Technical report, National Aeronautics and Space Administration - Geospatial Interoperability Office.

Caron C., S. Roche, M. Berthiaume and L. Gingras (2006). "The Strategic Impacts of Microgeomatics for Location Based Services in Business", *LI Conference 2006*, San Francisco, April 3-5.

Chrisman, N. (2005). Full circle: more than just social implications of GIS, *Cartographica*, 40(4): 23-35.

Cox, A. (1999). Power, Value and Supply Chain Management, *Supply Chain Management: An International Journal*, 4(4): 167-175.

Crompvoets, J. (2006). *National Spatial Data Clearinghouses, Worldwide Development and Impact*, PhD thesis, Wageningen University, Netherlands.

Crompvoets, J., A. Rajabifard, B. van Loenen, T. Delgado Fernández (2008). A Multi-View Framework to Assess Spatial Data Infrastructures, Melbourne, Australia: Digital Print Centre: The University of Melbourne, Australia, at: http://www.csdila. unimelb.edu.au/publication/books/mvfasdi.html.

Didier, M. (1990). *Utilité et valeur de l'information géographique*, Paris, France: Presses universitaires de France.

Craglia, M. and J. Nowak (European Commission, Directorate General Joint Research Centre) (2006). *Report of International Workshop on Spatial Data Infrastructures: Cost-Benefit / Return on Investment: Assessing the Impacts of Spatial Data Infrastructures*, Technical report, European Commission, Institute for Environment and Sustainability, Ispra, Italy.

European Commission (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

Gauthier, M. J. (1999). "La démocratisation des cartes ou les cartes pour tous et chacun", *Proceedings of the 19th International Cartographic Conference (ICC): Images du passé*, Vision d'avenir, 14-21 August 1999, Ottawa, Canada.

Genovese, E., G. Cotteret, S. Roche, C. Caron and R. Feick (2008). Evaluating the socio-economic impact of Geographic Information: A classification of the literature, *International Journal of Spatial Data Infrastructures Research*, (paper under review).

Goodchild, M. (1995). "Geographic Information Systems and Geographic Research", in J. Pickles (ed.). *Ground Truth: The Social Implications of Geographic Information Systems*, New York: The Guilford Press.

Grus, L., J. Crompvoets and A. Bregt (2007). Multi-view SDI assessment framework, *International Journal of Spatial Data Infrastructures Research*, 2: 33-53.

Halsing, D., K. Theissen and R. Bernknopf (2004). *A Cost-Benefit Analysis of The National Map*. Technical report, U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia.

Kaplinsky, R., and M. Morris (2001). *A Handbook for Value Chain Research,* Brighton : Institute of Development Studies.

Krek, A. and A. U. Frank (2000). The Production of Geographic Information - The Value Tree, *Geo-Informations-Systeme - Journal for Spatial Information and Decision Making,* 13(3): 10-12.

Krek, A. (2002). *An Agent-Based Model for Quantifying the Economic Value of Geographic Information,* PhD thesis, Institut for Geoinformation, Technical University Vienna.

Krek, A. (2003). "What are transaction costs and why do they matter?" *Proceedings of 6th AGILE Conference*, Lyon, April 2003, France.

Krek, A. (2004). "Requirements for an Efficient Value Pricing of Geoinformation", *Proceedings of 7th AGILE Conference*, Heraklion, Crete University Press.

Krek, A. (2006). "Geographic Information as an economic good", in Campagna M. (ed.). *GIS for sustainable development*, Taylor and Francis.

Longhorn, R. and Blakemore, M. (2008). *Geographic information: Value, pricing, production and consumption*, Boca Raton, FL: CRC Press.

Marion, G. (2005). Élaboration et expérimentation d'un outil visuel pour la gestion des aspects informels d'une organisation. Mémoire de M.Sc., Faculté d'administration, Université de Sherbrooke.

North, D. C. (1997). *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.

Noucher, M. and C. Archias (2007). "L'évaluation des Infrastructures de Données Spatiales - Application de divers cadres d'évaluation au CRIGE PACA", *Conference Geo-Evenement 2007*.

Oelschlager, F. (2004). Enterprise Information Integration; Enabling the Information Value Chain to Achieve Business Optimization, at: http://www.teamplay. primavera.com/partners/files/Enterprise_Information.pdf.

Onsrud, H. J. (1998). "A global survey of national spatial data infrastructure activities", at: http://www.spatial.maine.edu/~onsrud/gsdi/surveysum.htm.

Phillips, R. L. (2001). The management information value chain, *Perspectives*, 3.

Plante, M. (2006). *EcoGeo Phase I: Vers un portrait des retombées économiques dans le secteur de la géomatique au Québec*, Technical report, Université de Sherbrooke, Sherbrooke, Canada.

Roche, S. and C. Caron (2004). *Aspects organisationnels des SIG*, Paris: Hermès-Lavoisier.

Porter, M. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*, New York: Free Press.

Rodriguez-Pabon, O. (2005). *Cadre théorique pour l'évaluation des infrastructures d'information géospatiale*, Ph.D. Thesis, Département des Sciences Géomatiques, Faculté de Foresterie et de Géomatique, Laval University, Québec.

Samborski, B. (2007). *Building a Business Case for Shared Geospatial Data and Services: a practitioners guide to financial and Strategic Analysis for a Multiparticipant Program,* workbook of Geospatial Information Technology Association, U.S.A.

Samuelson, P. (1954). The Pure Theory of Public Expenditure, *Review of Economics and Statistics*, 36 (4): 387-389.

Shapiro, C, and H. Varian (1999). *Information Rules: A Strategic Guide to the Network Economy*, Boston: Harvard Business School.

Spataro, J. and B. Crow (2002). *A framework for understanding the information management market*. The Gilbane Report, at: http://gilbane.com/gilbane_report.pl/79/ .html.

Tulloch, D., E. Epstein, D. Moyer, B. Niemann, S. Ventura and R. Chenoweth (1998). "GIS and Society: A working paper", *UCGIS Annual Conference*, at: http://www. lic.wisc.edu/pubs/GIS_Society.pdf.

WEBSITES

Website 1: *http://www.quebecgeographique.gouv.qc.ca/*
Website 2: *http://www.lio.mnr.gov.on.ca*
Website 3: *http://www.geogratis.ca/*

Evaluation of Spatial Information Technology Applications for Mega City Management

Silke Boos and Hartmut Mueller

i3mainz, Institute for Spatial Information and Surveying Technology, FH Mainz
University of Applied Sciences, Germany
{boos, mueller}@geoinform.fh-mainz.de

Abstract

Administrations of mega cities face multiple challenges including informal settlements and air and water pollution control. Adequate mega city management requires a well-functioning spatial data infrastructure (SDI). The objective of this article is to evaluate the results of a comprehensive internet search concerning the use of spatial information technology in currently existing mega cities. The search starts from a nationwide view on the execution and the progress status of SDIs in the home countries of mega cities and zooms in to the specific aspects of spatial data management in the metropolitan areas of special interest. We conclude that current SDI development in mega cities covers the whole range from first stage conceptual ideas up to an almost complete operational SDI availability.

Keywords: Spatial data infrastructure (SDI), mega city, city management.

1. INTRODUCTION

City administrations of large cities, in particular of mega cities often are confronted with a multitude of key problems like informal settlements (land tenure, development approvals, building control), traffic management, natural hazards (floods, earthquakes, fires), unclear responsibilities and mandates (within or between administrations), uncoordinated planning, water management (fresh water supply and waste-water disposal), provision of continuous electrical power, visual pollution and garbage disposal, air and water pollution control (Kelly, 2008). To manage such problems adequate urban governance urgently needs comprehensive, reliable and easy accessible spatial data, in other words, a well-functioning spatial data infrastructure (SDI).

This article presents the results of an internet investigation, collecting information about current use of SDI in the world's largest metropolitan areas. A metropolitan area in this context is defined as an urban agglomeration with more than 10 million inhabitants (the mega cities), which applies to 26 cities in the world (http://www.citypopulation.de/world/Agglomerations.html). The following sections provide a short overview of general NSDI development for all countries of the world holding at least one mega city including the use of SDI or comparable initiatives in the associated metropolitan areas. After that, the results of the internet investigation are classified. Leaving legislative and organisational SDI aspects aside, the evaluation focuses on the technical aspects of the use of spatial information technology in mega city management. The classification is done on the basis of usability and accessibility of spatial data which was identified by the internet search.

2. APPLICATION OF SPATIAL INFORMATION TECHNOLOGY IN MEGACITIES AND THEIR HOME COUNTRIES

2.1 SDI application in the African region

NSDI in *Egypt* is still rudimental and has to deal with a number of bottlenecks such as weakness of partnerships, lack of digital data and metadata, absence of a clear institutional framework, shortage of access and sharing mechanisms to search for data, lack of national standards and a scarcity of qualified specialists (GSDI, 2007; Omran et al., 2006). Considering the underdeveloped NSDI of Egypt, it is no surprise, that for the city of *Cairo* no information concerning SDI development or comparable initiatives could be found.

Nigeria started the implementation of a National Geospatial Data Infrastructure (NGDI) in 2003 (Federal Ministry of Science and Technology of Nigeria, 2003). The policy statement to guide the operations of NGDI covers the following items:

− Facilitate cooperation and collaboration among stakeholders in generating spatial databases for development of an SDI at the National, State and local levels in Nigeria;
− Eliminate duplication in the acquisition and maintenance of spatial data;
− Establish institutional, legal, technical and administrative frameworks for:
  o a consistent and harmonised mechanism for spatial data distribution;
  o easy access to vital spatial datasets and their efficient sharing and exchange;
  o integration of datasets through the application of common standards;
− Promote investments in the production of spatial databases, and
− Promote research, training, education and capacity building related to spatial data production, management and usage.

In 2007, the government of *Lagos* constituted a committee for the provision of a fully digital mapping and enterprise GIS for Lagos State. The policy framework adopted by the administration for the development of Lagos should be reached by generation and sharing of information with organised private sector, developing skilled and knowledgeable workers. The mapping products that should be delivered as a result of this project should be at the scale of 1: 500 for metropolitan Lagos and at 1:1 000 for rural areas. Other scheduled products include: orthophotos (scale 1:2 000), contour lines (scale 1: 500 for urban and 1:1 000 for rural areas) and Digital Elevation Models.

2.2 SDI application in the Asian-Pacific region

In *Bangladesh* no official NSDI exists. SDI conform initiatives were initiated by the "Bangladesh Society of Geoinformatics" in 2006. Its mission is to build up capacity in geoinformatics within governmental and non-governmental agencies and to guide and assist the distribution of spatial information technology, sharing of ideas, information and knowledge among users, professionals and institutions. One of the objectives is to promote and assist establishment of the National Spatial Data Infrastructure (NSDI) in Bangladesh. In accordance with the rudimental national SDI initiatives in Bangladesh also in *Dhaka* neither city SDI nor any WebGIS application or similar was identified.

China has paid great attention to construct the Digital China Geospatial Framework (DCGF). This NSDI has four layers at National, Provincial, Municipal and County level. A series of fundamental spatial databases was completed as the kernel of DCGF. A

fully digital nationwide spatial data production system is widely established. The national coordinating mechanism is in action to strengthen the cooperation and data sharing and the national standards are getting more complete to support the DCGF (Li and Xuenian, 2008). In 2002, the *Shanghai* Municipal Government announced the "Digital City Shanghai" strategy. In this context a distributed WebGIS application for managing landscape resources was developed (Zhu et al., 2005), which allows the connection of all landscape bureaus of the city where data are kept locally for maintenance and updates. These data are also available online to the central bureau and other local bureaus. Beyond data exchange functions the GIS provides for spatial analysis functionality like distance-based spatial queries, for selection functions and for different types of buffering functions. In 2004, the city authority of *Guangzhou*, the capital city of south China, initiated the Digital Municipality of Guangzhou (DigiM.GZ) project (Cheng and Rao, 2006). The project aims to represent the Guangzhou metropolitan area as a digitalised virtual municipality by using a wide range of up-to-date GIS and telecommunications technologies. When in use, it shall provide for a universal platform to deal with all digital data relevant for city planning, management and maintenance, including water, gas and power supply, transport network, drainage and telecommunications. In *Beijing,* the Beijing Digital Green Management Information System is available, which consists of a GIS, remote sensing data, 3D virtual simulation, database, high-speed broadband networks and other hi-tech products. It integrates a database of Beijing landscaping areas and a database of social, economic, ecological and urban infrastructure. This system is constructed of components for integrated Management, system maintenance, dynamic garden inspecting, integrated query, planning, building maintenance, environmental benefits evaluation, 3D simulation, and other subsystems.

The NSDI scheme in *India* (established in 2001) aims at using GIS to merge satellite imagery and ancient topographic maps with data on water resources, flooding, rainfall, crop patterns, and civic layouts to produce 3-D digital maps. NSDI should, once ready, act as an online database to maintain spatial data layers and base maps in an easily retrievable from 40 major cities should be mapped at a scale of 1:1000, and in later phases the entire country should be covered. Another objective of the Indian NSDI is to achieve a national coverage of all forest maps, land use, groundwater and wasteland maps, pollution data, meteorological department's weather-info and department of ocean development's sea maps. The key elements for development of NSDI are: standards (to enable interoperability; standards for network, gateways, and protocols), evolving metadata, nodes (GIS-based spatial database servers), search and access protocols, electronic clearinghouse, creating user interfaces, and initiating an NSDI outreach and awareness program. For these purposes India has developed a geoportal. In 2005/06 in the Handni Chowk area of the walled city of *Delhi*, which covers an area of about 20 km² size, a pilot study on generating a 3D-GIS database was accomplished by the Department of Science and Technology and the Russian Academy of Sciences (RAS). The database was created by using a base map at scale 1:2500, high resolution satellite data, ground control points, video of the area, high resolution DEM from LiDAR/ ALTM and by 3D GIS data processing and analysis software (Kumarf 2007). In the future the database may be expanded for the entire city and provide for a basis for monitoring the city and for development of different applications for urban planning. In *Mumbai* various GIS applications for small areas with different aims have been made. The Mumbai Metropolitan Region Development Authority (MMRDA) recognised the usefulness of this technology and thus proposes in its Regional Plan (1996-2011) to build up a Regional Information System where the spatial and related attribute data should be organised and shared among the local authorities, planning agencies and other institutions working in the region. These developments may be stimulated by the Collective Research Initiative Trust (CRIT)' plans to generate an open-access SDI and

a set of simple tools and applications for knowledge transfer and participatory urban planning by communities and citizens in Mumbai. Until now the normal Internet user has only access to a demo version with some basic spatial data.

During the Survey and Mapping National Coordination Meeting in 2000, ISDI, the *Indonesian* SDI was declared to become a primary solution to solve the problems of the availability of and access to spatial data (Abdulharis et al., 2005). Bakosurtanal is the coordinating agency for the development of Indonesian NSDI (Arief Syafi'I, 2006). The NSDI aims at improvement of coordination mechanism, completion of spatial databases and national metadata developments, activation of national clearinghouse (Puntodewo and Nataprawira, 2007) and development of Digital Indonesia. Agency's spatial databases should be completed and should work within a nationally and globally integrated distributed system. A national clearinghouse prototype and a metadata gateway should be developed and metadata servers should be installed in key agencies. The city of *Jakarta* provides for a simple WebGIS application, which represents the road network of the city and enables different search functions to find streets and points of interest. No further SDI-activities in the city were identified.

In *Iran*, national organisations, ministry and municipal offices as well as private companies are active in the field of mapping and spatial data production. The national organisations concentrate their efforts on small-scale base mapping of the whole country. Governmental surveying offices and private companies are mostly involved in high resolution spatial data production needed for national and provincial projects (Baktash, 2003). Most research in the fields of photogrammetry, remote sensing, GIS and digital mapping is carried out in the national organisations, institutions and universities (Rad and Sarpoulaki, 2004). However, a few private companies also made remarkable research efforts for commercial products and services. The *Tehran* municipality, Public & International Relations Department committed to the development of a WebGIS with more than 140 layers, which should be launched before the end of the current Iranian year. The application should serve citizens and managers of various organisations and institutions as well as domestic and foreign tourists with needed information.

In *Japan,* the NSDI is implemented by the Geographical Survey Institute (GSI) and different ministries, who began their work on the Spatial Data Framework in 1995 and completed it in 2003. Over the period of development the institutions produced a collection of base maps, notably the topographical map series of 1:25 000, which covers the whole country. Those maps were used for generating several public and private sector maps like administration area maps, road maps and several thematic maps (Land Use Map, Land Condition Map, Volcanic Land Condition Map, and a Map of Active Faults in Urban Area). Beyond these maps also aerial photographs were published and the development of a national standard was established. The future work of the Japanese NSDI concentrates on a new infrastructure concept, which is promoted as "Digital Japan" and which shall lead to a virtual and real-time representation of the land realised by integrating geographic information of various kind and which shall be made accessible to anyone on the internet. Concerning the two Japanese mega cities *Osaka* and *Tokyo,* the internet investigation could not extract any specific SDI-initiatives. Both cities developed long-term master plans, where principal goals for city planning are formulated but no SDI strategy could be identified.

In *Pakistan* no official NSDI was established. Only some SDI-supporting-initiatives exist (Asmat, 2008), one of which was the Winner of GSDI Association Small Grant 2006-7. Under the aegis of the WWF this initiative develops an SDI for sharing environmental information. From the inception of the project large amounts of spatial data including

satellite imagery, digital vector data, and digital terrain models were acquired and developed. In its "Megacities Preparation Project" from 2005 (TA 4578, 2005). *Karachi's* government schedules the development of digital maps of the city by using GIS technologies. Yet this project is not finalised (Khan, 2007).

First official activities for establishing an NSDI in *Philippines* were initiated in 2001 under the umbrella of the National Geographic Information Council (NGIC) (Crisostomo, 2007). The central mapping agency of the government of the Philippines (NAMRIA) keeps all base maps such as topographic maps in different scales, aerial photographs and satellite images. NAMRIA also produces different thematic maps such as for land condition, land cover, land use, planimetric and administrative maps. As a member of a developing country Metro *Manila* has not yet a comprehensive SDI available. A Disaster Management Information System called "Metro Manila Map Viewer" was developed in 2004, which allows users to retrieve useful information and maps from datasets including hazards, transportation, public facilities, emergency services, elevation, land use/zoning, and high-resolution imagery.

The first phase of an NSDI Master Plan for *South Korea* was completed in 2000. The main purpose of the first phase was to establish basic GIS infrastructure by producing various kind of digital maps. The second phase of the NSDI, which started in 2001, concentrated on spreading GIS application for maintaining the digital maps and developing national standards (Han and Cho, 2001). The city of *Seoul* has at its disposal a widespread SDI on the technical base of several distributed GIS applications like Urban Planning Information System, Road Information System, Soil Information System, and other municipal affairs Information Systems. A Spatial Data Warehouse is available which provides for sharing and accessing the different spatial data of the GIS systems via a GIS Portal system (Choi et al., 2006). A map viewer program even allows analyses of the retrieved data.

In 2004, a feasibility study on NSDI was initiated by Geo-Informatics and Space Technology Development Agency (GISTDA) with grant support from the U.S. Trade and Development Agency (USTDA) for *Thailand*. The study could show various problems particularly concerning data sharing and data usage. Development of NSDI fits in very well with the Thai Government's scheme on a comprehensive utilisation of Information Technologies to support administration and public services. The key mechanism is the development of e-Government in which GIS forms a key component which plays an important role in providing for dynamic information to support better governance of the country (Silapathong, 2004). Collections of spatial data are available from the Royal Thai Survey, which provides for data in analogue and digital format (information available only in Thai language). For the city of *Bangkok* only a webpage in Thai language was found. This webpage seems to grant access to a comprehensive collection of spatial data in different GIS applications.

2.3 SDI application in the European region

As Francois Salgé states "*France* is creating an NSDI without knowing it. Thus NSDI is not per se an issue in the French context" (EC INSPIRE, 2006). Consequently there is no explicit overall governmental initiative to develop an NSDI in France even though a geoportal was launched in 2006 and a multitude of NSDI-like initiatives are undertaken. In *Paris* a WebGIS application gives access to the most important spatial data about the city. It is possible to access a series of thematic maps through a multiplicity of data layers.

Russia is just at the beginning in developing an NSDI. The concept dating from 2006 schedules a three stage process, which should be finalised by 2015 with the implementation of the national NSDI. The concept shall be transferred into a distributed system for collecting, processing, storage and delivery of basic spatial data and metadata. The system shall comprise subsystem levels of government and local governments and shall users grant remote access to digital databases of spatial data and metadata. For the city of *Moscow* no specific SDI solution information could be found during the internet investigation.

Currently, the Military Mapping Agency of *Turkey* is the main data producer of spatial data and has the most visible internet presence offering limited metadata for its own products. There are several persisting problems in the field of SDI in Turkey: lack of coordination between institutions; no standardisation, neither with regard to the spatial reference system, nor to data quality or data exchange; data duplication; the majority of large scale data not available in digital format; interoperability does not (yet) exist; lack of expert personnel and budget; and a lot of difficulties to share data (EC-INSPIRE, 2006). *Istanbul's* Water and Sewerage Administration (ISKI) developed the Infrastructure Information System (ISKABIS) to control and manage extensive water and wastewater facilities for the Istanbul Metropolitan Area. The system is based on a file server system application to achieve effective data sharing. Within the file server system various folders like maps, raster, infrastructure, superstructure, planning projects etc. are categorised in a similar way as a digital library. Each department in ISKI, such as mapping department, GIS department, Water Project department, Sewerage Project department, has to update exclusively the folder which it is responsible for. More than 30 applications are implemented in ISKABIS CAD/GIS program. Ultimate Map Management, Infrastructure Management, Projects Management, Address Query, Building Query, Cadastral Query, Geographical Information System Applications, and Easy Print Utility can be made via ISKABIS. The city administration of Istanbul provides for a WebGIS, which represents the road network for the metropolitan area of Istanbul containing a precise division into lots and house numbers, orthofotos of different years and a range of thematic information, as well.

Although in 1995 the National Geospatial Data Framework (NGDF) initiative was launched, there is yet no formal NSDI in the *United Kingdom*, or a single organisation with responsibility for its establishment and coordination. On the other hand, the country as a whole has a well developed GI sector, with extensive datasets available from both public and private sector sources (McLaren and Mahoney, 2000). Various efforts have been undertaken to implement a broad metadata service but these have not been sustainable. The government of the city of *London* provides for the City Online Maps Project Accessing Spatial Systems (COMPASS), which aims at improving access to information about the city of London through a unique access point so that residents and those visiting the city are better informed. A wide range of data is available on the site such as where to find your nearest services and information about planning policies affecting the city. One remarkable SDI conform application in London is the Newham Neighbourhood Information Management System (NIMS), where users gain access to data on economic, social and environmental conditions of the borough. Maps, charts, data download is available, as well as is generating of online reports and performance information.

2.4 SDI application in the Pan-American region

In 1998 the first activities concerning NSDI were initiated in the federal republic of *Argentina* by the SIGRA group (Geographic Information System of the Argentine Repub-

lic) and the National Mapping Agency (IGM) leading to the NSDI implementation in 2001. In 2004 the National Geographic Information System of the Republic of Argentina (PROSIGA) started as an Internet distributed GIS, in which seven specific SDI working groups are present: Institutional framework, Policy and Agreements, Fundamental and Basic Data, Metadata and Catalogues, Diffusion and Communication, Training, Search Engine for Geographic Names and IT for SDIs (Machuca and Rickert, 2008). The department of Geographic Information Systems of the city administration of *Buenos Aires* developed a widespread WebGIS application built up on open source components and integrating a multiplicity of spatial data of the city. The GIS covers a range of applications like health, education, tourism, sports, culture, leisure, green spaces, social services, transportation etc. and enables access to information up to parcel units (it is possible to view for most of the parcels a photograph showing the parcel-related buildings). The department also provides for thematic maps, which are based upon the GIS data and can be ordered in digital or analogue format.

In *Brazil* the Ministry of Budget Planning and Management is responsible for the Brazilian NSDI, with strong participation of the Brazilian Institute of Geography and Statistics (IBGE) and the National Institute of Space Research (INPE). The Brazilian cartographic community, in particular Federal Government agencies, made great efforts to constitute an NSDI in Brazil (Camara et al., 2006). The IBGE launched map servers offering diverse information and providing for geodata of the whole country. The department for planning of the city of *Sao Paulo* makes an internet portal available, which enables access to a multiplicity of statistical data, thematic maps and allows for the visualisation of infrastructural data in a WebGIS client. For *Rio de Janeiro* the department of city planning offers digital maps and databases of the municipality of Rio in a geoportal and allows for download of statistical tables, maps and spatial data.

Mexico's NSDI initiative is called the "Infraestructura de Datos Espaciales de México" or IDEMEX (Hyman et al., 2002). The Mexican NSDI implementation is led by the National Institute of Geography, Statistics and Informatics (INEGI) since 1997 (Albites, 2008). INEGI developed an internet presence (geoPortal), where users can view and download a series of spatial data, including appropriate metadata (Ramírez 2005). The Interactive Atlas Nacional de Mexico (ANIM) on this website shows in an exemplary way the provision of public information. The user is capable of viewing geographical information from various sources through a single interface. For the Mexican mega city *Mexico City* the internet investigation did not extract any specific SDI-like-initiative.

The *United States* clearinghouse was established in 1994 with the US Federal Geographic Data Committee (FGDC) responsibility of NSDI implementation (Clinton, 1994). In 2004 still the NSDI major development focus was almost completely restricted to the United States federal level (Steven 2005). Spatial data are provided in a nationwide geoportal offering a multiplicity of functions to access, publish and share spatial data in a widespread number of categories. Concerning city SDI initiatives, in 2008 the *New York City* government has published its IT strategy for the next years (NYC PlanIT). The strategic plan describes a framework for how the City will leverage general information technology in the years ahead to improve New Yorkers' lives. The plan discusses the utilisation of spatial data. An Interactive City Map of New York provides information on transportation, education, public safety, resident service and city life. The office of Emergency Management operates a GIS, which maps and accesses data - from flood zones and local infrastructure to population density and blocked roads - before, during, and after an emergency case. Beyond that the City government runs a spatially-enabled public website called ACCESS NYC, which has the capability to identify and to screen for over 30 City, State, and Federal human service benefit programs

to explore appropriate services for the individual users needs. The *Los Angeles* government publishes a collection of interactive maps containing information on traffic, parcels, flooding, city services, leisure, among other information.

3. ANALYSIS OF SEARCH RESULTS

3.1 Valuation method for results classification

From the internet investigation a wide range of different development stages of spatial data handling in the examined countries and their associated mega cites emerged. Obviously this is determined by large deviations in terms of social, economic and political conditions of different countries and cities. The most striking fact is that many larger cities of the southern hemisphere are located in developing countries where obviously the most unfavourable conditions are given which hinder the building of a working SDI. In the given inhomogeneous context the definition of global comparison criteria seems to be difficult. Thus, formal criteria were defined to set an objective evaluation framework. The main focus of the evaluation concentrates on the technical part of spatial data processing while omitting the institutional and legislative SDI aspects. The evaluation framework consists of five categories which are designed to classify all investigated items. The purpose of the used classification is to reflect the requirements to be met by a working SDI. That is why the items of content planning, of content provision and of content accessibility are addressed separately. The list of items not only addresses the mega cities themselves but also their home countries, because a city is part of a country and, therefore, is assumed to be part of the NSDI of its home country, as well. The five categories of the evaluation framework were defined as described in detail in the following section.

If, for whatever reason, little information on an item could be found on the web, the corresponding item was marked with 'SDI development status unknown'. If initial activities towards SDI development were observed the status 'SDI master plan available' was given. Further definition of the classification schema differentiates primary from secondary spatial data. Primary spatial data are original data, like survey data, data with limited need of interpretation like water bodies or boundaries, which are obtained without analysis or less interpretation. Secondary data are thematic data which are derived from the analysis of primary data, statistical data collection and/or image interpretation. This differentiation is in accordance with the GSDI Cookbook (Nebert, 2004), with the guidelines of the European INSPIRE initiative and with the Australian Spatial Data Infrastructure (http://www.anzlic.org.au/policies.html) which all define primary data in terms of 'Fundamental Data' or similarly 'Global -', 'National -', 'Framework -', 'Base -' , 'Reference -', and 'Core Data'. Even if the requirements concerning spatial information are considerably different at national and urban level, the overall differentiation in 'Primary spatial data available' and 'Secondary spatial data available' provides for a common basis for classification. Another important finding of the internet investigation was the fact that the process of SDI development in many of the searched countries and cities currently is in the stage of digital data production. However, the captured data often are not yet available via a Geoportal or a similar distributed web application. To reflect this finding the classification schema differentiates between spatial data availability and 'Spatial data accessible'. Therefore, the final classification schema consists of five categories:

– SDI development status unknown;
– SDI master plan available;
– Primary spatial data available;

– Secondary spatial data available, and
– Spatial data accessible.

3.2 Application of spatial information technology in the home countries of mega cities

The progress of spatial data handling in the home countries of mega cities in terms of the internet investigation results shows a large diversity (Table 1). Some countries like Russia or Nigeria are just at the beginning of developing an NSDI, while other countries are at the stage of producing primary data (e.g. Iran, Pakistan) and secondary data (e.g. China, Japan). It also can be shown, that the progress in developing an NSDI is well-advanced in Europe and Pan-America and India, where users already have access to spatial data via distributed web applications.

Table 1: Application of SDI in the home countries of mega cities.

	SDI development status unknown	SDI master plan available	Primary spatial data available	Secondary spatial data available	Spatial data accessibility available
Argentina					●
Bangladesh		●			
Brazil					●
China				●	
Egypt	●				
France					●
Indonesia		●			
India					●
Iran			●		
Japan				●	
South Korea		●			
Mexico					●
Nigeria		●			
Pakistan			●		
Philippines			●		
Russia		●			
Thailand			●		
Turkey			●		
United Kingdom					●
United States					●

3.3 Application of spatial information technology in the world's mega cities

The internet investigation of the status of SDI in the mega cities proved to be more difficult than for the counties hosting mega cities because less publication does exist concerning the related items. In both cases, sometimes the analysis has to rely on older publications found on the web which, although describing the planned SDI activities at the date of publication may not necessarily represent the most current development status. Moreover, some of the cities only provide information in their national language, which, due to lack of language ability of the authors, in some cases could not be analysed. Particularly with regard to the search results obtained for Japan we can not exclude that relevant information is available elsewhere. A general restriction concerns the fact that there is no guarantee that information not covered by the search is provided by any smaller local organisations.

Nevertheless it can be stated that, like in the home countries of the mega cities, the application of spatial information technology in the mega cities of the world is largely diverse. Table 2 shows the availability of digital spatial data in the mega cities under review. The application of spatial information technology in the cities under consideration varies considerably. It starts from the provision of simple WebGIS applications which only show the road network and some less basic data like in Jakarta or Mumbai, it comprises advanced applications which enable the presentation of social, economic, ecological and urban information related to the city (e.g., Buenos Aires, Los Angeles, Paris) and it ends up with highly elaborated comprehensive distributed information systems which can be found in Seoul, London and New York City.

Table 2: Application of SDI in the world's mega cities.

	SDI development status unknown	SDI master plan available	Primary spatial data available	Secondary spatial data available	Spatial data accessibility available
Bangkok	●				
Beijing					●
Buenos Aires				●	
Cairo	●				
Delhi			●		
Dhaka	●				
Guangzhou				●	
Istanbul					●
Jakarta			●		
Karachi		●			
Lagos		●			
London					●
Los Angeles				●	
Manila				●	
Mexico City	●				
Moscow	●				
Mumbai			●		
New York					●
Osaka	●				
Paris				●	
Rio de Janeiro					●
Sao Paulo					●
Seoul					●
Shanghai					●
Tehran		●			
Tokyo	●				

3.4 Synoptic view of Spatial Information Technology applications in the world's mega cities and their home countries

Figure 1 shows that in many cases a correlation between NSDI development and urban SDI development exists. The more advanced is the NSDI of a country, the more advanced is the SDI of its cities. In terms of the defined classification categories it can be seen that in many cases the national SDI development of the countries is one step ahead the SDI development of its largest cities.

Figure 1: Technical progress of SDI development in the world's mega cities and their home countries.
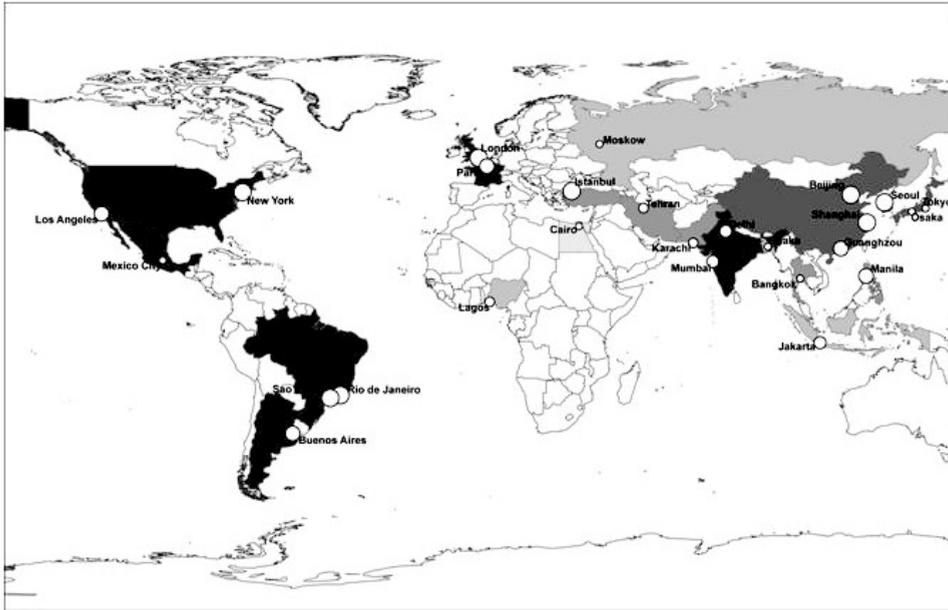
Technical progress of SDI (2008)
Home countries and related mega cities

Megacity		Homecountry	
○	SDI development status unknown		SDI development status unknown
○	SDI master plan available		SDI master plan available
○	Primary spatial data available		Primary spatial data available
○	Secondary spatial data available		Secondary spatial data available
○	Spatial data accessibility available		Spatial data accessibility available
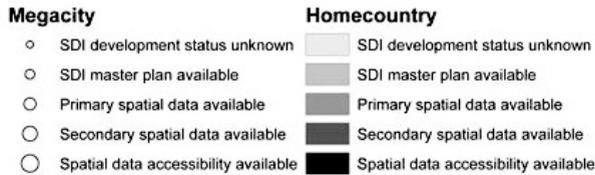
4. CONCLUSIONS AND FURTHER WORK

The investigation of current applications of state-of-the-art SDI technology in the world's existing mega cities including NSDI development in their home countries shows a large diversity in terms of stage of development. Whilst for some countries and cities almost no useful information was retrieved from the web others are in the conceptual phase of SDI development. Often NSDI progress and urban SDI development is correlated. In some regions primary and secondary data production is in progress. The most advanced SDI implementations are to be found in some countries and cities where web based services for broad access to comprehensive distributed spatial data pools are already in operation. In any case, the technology oriented approach of this study should be completed by other investigations which are to explore the organisational and legislative aspects of SDI implementation including their interaction with planning and other management activities in mega cities.

REFERENCES

Abdulharis, R., van Loenen, B., Zevenbergen, J. (2005). "Legal Aspects of Access to Geo-Information within Indonesian Spatial Data Infrastructure", *ISPRS Workshop on Service and Application of Spatial Data Infrastructure*, XXXVI (4/W6), Hangzhou, China.

Albites, F.A.H. (2008). "Relationships of Cartography, Geographic Information Systems and the Mexican Spatial Data Infrastructure (IDEMEX)", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* XXXVII. Part B4. Beijing.

Arief Syafi'I, M. (2006). "The Integration of Land and Marine Spatial Data Set as part of Indonesian Spatial Data Infrastructure Development", *Seventeenth United Nations Regional Cartographic Conference for Asia and the Pacific.* Bangkok.

Asmat, A. (2008). Potential of Public Private Partnership for NSDI implementation in Pakistan. Master-Thesis, Enschede.

Baktash, P. (2003). "National and Regional Spatial Data Infrastructure (NSDI & RSDI) and National Cartographic Center of Iran's Activities about it", *Proceedings of the 2nd FIG Regional Conference*, Marrakech, Morocco.

Omran, E.S.E., Crompvoets, J., and Bregt, A. (2006). Benefits and bottlenecks of SDI development in Egypt. *GIS@development*, 2(1): 32-35.

Camara, G., Fonseca, F., Moneiro, A.M., Onsrud, H. (2006). Networks of Innovation and the Establishment of a Spatial Data Infrastructure in Brazil, *Information Technology for Development*, 12(4): 255-272.

Cheng, P., Rao, J. (2006). "Digim.GZ -- The Digital Municipality of Guangzhou Project", *Proceedings of Map Asia 2006*, Bangkok, Thailand.

Choi, Y, Han J.-H., Lee, I. (2006). "The Spatial Data Warehouse of Seoul", *Proceedings of the ESRI user conference 2006*, at: http://gis.esri.com/library/userconf/ proc06/papers/papers/pap_2000.pdf.

Clinton, William (1994). Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure. Executive Order 12906, *Federal Register*, 59 (71), pp.17671-17674.

Crisostomo B.A. (2007). "The Philippines Experience in the Development of a National Spatial Data Infrastructure (1993 – 2007)", *Workshop on Spatial Data Infrastructure"hosted by AIT & SCOSA*, Thailand.

EC-INSPIRE (2006). Spatial Data Infrastructures in France: State of play Autumn 2006. Spatial Application Division, K.U.Leuven Research & Development, Leuven.

Federal Ministry of Science and Technology of Nigeria (2003). National Geoinformation Policy (Draft), at: http://www.rectas.org/NigeriaGIpolicy.pdf.

GSDI (2007). GSDI Small Grant Summaries 2006-2007, at: http://www.gsdi.org/ docs2007/GSDISmGrtSums06-07.pdf.

Han, S., Cho, H. (2001). "Development of National Spatial Data Infrastructure in Korea". *Proceedings of the Technical Conference during the FIG Working Week*. Seoul, South Korea.

Hyman, G., Lance, K., Perea, C., Rey, D. I. (2002). "Survey of the Development of National Spatial Data Infrastructures in Latin America an the Caribbean", *Proceedings of the ESRI user conference 2002*, at: http://gis.esri.com/library/userconf/ proc03/p1140.pdf.

Khan, J. (2007). Karachi Master Plan 2020: Only 15 GIS town maps completed as plan constantly delayed. *Daily Times*, 25 September, at: http://www.dailytimes.com.pk/default.asp?page=2007\09\25\story_25-9-2007_pg12_3.

Kelly, P. (2008). Current Problems and Issues in Using SDI in Mega cities. Report of FIG Working Group 3.2: Spatial Data Infrastructures in Mega cities, unpublished.

Kumar, R.S. (2007). 3D Mapping of Walled City, *GIS@development.net*, 11(4).

Li, P., Lan W., Xuenian X. (2008). "SDI in China: Progress and issues", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII. Part B4. Beijing.

Machuca, J.H., Rickert, J.E. (2008). "Argentina, Geographic Information System on the Internet (PROSIGA)", *Proceedings of the GSDI-10 Conference*. Trinidad, St. Augustine.

McLaren, R., Mahoney, R. (2000). "NSDI in the UK", *Proceedings of the Quo Vadis-International Conference FIG Working Week 2000*, Prague, at: http://www.fig.net/pub/proceedings/prague-final-papers/mclaren.htm.

Nebert, D.D. (2004). GSDI Cookbook v. 2.0, at: http://gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf.

Puntodewo, Nataprawira, R. (2007). "Indonesian Geospatial Data Clearinghouse", *Proceedings of the 3rd FIG Regional Conference*, Jakarta, Indonesia.

Rad, A.E., Sarpoulaki, M. (2004). "Islamic Republic of Iran National Report for Photogrammetry and Remote Sensing 2000-2004", *Proceedings of ISPRS Commission VI, ISPRS Congress,* Istanbul.

Ramírez, J.O. (2005). "The Spatial Data Infrastructure of México "IDEMex". From Pharaohs to Geoinformatics", *Proceedings of FIG Working Week 2005 and GSDI-8*, Cairo, Egypt.

Silapathong, C. (2004). "Development and Implementation of NSDI in Thailand", *Conference Proceedings of Map Asia 2004*, Beijing, China.

Stevens, A.R. (2005). "The US National Spatial Data Infrastructure: What is new?" *ISPRS Workshop on Service and Application of Spatial Data Infrastructure*, XXXVI(4/W6), Hangzhou, China.

TA 4578 – Pakistan: Karachi Mega Cities Preparation Project (2005). Final Report, at: http://www.adb.org/documents/produced-under-ta/38405/38405-pak-dpta.pdf.

Zhu, Y., Yang, C., Wong, D.W., Kafatos, M. (2005). A Distributed GIS for Managing Shanghai Landscape Resources. *Geographic Information Sciences*, 11(1): 29-39.

INVESTIGATED WEBSITES

Argentina
http://www.sig.gov.ar/
http://www.gsdi.org/newsletters/sdilacv4n12english.pdf
Buenos Aires
http://mapa.buenosaires.gov.ar/sig/index.phtml

Bangladesh
http://www.bsgi-bd.org/Index.html

Brazil
http://www.gisdevelopment.net/policy/gii/gii0024.htm
http://www.geominas.mg.gov.br/
http://mapas.ibge.gov.br/
Sao Paulo, Rio de Janeiro
http://sempla.prefeitura.sp.gov.br/mapasedados.php
http://www.armazemdedados.rio.rj.gov.br/

China
http://www.sbsm.gov.cn/pcgiap/tsukuba/seminar/paper_cn.pdf
http://sedac.ciesin.org/china/
Guangzhou, Beijing
http://www.otitan.com/info/20071219/20071219135109.shtml

France
http://www.geoportail.fr/
Paris
http://paris-a-la-carte-version-pl.paris.fr/carto/mapping

India
http://gisserver.nic.in/nsdiportal/gotogos.jsp
http://www.gisdevelopment.net/news/viewn.asp?id=GIS:N_npqesowhvy
Delhi, Mumbai
http://www.gisdevelopment.net/news/viewn.asp?id=GIS:N_xplqdkajin
http://www.gisdevelopment.net/magazine/years/2007/april/32_1.htm
http://www.mmrdamumbai.org/planning_information.htm
http://crit.org.in/category/mapping/
http://mumbai.freemap.in/

Indonesia
Jakarta
http://map.yellowpages.co.id/Default.aspx

Iran
Teheran
http://www.tehran.ir/Default.aspx?tabid=5215&ctl=Details&mid=22486&ItemID=23030

Japan
http://www.nsdipa.gr.jp/english/pof.html
http://www.geoinfo.ait.ac.th/download/SCOSA2007/2_MrKawase/NSDI.pdf
Osaka, Tokio
http://www.city.osaka.jp/english/more_about_osaka/city_concept/index.html
http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN015060.pdf

Mexico
http://www.inegi.gob.mx/inegi/default.aspx

Nigeria
Lagos
http://www.tundefashola.com/archives/news/2008/07/15/20080715N01.html

Philippines
http://www.geoinfo.ait.ac.th/download/SCOSA2007/4_SDI_Activities_by_countries/6_P
hilippines.pdf
http://www.geom.unimelb.edu.au/research/SDI_research/Integrated/Int_Template_Phili
ppines.pdf
http://www.namria.gov.ph/home.asp

Manila
http://www.pdc.org/mmeirs/html/mmeirs-home.jsp

Russia
Moscow
http://www.gisa.ru/

South Korea
Seoul
http://english.seoul.go.kr/government/ICSFiles/afieldfile/2005/03/25/GIS.pdf

Thailand
http://www.gisdevelopment.net/policy/international/ma04013pf.htm
http://www.rtsd.mi.th/service/
Bangkok
http://www.bangkokgis.com/

Turkey
Istanbul
http://iskabis.iski.gov.tr/ENGLISH/be_awards2007/awards.html
http://sehirrehberi.ibb.gov.tr/MapForm.aspx?&rw=1E7&cl=4F8

United Kingdom
http://www.dnf.org/Pages/about%20dnf/
http://www.ordnancesurvey.co.uk/oswebsite/
http://www.planningportal.gov.uk/england/government/en/
http://www.gigateway.org.uk/
London
http://www.cityoflondon.gov.uk/Corporation/maps/Interactive+City+maps.htm
http://www.newham.info/iads/

United States of America
http://gos2.geodata.gov/wps/portal/gos
http://www.fgdc.gov/nsdi/nsdi.html
New York, Los Angeles
http://www.nyc.gov/html/doitt/downloads/pdf/nyc-plan-it-08.pdf
https://a858-ihss.nyc.gov/ihss1/en_US/IHSS_languageSelectionPage.do
http://gis.nyc.gov/doitt/cm/CityMap.htm
http://www.lacity.org/lacity197.htm

Local Government and SDI – Understanding their Capacity to Share Data

Kevin McDougall[1], Abbas Rajabifard[2] and Ian Williamson[2]

[1]University of Southern Queensland, Australian Centre for Sustainable Catchments
mcdougak@usq.edu.au
[2]University of Melbourne, {abbas.r, ianpw}@unimelb.edu.au

Abstract

Local government has been recognised as an early leader in the development, deployment and innovation in spatial information systems. The introduction of corporate wide spatial data portals within local government was as significant as the release of Google Earth to the wider public. Although these information systems continue to expand and mature, the potential for these local spatial data infrastructures (SDIs) to contribute to higher level SDI initiatives remain largely unrealised. This article explores local government SDI within Australia to assess its capacity to contribute to higher level SDI initiatives. A comprehensive survey of over 100 local government authorities was undertaken to assess their SDI capacity and collaborative initiatives. The results were analysed to identify factors that contribute to their successful local SDI development and, more widely, to the development of higher level SDI initiatives through data sharing partnerships. The findings from the analysis indicate that suitable policy frameworks, an understanding of business needs, organisational support and ability to access data through equitable sharing arrangements are critical drivers in building and developing SDI from the local level.

Keywords: local government, spatial data infrastructure (SDI), data sharing.

1. INTRODUCTION

The exchange of fundamental spatial data between local and state jurisdictions continues to be problematic for a variety of technical, institutional, political and economic reasons (Harvey and Tulloch 2006; McDougall et al., 2005; Nedovic-Budic and Pinto, 2000; Onsrud and Rushton, 1995; Pinto and Onsrud, 1995). This impacts on the development of spatial data infrastructures, particularly at local and state levels, and hence the efficient delivery of government and community services (McDougall, 2006; McDougall et al., 2002; Warnecke et al., 2003). It is recognised that local-state government SDI environments are critical because it is within these environments where the most useful operational spatial data resides.

At the country or national level some progress has been made in describing SDI development (Masser, 1999) and spatial data clearinghouses (Crompvoets et al., 2004), however it is difficult to translate the outcomes of these studies to a local level. Some efforts have been made to understand the Australian SDI environments (McDougall, 2006; Warnest, 2005), particularly with respect to the sharing of spatial data and the models of collaboration between Australian jurisdictions.

The technical issues of data integration and interoperability are progressively being advanced (Abel et al.,1999; Dangermond and Brown, 2003), however it is the organisational, legal and economic issues that continue to impede the integration of spatial data in heterogeneous data sharing environments (Masser, 1998; Masser and Campbell, 1994; Nedovic-Budic and Pinto, 2001; Onsrud and Rushton, 1995). In particular, the

vertical integration of multiple levels of data across multiple levels of government is recognised as a major impediment to a fully robust National Spatial Data Infrastructure (NSDI) (Harvey et al., 1999). Masser (2005) has identified that the vertical integration of data is not well understood and that greater efforts are needed to explore the nature of spatial data sharing and its effectiveness in a multilevel SDI environment, particularly with respect to the organisational issues.

Our knowledge of SDI frameworks has come from the first generation of SDIs which emerged from national mapping and land administration authorities in the mid 1990s. Countries that developed the first generation of SDIs had a limited knowledge of the different dimensions and issues relating to the SDI concept (Rajabifard et al., 2006). The major objectives of these initiatives were to promote economic development, to stimulate better government and to foster sustainable development (Masser, 1998). So, like the national information infrastructure visions espoused by governments in the early to mid 1990s, SDI has developed in all shapes and sizes (Masser, 1999) and is viewed differently by different stakeholders. In recent years we have seen the second generation of SDIs emerge. Craglia and Signoretta (2000) identified in their case studies of local municipalities that because of the heterogenous nature of this level of government, framework models to describe SDI cannot be easily replicated.

Increasingly, partnerships are considered essential for SDI development because they provide the mechanism to allow organisations to work together to achieve SDI goals and to share the implementation responsibilities and eventual partnership benefits (Wehn de Montalvo, 2001). Experience in several countries, including Australia, has identified a number of problems with establishing partnerships at every level of government. These problems include poor structure of the partnerships, lack of awareness of the benefits of the partnership, lack of clear responsibilities of each partner, fear of losing of control of data, funding and buy-in (Wehn de Montalvo, 2001). Although these issues have been identified, the key problem remains of evaluating or measuring the impact of these issues (Dangermond and Brown, 2003; Nedovic-Budic and Pinto, 2001).

The successful implementation of the next generation of SDIs will, to a large extent, depend on the ability of SDI coordinators to comprehend and build on the success or failure of previous SDI initiatives (Giff, 2006). Although the SDI community continues to promote the benefits of spatial data infrastructures to society, no methodologies currently exist to measure the performance and outcomes of these infrastructures. Future investment in these increasingly critical infrastructures and guiding of government policy on the access to spatial data will depend on the availability of appropriate performance measures to justify further funding and development.

2. SDI DEVELOPMENT AT THE LOCAL LEVEL IN AUSTRALIA

The majority (77%) of Australia's 20.1 million people is located in the eastern states (Queensland, Victoria and New South Wales), although these three states represent only approximately 36% of the total land area. Although the majority of land management is undertaken by the state governments, it is local government that services the general community with respect to day-to-day property management issues. In September 2005, there were 673 local governments (councils) consisting of cities, towns, municipalities, boroughs, shires, districts, and in the Northern Territory, a number of rural Aboriginal communities.

Australia, like many developed countries, has progressively established a capacity to build, manage and distribute its spatial data across the government and non-government sectors. Local government in Australia is a system of government established under state government legislation and is governed by a council, elected directly by, and accountable to, the various communities which they serve. Local government authorities (LGAs), or councils as they are commonly termed, are multifunctional and provide a wide range of services through a single administrative structure for the governance and good management of towns, cities and communities (Hullick and Cooper, 1993).

Most local governments control or oversee land development and planning, parks, community facilities, environmental compliance, water supply, sewerage and community health amongst other responsibilities. The land related information and mapping that supports their decision-making is typically at a detailed level or large scale (1: 100 to 1:5 000). Local government in Australia was an early adopter of land information and geographic systems, both as a user of the early digital map products such as the digital cadastral data bases (DCDB) and also a prominent information contributor (McDougall and Perret, 1987; Williamson and Blackburn, 1985). Many of these developments were driven by the need for improved land use planning (Nash and Moll, 1976) and better financial management of the organisation and their assets (Cushing et al., 1975).

By the late 1970s, many local governments in Australia had computerised records of their properties for the purpose of rating and taxation, however these systems constituted financial management systems rather than spatial information systems. Even at this early stage of land information systems development, the problems of dealing with the complex nature of address, property and land parcels were recognised, and the concept of a unique property identifier was considered (Moyer and Fisher, 1973). The local government developments in Australia parallelled efforts in other countries such as the United Kingdom, where the development of systems such as the Local Authority Management Information Systems (LAMIS) were undertaken by local governments in conjunction with mainframe computer vender ICL (Mayr, 1992). Traditional computer applications for planning began to make way for more spatially demanding and accuracy specific applications such as engineering infrastructure, transport planning, property management and facilities management (Bomberger, 1983).

In the late 1980s to mid 1990s with the maturing of GIS software and the affordability of computer systems, GIS was adopted widely across both large and small local governments (Wadlow, 1989). This period was characterised as a time of system consolidation and data collection. It also coincided with the completion of many of the state government cadastral data bases which became a critical base data set for most local governments. Trends on adoption and diffusion of GIS and geographic information technology in local government have been explored in the USA (Budic, 1994; Budic and Godschalk, 1994; Warnecke, 1995), the UK (Campbell, 1993; Masser, 1993; Masser and Campbell, 1995) and Europe (Masser and Campbell, 1996). Although GIS technology has been adopted widely across local government in Australia, there is little documented evidence on its growth or diffusion within this sector of government in Australia.

The late 1990s and the early 2000s saw the improvement in cost efficiency of GIS technology and greater utilisation of the spatial data within local government. GIS now supports many activities including front counter enquiries, land planning, asset management, local health, environmental compliance and animal registration amongst others. GIS had become a tool and the information that it provided to the organisation

went from being "nice to have" to being "critical". Web mapping introduced spatial data to a broad base of LGA users and also improved community access to basic land and spatial data. Local governments have continued to be leaders in the application of spatial data and technology through the use of web mapping applications and location based services.

Compared with many countries, local government in Australia has a relatively narrow range of functions. For instance, it does not take general responsibility for the provision of services such as education and policing (United Nations Economic and Social Commission for Asia and the Pacific, 2003). The size of local governments in Australia reflects the diversity and often complexity of this tier of government. Approximately 36% of local governments are populated by less than 3 000 people and almost three quarters have a population of less than 30 000 people. Many of these sparsely populated local governments are located in the rural areas of Australia and provide critical infrastructure including roads, housing, water and sanitation.

3. A RESEARCH METHOD FOR UNDERSTANDING SDI CAPACITY AT THE LOCAL LEVEL

To better understand the complexity of the heterogeneous nature of local government and their capacity to contribute to SDI development, a survey of local governments was conducted across the three Australia states namely Victoria, Queensland and Tasmania. Contact was made with 183 LGAs across three states comprising: 74 in Victoria, 89 in Queensland and 20 in Tasmania). The states were selected on the basis of a variety of characteristics including geographic area, population and the number of local governments. These three states represent almost 50% of Australia's population base, approximately 35% of the total number of local governments and about 25% of the geographic land area, thereby providing a contrasting mixture of local governments, geography and institutional arrangements.

In order to ensure a high response rate to the survey, direct telephone contact was initially made to each of the local governments in the first instance. This enabled a contact person in each LGA to be identified so that the questionnaire was directed to the relevant person. After the telephone contact an email containing the URL for the web based survey was then sent to each LGA contact. After two weeks a reminder email was sent to follow up and improve the response rate. A total of 110 responses were received including seven responses which were rejected as either incomplete or invalid. The remaining 103 valid returns represent a response rate for the survey of 56%.

The LGA questionnaire was arranged in eight parts and included questions on each LGAs organisation, information policies, access to data, data holdings and maturity, use of standards, personnel, existing collaborations and outcomes from data sharing partnerships. Table 1 summarises the structure of the LGA questionnaire. Parts 1 to 7 investigated the capacity of each LGA across the components of an extended SDI model, whilst part 8 of the questionnaire examined the outcomes and overall level of satisfaction of LGAs with the data sharing partnership.

The LGAs surveyed across the three states varied dramatically in terms of the number of properties they manage and their capacity. The largest local government to respond was Brisbane City Council, with approximately 400 000 properties in its local government area. The smallest LGA to respond was also from Queensland, Nebo Shire Council, which has approximately 1 500 properties, but spread over an area of almost 10 000 square kilometres.

Table 1: Structure of the LGA questionnaire.

LGA Questionnaire Component	Topics Covered
Part 1: LGA Organisation	Number of properties, staffing, ICT capacity, GIS capacity, management support
Part 2: Policy on Use of Spatial Data	Internal and external policies, cost recovery, attitudes towards privacy, copyright and legal liability.
Part 3: Accessing Spatial Data/ Technology	Locating LGA data, technology and mechanisms to access spatial data
Part 4: About LGA Spatial Data	Importance of property data, use of state government data, requests for their data, completeness of their data
Part 5: Spatial Data Standards and Integration	Attitudes towards standards, use of metadata and level of data integration
Part 6: About People	Profile of staff in spatial management area, organisational change, training
Part 7: Collaboration with organisations	Level of collaboration, barriers and drivers, preferred models, expectations from data sharing and collaboration
Part 8: Outcomes from Specific Data Sharing Partnerships	Outcomes in terms of value, improved quality, improved communication, updates, overall satisfaction

A range of quantitative and qualitative analysis was undertaken on the questionnaire data. Analysis of the variations between the LGAs in each of the state government jurisdictions was determined by using statistical significance testing of the mean state results for each variable. Finally, factor and multiple regression analysis was undertaken to determine areas across the SDI framework that impact on the capacity of the LGAs to contribute to higher levels of SDI development.

4. RESULTS AND ANALYSIS

The results are presented in a summary form across the broad SDI component areas of data, policy, access arrangements, standards, organisational capacity, people and partnerships. Only the key findings are reported here for the sake of brevity.

4.1 Key Findings on Local Government Capacity

Although Australia's ICT infrastructure is poor in many remote areas of the country, its overall ICT infrastructure and capacity is comparable to other developed nations. In 2005, Australia was rated 11th out of 115 countries based on the network readiness index, and has generally improved its position since 2002. The survey results reflect these findings with most LGAs indicating that the ICT infrastructure was adequate.

The technical capacity within the local government sector is in part reflected by its ability to provide online services or e-business to their customers. The survey results indicated that 39% of LGAs are already providing online services to customers, whilst another 22% were in the process of developing these services. The states of Queensland and Tasmania were found to have the most mature spatial information systems with over a third of the LGAs having had a GIS established for 10 years or longer. This contrasted significantly with Victoria where only 7% of LGAs indicated that their GIS had been in place for more than 10 years whilst more than 53% of LGAs identified their GIS was less than six years old. The major area of resource deficiency identified was in the area of staffing. Most LGA respondents (84%) indicated that their GIS unit was staffed by three staff or fewer. In some cases, the officers responsible for managing the GIS were undertaking the GIS management in conjunction with other activities.

Approximately 59% of LGAs indicated that they did not have any formal policies on the use of their spatial data by external organisations or users. These findings agree with a 2004 survey of LGAs by Australian Local Government Association which found that approximately 60% of LGAs were found to have no formal policies on information.

The questionnaire findings found that approximately 30-40% of staff in LGAs having access to GIS at their desktops. This indicates the growing level of importance of spatial data to the organisations. Only 13% of LGAs indicated that external clients used the internet to find data which generally identified that the LGA had a web portal open to the public. However, this trend is most likely to change over the next few years as more LGAs begin to provide web access to their spatial data. The following comment was indicative of the general experience of LGAs to opening up their spatial data to the public:

Making the common property-based data freely available to the public via web-mapping has resulted in a sharp decline in ad-hoc queries and resulted in significant savings on staff time.

Local governments obtain a significant amount of their spatial data from the state government. LGAs from all three states indicated that the cadastral mapbase, property valuations data, orthophotography and topographic data were the most commonly sourced state datasets. When asked the question if the data they required from the state government agencies was easily accessible, 76% LGAs agreed or strongly agreed. On the separate question on the pricing of the data, 63% agreed that the cost of acquiring this data was acceptable. Additionally, the average level of completeness or maturity of LGA data sets is generally quite high, which indicates that LGAs have significant holdings of data which may potentially be available for sharing.

Although standards and formats were identified as being problematic by some LGAs, most indicated that standards were not a significant issue. Metadata is considered an important component of spatial data and identifies its source, currency and quality. However, only 42% of LGAs currently enter metadata within their GIS. These findings are supported by other documented studies such the Local Government and the Australian Spatial Data Infrastructure Project which identified that only 44% of LGAs stored metadata (Hawkesbury-Nepean Catchment Management Trust, 2000). Not surprisingly this study also identified that the majority of the metadata collected by these LGAs was not compliant with national metadata standards, which may inhibit future state and national efforts to exchange data.

The results indicate that almost 66% of LGAs have only one GIS staff member, a further 18% have either two or three GIS staff, and remaining 16% of LGAs have four or more GIS staff. This highlights the difficult situation faced by many LGAs in participating in data sharing partnerships. With only one staff member to manage the organisation's GIS work, the time available to undertake extra duties, such as partnership participation, is often limited.

Local government collaboration was examined across a range of possible partners including state and federal government, private sector, academic institutions and local government associations/groups. A number of general trends were evident. Firstly, LGAs were most likely to collaborate with State governments, followed closely by the relevant state local government association or regional local government group. Secondly, the difference in the level of collaboration/co-operation of LGAs with the state governments in Tasmania and Victoria in comparison to Queensland was significant

(p<0.001). This significant variation between Queensland and the other two states provides a useful barometer of the degree of trust and interaction between local and state government in each of these states.

LGAs identified that the greatest barriers to collaborating with state government agencies for spatial data sharing were legal liability, data standards, accessing of data, copyright and privacy. Motivations for the sharing of spatial data were found to be closely aligned with improving decision making and the delivery of services which emphasise the strong business basis for the exchange of data. The questionnaire found that 83% of the total respondents who had signed a data sharing partnership arrangement either agreed or strongly agreed that it had been worthwhile for their organisation. The levels of agreement were highest in Tasmania and Victoria whilst there was a lower level of agreement (approximately 60%) in Queensland.

A similar trend was observed on the question on whether the data sharing partnerships had improved their organisation's data quality. Again, the overall level of agreement to this question was high (71%) across the aggregated state data, with both Victoria and Tasmania responding positively. However, the level of agreement from Queensland LGAs was only 36% which reflects that the initial data sharing arrangement had done little to improve the quality of the LGAs data.

The overall level of satisfaction of the LGAs with the data sharing partnerships being investigated in each of the states revealed that across the three states, 73% of LGAs were either mostly satisfied or very satisfied with the outcomes of the data sharing partnership. The individual levels of positive satisfaction found for each state were Tasmania (92%), Victoria (91%) and Queensland (52%). Again the difference between Queensland and the other two states was significant.

4.2 Inter-Jurisdictional Differences

An analysis of the average responses on a state by state grouping was undertaken to determine areas of key inter-state difference. The results of the difference analysis are summarised in Table 2. The variables used in the questionnaire included a range of measurement types including continuous numeric values (e.g. number of properties), descriptive ordinal/internal values (e.g. Likert scale – agree, strongly agree) and categorical or nominal values. The categorical or nominal variables were not suitable for factor analysis and were therefore not utilised. Prior to the analysis the continuous numeric variables and the ordinal Likert variables were transformed to numerical interval classes between 0 and 5. Only the variables that illustrated significant inter-state variation (p<0.05) are tabulated. The variables highlighted in Table 2 have the highest level of significance with respect to inter-state variation with p-values <0.01 for both the ANOVA and Kruskal Wallis tests.

The length of time having a GIS was identified as a significant inter-state variation. This is partially explained by a recent Victorian state government data sharing partnership which assisted in the establishment of a large number of geographic information systems at the start of the project. A significant difference was identified in the cost recovery policy between the states which reflected the more restrictive pricing and access arrangements in Queensland at the time.

Table 2: Variables that illustrate significant inter-state difference (p <0.05).

Variables	State			ANOVA Significance p<0.05	Kruskal Wallis Significance p<0.05
	QLD	VIC	TAS		
	Mean	Mean	Mean		
Length of time having GIS	3.21	2.50	3.62	0.000	0.000
Cost recovery policy	2.02	1.31	1.62	0.000	0.000
Cost of state data is acceptable	3.31	3.74	4.23	0.001	0.002
Access to training	3.52	4.10	3.00	0.001	0.002
Average level of collaboration across organisations	2.93	3.25	2.52	0.002	0.001
Data sharing partnerships is worth-while	3.67	4.36	4.54	0.000	0.000
Data sharing has improved quality	3.21	4.29	4.38	0.000	0.000
Believe that the benefits are equal	3.40	4.05	3.31	0.001	0.001
Are provided updated data regularly	3.17	4.69	4.23	0.000	0.000
Communication frequency	3.35	4.12	3.92	0.000	0.000
Overall level of satisfaction	3.48	4.17	4.31	0.000	0.000

Differences between the states were identified in the trends on collaboration. Tasmania and Victoria appear to have developed a higher level of trust and intergovernmental relations than Queensland. The overall level of satisfaction with the data sharing partnerships was also highest in Victoria and Tasmania. Queensland and Victoria showed significant differences towards web mapping and external accessibility of data when compared to Tasmania. The difference between the LGAs' perspectives on the cost of state government data is most evident between Queensland and Tasmania, with Queensland LGAs less satisfied with the pricing arrangements and Tasmanian LGAs generally very satisfied. A similar trend can be seen on the LGAs view on the limitations placed on the use of state data by the data custodians.

Tasmania showed significant differences in the overall maturity levels of their spatial data holdings, generally being higher than Queensland and Victoria. There was also evidence of greater staff stability in this state than in the other states, perhaps related to the relative isolation of this island state and the smaller size of their LGAs in comparison to Queensland and Victoria. Victoria showed strong differences from the other two states in the areas of access to training and the level of positive collaborations with other organisations. It is suggested that the smaller state size and access to training provided by the state government in return for sharing data is responsible for these results.

4.3 Factor and Regression Analysis

A factor and regression analysis was undertaken to identify which variables or groups of variables were contributing towards the success of the data sharing partnerships and hence SDI development. Factor analysis is a well documented technique that assists in identifying clusters of variables that may be logically grouped into a smaller set of these variables which have common underlying constructs or factors (Brace et al., 2006). The factor analysis was undertaken using the standard principal component analysis method to reduce the total number of independent variables from 36 to 13 grouped factor components.

Table 3: Results of multiple regression modelling.

Factor	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	2.202	.895		2.461	.016
Size	-.015	.076	-.025	-.203	.839
Organisational Support and Attitudes	.294	.156	.221	1.883	.063
Data Accessibility/Maturity	-.164	.149	-.148	-1.100	.274
Internal Accessibility	-.103	.127	-.102	-.811	.420
Access to State Data	.372	.115	.343	3.244	.002
Level of concern on data restrictions	.110	.089	.111	1.240	.218
Standards and Metadata	-.067	.092	-.068	-.726	.470
Use of State Data and Restrictions	.104	.114	.095	.914	.363
Organisational Change	-.172	.132	-.128	-1.301	.197
Staff Growth and Training	.057	.126	.047	.456	.650
Business Needs	.266	.098	.247	2.705	.008
Policy on External Access to Data	-.237	.077	-.260	-3.056	.003
Length of Collaboration	-.036	.041	-.076	-.898	.371

A multiple regression model using the simultaneous technique was then applied using the 13 grouped components from the factor analysis as the independent input variables and the satisfaction with existing data sharing partnerships as the as the dependent variable. The analysis yielded a model (see Table 3) that was significant: $F(13,88) = 4.659$, $p<0.005$, with an Adjusted $R^2 = 0.32$, which indicates that the model has accounted for approximately 32% of the variance in the criterion variables.

The highlighted component factors in Table 3 namely, organisational support, awareness of state data, external access policy and the business needs are identified as significant to the partnership outcomes. The organisational support factor importantly encompasses ICT capacity, management support and attitudes to making data and resources available. This emphasises the importance of assessing a potential partner's capacity during partnership development to better understand the ability of the organisation to contribute to the partnership outcomes.

Policies on access and pricing were again identified as important to the outcome of the partnerships. Policies at state and local level should be aligned to ensure that there is minimal conflict. Local government are more likely to follow state government policy direction due to their limited capacity to resource their own policy development. External access policies and the use of the internet are identified as important considerations for partnership development.

The business needs factor underlines the importance of maintaining a business focus for the data sharing initiative to be sustainable. If the data sharing initiative is linked to important business processes, it is more likely to receive priority and be incorporated within mainstream operations. Wehn de Montalvo (2003) in her study on the willingness to share data, found that attitude and social pressure were the strongest determinants of willingness to share spatial data. In particular, organisational pressure, GIS community pressure, knowledge creation and social outcomes were identified by as key determinants.

5. DISCUSSION AND IMPLICATIONS FOR THEORY AND PRACTICE

This analysis of local governments examined their capacity, characteristics and outcomes of the data sharing partnerships in the states of Queensland, Victoria and Tasmania. A number of significant trends and differences were identified amongst the variables and across the three states. The initial analysis of the questionnaire data has identified a number of important characteristics of local governments including their capacity across a number of the identified SDI component areas, existing preferences for collaboration and their level of satisfaction with the existing data sharing partnerships. The organisational analysis identified that the ICT capacity of LGAs was significantly better than expected and management support for GIS was generally satisfactory. Policies on access and pricing are not well developed in local government, as small staff numbers and other activities take priority. It is therefore important that state government agencies continue to lead and support LGAs to develop their policy frameworks. LGAs appear more likely to adopt or mimic the state government policies on access and pricing, although this has not been proven conclusively.

The findings from the factor analysis underscore the key motivations for sharing of data, particularly at the local government level. LGAs are very tightly resourced and highly business driven. Therefore, the linkage of data sharing initiatives to the business processes of LGAs is more likely to result in more successful and sustainable outcomes. The research also indicates that policies at that state and local level should be aligned where possible to ensure that there is minimal conflict. Local governments are more likely to follow the lead of state agencies on policy development due to their limited capacity to develop their own specific spatial data access and pricing policies.

Often, LGAs are at the cutting edge of spatial data access and provision through the use of the internet and web mapping. Because of the closeness of LGAs to their customers, they see immediate and significant benefits through providing information access to the local community. Information access facilitates better service and evidence indicates that it reduces the number of general enquiries. Organisational support and leadership were also rated highly and agree with previous theoretical and empirical research.

Local government data is increasingly available over the web and indications are that it will be a strong driver for facilitating business and reducing the number of over-the-counter enquiries for LGAs. The level of completeness of core data sets was very high for most local governments which should provide an excellent basis for exchanging digital data. Standards and metadata were identified as issues that will continue to demand attention and strategies to improve compliance in these areas. Integration of data across the LGAs is well advanced, but full interoperability is still some way off.

6. CONCLUSIONS

Spatial data is widely utilised across all levels of government, business and the general community. The objectives of SDI initiatives are to create, maintain and disseminate spatial data for the benefit of society. However, the co-operation and exchange of information has continued to be problematic with detrimental impacts on government business and areas such as emergency services. This research has found that local governments have mature spatial data holdings and the ICT infrastructure to facilitate SDI development through the wider sharing of data.

The role of local government in building and developing SDI at the local level is critical. Although the local government environment is complex, a number of important trends have emerged from this research. Firstly, LGAs have a strong focus on meeting their business needs and, therefore, SDI development should be considered as a significant business enabler. Secondly, LGA capacity to develop information policy frameworks is often limited, so it is essential that strong and positive information policy is provided and disseminated at the state level. Finally, local government must be viewed as an equal partner in SDI development to engender trust and facilitate data sharing on an equitable basis.

ACKNOWLEDGEMENTS

REFERENCES

Abel, D.J., V.J. Gaede, K.L. Taylor and X. Zhou (1999). SMART: Towards Spatial Internet Marketplaces, *Geomatica,* 3(2): 141-64.

Bomberger, D. (1983). "Local government's use of computer graphics: current and future applications to benefit communities." *Proceedings of 11th Australian Conference on Urban and regional Information Systems*, 22-25 November, 1983, Brisbane. pp. 11-18.

Brace, N., R. Kemp and R. Snelgar (2006). *SPSS for Psychologists,* 3rd ed. New York: Palgrave Macmillan.

Budic, Z.D. (1994). Effectiveness of geographic information systems in local planning, *Journal of the American Planning Association,* 60(2): 244-63.

Budic, Z.D. and D.R. Godschalk (1994). Implementation and management effectiveness in adoption of GIS technology in local governments, *Computers, Environment and Urban Systems,* 18(5): 285-304.

Campbell, H. (1993). "GIS implementation in British local government", in I Masser and HJ Onsrud (Eds). *Diffusion and Use of Geographic Information Technologies.* Dordrecht, Netherlands: Kluwer Academic publishers, pp. 117-46.

Craglia, M. and P. Signoretta (2000). From global to local: the development of local geographic information strategies in the United Kingdom, *Environment and Planning B: Planning and Design,* 27(5): 777-88.

Crompvoets, J., A. Bregt, A. Rajabifard and I.P. Williamson (2004). Assessing worldwide developments of national spatial data clearinghouses, *International Journal of Geographic Information Science,* 18(1): 1-25.

Cushing, W., L.W. Parish and C.J. Warren (1975). "Local government finance in Australia: Towards a uniform information system", *Proceedings of URPIS 3*, 27-28 November 1975, Newcastle, pp. 1-21.

Dangermond, J. and C. Brown (2003). GIS: infrastructure underpinnings for the National Map, *Photogrammetric Engineering and Remote Sensing,* 69(10): 1159-64.

Giff, G. (2006). "The value of performance indicators to spatial data infrastructure development", *Proceedings of International Conference on Global Spatial Data Infrastrucure (GSDI9)*, 3-11 November, Santiago, Chile.

Harvey, F.J. and D. Tulloch (2006). Local-government data sharing: Evaluating the foundations of spatial data infrastructures, *International Journal of Geographical Information Science,* 20(7): 743-68.

Harvey, F.J., B.P. Buttenfield and S.C. Lambert (1999). Integrating geodata infrastructures from the ground up, *Photogrammetric Engineering and Remote Sensing,* 65(11): 1287-91.

Hawkesbury-Nepean Catchment Management Trust (2000). *Local government and the Australian Spatial Data Infrastructure - Spatial data and metadata workshop report*.

Hullick, J.M. and M.J. Cooper (1993). "The role of local government in urban management in the 1990", *Proceedings of AURISA 93*, 22-26 November, Adelaide. pp. 26-34.

Masser, I. (1993). "Diffussion of GIS in the British local government", in I. Masser and H.J. Onsrud (Eds.). *Diffusion and Use of Geographic Information Technologies.* Dordrecht, Netherlands: Kluwer Academic publishers, pp. 99-115.

Masser, I. (1998a). *Governments and geographic information,* London: Taylor and Francis Ltd.

Masser, I. (1998b). "The first generation of national geographic information strategies", *Proceedings of GSDI 3*, 17-19 November 1998, Canberra, Australia.

Masser, I. (1999). All shapes and sizes: the first generation of national spatial data infrastructures, *International Journal of Geographic Information Science,* 13(1): 67-84.

Masser, I. (2005). *GIS worlds - creating spatial data infrastructures,* Redlands, CA: ESRI Press.

Masser, I. and H. Campbell (1994). "Information sharing and the implementation of GIS: some key issues", in M.F. Worboys (Ed.). *Innovations in GIS.* Taylor and Francis, pp. 217-27.

Masser, I. and H. Campbell (1995). "Information sharing: the effects of GIS on British local government", in H.J. Onsrud and G. Rushton (Eds.). *Sharing Geographic Information.* New Brunswick, NJ: Centre for Urban Policy Research, pp. 231-49.

Masser, I. and H. Campbell (1996). "Great Britian: the dynamics of GIS diffusion", in I. Masser, H. Campbell and M. Craglia (Eds.). *GIS diffusion: the adoption and use of geographical information systems in local government in Europe.* London: Taylor and Francis, pp. 49-66.

Mayr, W. (1992). "Desktop GIS - power to the workers", *Proceedings of 20th Australian Conference on Urban and Regional Information Systems*, 25-27 November, Gold Coast. pp. 1.01-1.47.

McDougall, K. (2006). *A Local-State Government Spatial Data Sharing Partnership Model to Facilitate SDI Development,* PhD Thesis, The University of Melbourne.

McDougall, K. and P. Perret (1987). The Australian Key Centre in Land Information Studies - its developments in LIS and local government. *Association of Consulting Surveyors Queensland Newsletter.* 7 pages.

McDougall, K., A. Rajabifard and I.P. Williamson (2002). "From little things big things grow: building the SDI from local government up", *Proceedings of AURISA 2002 and 3rd Trans Tasman Surveyors Conference*, 25-30 November, Adelaide.

McDougall, K., A. Rajabifard and I.P. Williamson (2005). "Understanding the motivations and capacity for SDI development from the local level", *Proceedings of 8th International Conference on Global Spatial Data Infrastructures (GSDI 8)*, 16-21 April, Cairo, Egypt.

Moyer, D. and K.P. Fisher (1973). *Land Parcel Identifiers For Information Systems,* Chicago: American Bar Foundation.

Nash, K.R. and A.P. Moll (1976). "Council of the City of Sydney land information system: description of planning applications", *Proceedings of 4th Australian Conference on Urban and regional Information Systems*, 25-26 November, Melbourne. pp. 6.01-6.20.

Nedovic-Budic, Z. and J.K. Pinto (2000). Information sharing in an interorganizational GIS environment, *Environment and Planning B: Planning and Design,* 27(3): 455-74.

Nedovic-Budic, Z. and J.K. Pinto (2001). Organizational (soft) GIS interoperability: lessons from the U.S., *International Journal of Applied Earth Observation and Geoinformation,* 3(3): 290-8.

Onsrud, H.J. and G. Rushton (1995a). "Sharing geographic information: an introduction", in H.J. Onsrud and G. Rushton (Eds.). *Sharing geographic information.* New Brunswick, New Jersey: Centre for Urban Policy Research, pp. xiii-xviii.

Onsrud, H.J. and G. Rushton (1995b). "Fundamental questions and future research", in H.J. Onsrud and G. Rushton (Eds.). *Sharing Geographic Information.* New Brunswick, NJ: Centre for Urban Policy Research, pp. 493-502.

Pinto, J.K. and H.J. Onsrud (1995). "Sharing geographic information across organisational boundaries: a research framework", in H.J. Onsrud and G. Rushton (Eds.). *Sharing geographic information.* New Brunswick, New Jersey: Centre for Urban Policy Research, pp. 45-64.

Rajabifard, A., A. Binns, I. Masser and I.P. Williamson (2006). The role of sub-national government and the private sector in future spatial data infrastructures, *International Journal of Geographical Information Science,* 20(7): 727-41.

United Nations Economic and Social Commission for Asia and the Pacific (2003). Local Government in Asia and the Pacific: A Comparative Study, at http://www. unescap.org/huset/lgstudy/country/australia/australia.html.

Wadlow, R.F. (1989). "The impact of land information systems in local government", *Proceedings of the 17th Australasian Conference in Urban and Regional Planning Information Systems*, Perth, WA.

Warnecke, L. (1995). *Geographic information/GIS institutionalization in the 50 states: users and coordinators*, Report, National Center for Geographic Information and Analysis, Syracuse, NY.

Warnecke, L., D. Decker, L. Pelch, S. Davis and J. Gilligan (2003). *Statewide Leadership and Coordination of Geographic Information and Related Technology in the 50 States - NSGIC State Summaries*, Report to National States Geographic Information Council.

Warnest, M., A. Rajabifard and I.P. Williamson (2005). "A collaborative approach to building national SDI in federated state systems: case study of Australia", *Proceedings of GSDI-8*, 16-21 April 2005, Cario, Egypt.

Wehn de Montalvo, U. (2001). "Strategies for SDI implementation: a survey of national experiences", *Proceedings of 5th Global Spatial Data Infrastructure Conference*, 21-25 May, 2001, Cartagena de Indias, Columbia.

Wehn de Montalvo, U. (2003). *Mapping the determinants of spatial data sharing,* Aldershot, England: Ashgate Publishing Ltd.

Williamson, I.P. and J.W. Blackburn (1985). "Land information systems in Australasia: lessons and experiences", *Proceedings of 1985 AURISA National Conference (URPIS 13)*, Adelaide, SA. pp. 263-77.

Changing Notions of a Spatial Data Infrastructure

Ian Masser

Town End House, Taddington, Buxton SK17 9UF, England
ian.masser@yahoo.co.uk

Abstract
This article examines some of the changes that have taken place in the notion of a spatial data infrastructure (SDI) over the last 15 years. The discussion is divided into five parts. The first of these considers the impacts of innovations in communications and information technology during this period on the nature of SDIs. The second examines the changes that have taken place in the conceptualisation of SDIs while the third discusses the changing nature of SDI implementation in the context of the concepts of multi level governance that have been developed by political scientists. Underlying this discussion is the realisation that SDI development and implementation is very much a social process of learning by doing. This process is explored in the fourth section of the article with reference to the experience of the State of Victoria in Australia. The concluding section of the article considers the challenges facing SDI implementation and identifies a number of dilemmas that have yet to be resolved.

Keywords: Spatial data infrastructure (SDI), SDI implementation, multi level governance, learning by doing.

1. INTRODUCTION

The term 'National Spatial Data Infrastructure' was first used in a paper presented by John McLaughlin at the 1991 Canadian Conference on Geographic Information Systems entitled 'Towards a national spatial data infrastructure' (McLaughlin, 1991). The ideas contained in this paper were subsequently developed by the United States National Research Council's Mapping Science Committee in their report on 'Toward a coordinated spatial data infrastructure for the nation' (National Research Council, 1993). This recommended that effective national policies, strategies, and organisational structures should be established for the integration of national spatial data collection, use and distribution.

These concepts were expanded and developed during the following year in the Executive Order 12906 signed by President Clinton entitled 'Coordinating Geographic Data Acquisition and Access: the National Spatial Data Infrastructure' (Executive Office of the President, 1994). The Executive Order significantly raised overall awareness of the need for governmental strategies that facilitate geospatial data collection, management and use not only among Federal agencies in the United States, but also nationally and internationally (Masser, 2005).

Since then the number of SDI initiatives has increased dramatically in all parts of the world to the extent that Crompvoets et al.'s (2004) work on the development of clearinghouses suggests that as many as half the world's countries were considering SDI related projects. These figures must be treated with some caution as they do not necessarily imply that all these countries are actively engaged in SDI formulation or implementation. It is also likely that many of them may be engaged in some aspects of SDI development without necessarily committing themselves to a comprehensive SDI

programme. Nevertheless the term 'SDI phenomenon' seems to be a reasonable description of what has happened in this field over the last fifteen years.

With these considerations in mind this article examines some of the changes that have taken place in the notion of a SDI during this time. The discussion is divided into five parts beginning with technological developments and then moving on to institutional matters. The first of these considers the impacts of innovations in communications and information technology during this period on the nature of SDIs. The second examines the changes that have taken place in the conceptualisation of SDIs while the third discusses the nature of SDI implementation with particular reference to the concepts of multi level governance that have been developed by political scientists. Underlying a great deal of this discussion is the notion that SDI development and implementation is very much a social process of learning by doing. Some of the main features of this process are examined in the fourth section of the article with reference to the experience of the State of Victoria in Australia. The concluding section of the article considers the challenges facing SDI implementation and identifies a number of dilemmas that have yet to be resolved.

2. THE IMPACT OF INNOVATIONS IN INFORMATION AND COMMUNICATIONS TECHNOLOGIES

New technologies have played an important role in the evolution of the SDI concept. The earliest SDIs were conceived before the Internet and the World Wide Web (WWW) came into being and the opportunities opened up by their development have dramatically transformed the way that way that data is delivered to users. This was recognised by the US Mapping Sciences Committee in their report on Distributed Geolibraries (National Research Council, 1999). In their view, "the WWW has added a new and radically different dimension to its earlier conception of the NSDI, one that is much more user oriented, much more effective in maximizing the added value of the nation's geoinformation assets, and much more cost effective as a data dissemination mechanism."

The WWW has developed very rapidly over the last few years and the term 'Web 2.0' was introduced around 2005 to highlight the changes that had taken place since the emergence of Web 1.0 in the 90s (O'Reilly, 2005). The most important differences between the two can be seen from some contrasting examples which illustrate the interactive and participatory nature of Web 2.0. The Web 1.0 consisted largely of static sites such as the Encyclopaedia Britannica online whereas the Web 2.0 hosts dynamic sites such as Wikipedia that are constantly being revised and enlarged by the contributions from users. Similarly the personal websites that characterised the Web 1.0 have been replaced by the interactive blogs that are an important feature of the Web 2.0. One of the standard bearers for Web 1.0 was the Netscape server while Google can be seen as the standard bearer for Web 2.0. Unlike Netscape, Google began life as a web application that was delivered as a service with customers paying directly or indirectly to use that service.

These differences are reflected in the development of the GeoWeb that underpins the emergence of SDIs. The most important of these from a user perspective have been summarised in Table 1. From this it can be seen that the GeoWeb 2.0 is essentially dynamic, participatory, user centric, distributed, loosely coupled and rich in content in contrast to the static, producer driven and producer centric, centralised and closely coupled basic content of the GeoWeb 1.0.

Table 1: Differences between GeoWeb 1.0 and 2.0 (Maguire, 2005).

GeoWeb 1.0	GeoWeb 2.0
Static	Dynamic
Publishing	Participation
Producer centric	User centric
Centralised	Decentralised
Close coupling	Loose coupling (e.g. mash ups)
Basic	Rich

The launch of Google Earth in June 2005 brought many of the elements of the Ge-oWeb 2.0 within reach of millions of users. Google Earth combined the powerful search engines developed by Google with the ability to zoom rapidly in or out from space to the neighbourhood street level. It also created new opportunities for these users to overlay their own spatial data on the top of Google Earth's background imagery. As Butler (2006, 776) pointed out in an article in the science journal, Nature: 'By offering researchers an easy way into GIS software, Google Earth and other virtual globes are set to go beyond representing the world, and start changing it.' For this reason they must be regarded as 'disruptive technologies' that are transforming the GIS industry in ways that the market does not expect.

A position paper from the Vespucci initiative (Craglia et al., 2008) highlights some of the impacts of the developments in information technology, spatial data infrastructures and earth observation that have taken place since the launch of Vice President Gore's (1998) vision of Digital Earth. It points out that many elements of his vision are now regularly being used by large numbers of people throughout the world and that geography has become an important way of organising many different kinds of digital spatial data that are now regularly collected by sensors that provide multi level spectral information about the earth's surface in large scale intergovernmental initiatives such as the Global Earth Observation System of Systems (GEOSS). The paper also sets out its own vision for the next five to ten years. Elements of this vision include the development of multiple connected infrastructures addressing the needs of different audiences, and the possibility of searches through time and space to find analogous situations with real time data from both sensors and individuals.

3. FROM PRODUCERS TO USERS - THE GENERATION ANALOGY

There are interesting parallels between the shift from producers to users that has occurred as a result of emergence of the WWW and the changes that have taken place in the governance of SDIs over this time. A good example of the latter can be found in the typology of SDIs that has been developed in the course of the State of Play studies that have been carried out by the Spatial Application Division at the University of Leuven for the European Commission over the last five years (Vandenbroucke et al., 2008). This typology is based on the coordination aspects of national SDI initiatives. Matters of co-ordination have been emphasised because 'it is obvious coordination is the major success factor for each SDI since coordination is tackled in different ways according to the political and administrative organisation of the country' (SADL, 2003). A basic distinction is made between countries where a national data producer such as a mapping agency has an implicit mandate to set up a SDI and countries where SDI development is being driven by a council of Ministries, a GI association or a partnership of data users. A further distinction is then made between initiatives that do and do not involve users in the case of the former and between those that have a formal mandate and those that do not in the case of the latter.

This distinction is also reflected in the generation analogy that has been used to high-light the main structural changes that have taken place in the notion of spatial data infrastructures over the last fifteen years. Some features of the first generation of eleven SDIs that had emerged during the first half of the 1990s were described by Masser (1999). What distinguished these from other GI policy initiatives was that they were all explicitly national in scope and their titles all referred to geographic information, geo-spatial data or land information and included the term 'infrastructure', 'system' or 'framework'.

The development of a second generation of SDIs began around 2000 (Rajabifard et al., 2003). The most distinctive feature of the second generation of SDIs was the shift that was taking place from the product model that characterised most of the first generation to a process model of a SDI (Table 2). Database creation was to a large extent the key driver of the first generation and, as a result, most of these initiatives tended to be data producer, and often national mapping agency, led. The shift from the product to the process model is essentially a shift in emphasis from the concerns of data producers to those of data users.

This shift had profound implications for this involved in SDI development in that it has resulted in data users becoming actively involved in SDI development and implementation. The main driving forces behind the data process model are data sharing and reusing data collected by a wide range of agencies for a great diversity of purposes at various times. Also associated with this change in emphasis is a shift from the centralised structures that characterised most of the first generation of national SDIs to the decentralised and distributed networks that are a basic feature of the WWW.

Table 2: Current trends in SDI development (Masser, 2005, p. 257).

From a product to a process model	From formulation to implementation
From data producers to data users	From coordination to governance
From database creation to data sharing	From single to multilevel participation
From centralised to decentralised structures	From existing to new organisational structures

There has also been a shift in emphasis from SDI formulation to implementation as those involved gained experience of SDI implementation and a shift from single level to multi level participation, often within the context of an administrative hierarchy of SDIs. As a result of these developments the coordination models that had emerged for single level SDIs have been substantially modified and more complex and inclusive models of governance have emerged. They may also require the creation of new kinds of organisational structure to facilitate effective SDI implementation.

In the last few years there are also signs that a third generation of SDIs is emerging. The most important difference between the second and third generation is that the balance of power in the latter has shifted from the national to the sub national level (Rajabifard et al., 2006). Most large-scale land related data is collected at this level where it is used for collecting land taxes, land use planning, road and infrastructure development, and day-to-day decision making. Alongside these developments there has been a shift from government led approaches to whole of industry models where the private sector operates on the same terms as its government partners. One consequence is that national SDI activities are likely to be increasingly restricted to the strategic level while most of the operational level decisions are handled at the sub national levels by local government agencies in conjunction with the private sector.

The concept of spatially enabled government that is emerging as a result of these trends presents important challenges for those involved. The initial development of SDIs was largely in the hands of small elite of spatially aware professionals from the fields of geography, planning, surveying, land administration and environmental science. This elite not only dominated the production of geographic information, but were also its main users. In recent years, as a result of the development of location based services and the expansion of eGovernment activities the position has substantially changed to the extent that the vast majority of the public are users, either knowingly or unknowingly, of spatial information (Masser et al., 2008). As a result many traditional professional practices must be drastically altered to ensure that SDIs develop in such a way that they provide an enabling platform that will serve the wider needs of society in a transparent manner.

4. SDI IMPLEMENTATION IN THE CONTEXT OF MULTI LEVEL GOVERNANCE

Many national SDI documents seem to abide by the principle of 'one size fits all'. They suggest that the outcome of SDI implementation will lead to a relatively uniform product at the sub national level. However, there is both a top down and a bottom up dimension to national SDI implementation. National SDI strategies drive state wide SDI strategies and state wide SDI strategies drive local level SDI strategies and the outcomes of these processes are likely to be that the level of commitment to SDI implementation will vary considerably from state to state and from local government to local government.

The top down vision of a SDI emphasises the need for standardisation and uniformity whereas the bottom up vision stresses the importance of diversity and heterogeneity given the different aspirations of the various stakeholders and the resources that are at their disposal. Consequently the challenge to those involved in SDI initiatives is to find ways of ensuring some measure of standardisation and uniformity while recognising the diversity and the heterogeneity of the different stakeholders. This is likely to become increasingly important as sub national agencies take over the operational activities associated with SDI implementation.

The SDI that emerges from this process will have many features in common with a patchwork quilt or a collage of similar, but often quite distinctive components. The patchwork quilt analogy assumes that the SDI outcome will be like the product of similar pieces of cloth of various colours sewn together to form a bedcover. This is a particularly useful where the SDI participants are largely administrative regions with similar functions in the hierarchy. The collage analogy, on the other hand, is based on the notion of a picture that is built up from different pieces of paper and other materials. This is most useful where the participants such as transportation and environmental agencies cover overlapping administrative districts (Masser, 2007, p. 80-82).

These two analogies broadly correspond to the two types of multi level governance identified by political scientists such as Hooghe and Marks (2003) whose key features are summarised in Table 3. Type 1 governance describes jurisdictions at a limited number of levels as in the patchwork quilt model. These jurisdictions are essentially general purpose in that they bundle together many different functions such a housing, education, roads and environmental affairs. Membership of such jurisdictions is usually territorial in terms of nation, region or community and they are characterised by non intersecting memberships between different jurisdictions at the same level. In other words, a citizen may belong to only one of these jurisdictions. A limited number of levels are involved in these jurisdictions which are intended to be stable for periods of

several decades or more. In essence, every citizen is located in a Russian Doll of nested jurisdictions where there is only one relevant jurisdiction at each level of the administrative hierarchy.

Table 3: Types of multi level governance (Hooghe and Marks, 2003, p. 236).

Type 1	Type 2
General purpose jurisdictions	Task specific jurisdictions
Non intersecting memberships	Intersecting memberships
Jurisdictions at a limited number of levels	No limit to the number of jurisdictional levels
System wide architecture	Flexible design

Type 2 governance, on the other hand, is composed of specialised task specific jurisdictions such as school catchment areas, watershed management regions, and travel to work areas. Like a collage it is fragmented in nature with every piece fulfilling its own function. In type 2 governance there is no reason why smaller jurisdictions should be neatly contained in larger ones while others may define a small segment of a larger area as is the case with a site of special scientific interest within a National Park. There is no limit to the number of jurisdictional levels that are designed to respond flexibly to new needs and circumstances.

Generally, type 2 governance activities are embedded in type 1 structures, but the way that this works out varies considerably. Type 1 jurisdictions are rooted in community identities whereas type 2 structures are more pliable. The main benefit of multi level governance lies in its scale flexibility. Its chief cost lies in the transaction costs of coordinating multiple jurisdictions. The coordination dilemma confronting multi level governance is described by Hooghe and Marks (2003, p.239) in the following terms: "To the extent that policies of one jurisdiction have spillovers (i.e. negative or positive externalities) for other jurisdictions, so coordination is necessary to avoid socially perverse outcomes".

One strategy for dealing with the coordination dilemma that underpins type 1 governance is to limit the number of autonomous actors who have to be coordinated by limiting the number of autonomous jurisdictions. An alternative approach is to limit coordination costs by constraining interaction across jurisdictions. Type 2 governance sets no ceiling on the number of jurisdictions, but may spawn new ones along functionally differentiated lines to minimise externalities across jurisdictions. Both these strategies have important implications for those concerned with SDI implementation.

5. SDI IMPLEMENTATION AS A SOCIAL LEARNING PROCESS

The old adage that Rome wasn't built in a day is equally applicable to SDIs. The creation of SDIs is a long term task that may take years or even decades in some cases before they are fully operational. This process is likely to be an evolving one that will also reflect the extent to which the organisations that are involved are changing themselves over time. As a result major changes in the form and content of SDIs can be expected over time as they reinvent themselves. In some instances this process may even lead to the closing down of a SDI as was the case with the British National Geospatial Data Framework in 2002 (Masser, 2005).

The experiences of the State of Victoria in Australia provide a good example of learning by doing during the implementation process at the sub national level. It is worth considering because Victoria has been particularly innovative in recent years both in the

applications field and also in terms of the steps that it has taken to promote spatially enabled government. It is also one of the few SDIs that have published regular reports during the implementation process and cross referenced new developments in relation to previous work. This makes it possible to trace SDI evolution in more detail than is usually the case.

The main stages in the evolution of Victoria's Spatial Information Strategy are summarised in Table 4. From this it can seen that they date back at least twenty years. Concerns about duplication in maintaining computerised databases led the Victorian government to set up LANDATA to coordinate the development of a common land information system for the state as far back as 1984. This body turned out to be both under resourced and capable of producing digital maps only in a format that was unsuitable for modern GIS applications. As a result the state commissioned Tomlinson Associates to carry out a comprehensive GIS planning study in 1991. This study examined the work carried out by 40 state agencies and reviewed 270 data sets. It was a seminal work for the spatial information industry in the state which demonstrated both the strategic importance and the economic potential of land information (Thompson et al., 2003).

Table 4: The evolution of Victoria's spatial information management framework (Department of Sustainability and Environment, 2008, p.13).

1993 Core spatial information identified Government wide planning methodology introduced
Victorian Geospatial Information Strategy 1997-2000 Core data improved Spatial management framework put in place - policy, infrastructure, awareness, distribution, business systems Core principles for managing spatial information introduced - metadata, quality management, privacy, liability, licensing, pricing, custodianship Coordinating and cooperative arrangements between key stakeholders established
Victorian Geospatial Information Strategy 2000-2003 Spatial information management principles further codified Introduction of the concepts of 'framework' and 'business' information Role of custodians defined Framework datasets identified and custodians assigned
Victorian Spatial Information Strategy 2004-2007 Best practice principles for spatial information management extended to custodians of all spatial datasets Custodianship formally identified as the basis for spatial information management Holistic spatial information management framework defined Victorian Spatial Council established

This was followed by another strategy for the period from 2000 to 2003. This further consolidated the creation of high quality fundamental data sets for the state and set out best practice management principles for custodianship, metadata, access, pricing and licensing and spatial accuracy (Thompson et al., 2003). The concept of custodianship lies at the heart of the State's spatial information management strategy.

While the proposals for the next strategy were under consideration the Land Information Group with its 70 staff took a new position in early 2004 within the Department of Sustainability and Environment as the Spatial Information Infrastructure component of its Strategic Policy and Projects group. This move made it possible for those involved

to play a larger part in the development of spatial information policy for the Department as a whole.

The Victorian Spatial Information Strategy for 2004 to 2007 (VSIS) differed from its predecessor in several important ways. The VGIS strategies had focussed largely on issues of management and custodianship associated with the eight fundamental data sets that are the core of the State's SDI, but the VSIS was much broader in scope and presented a whole of industry approach rather than a governmental model. The implementation of the VSIS and the preparation of the latest strategy for 2008-2010 have been overseen by the whole of industry body, the Victorian Spatial Council, that was set up in 2004. Its membership includes representatives drawn from state government (3), local government (2), federal government (1), academia (2), the professions (2) and the private sector (2). An independent chairman of the Council has been appointed by the Secretary of the State Department of Sustainability and Environment.

The experiences of Victoria indicate that a combination of internal and external factors affects the evolution of SDIs over time. Internally, those involved participate in a process of learning by doing that takes account of the experiences of earlier stages of SDI implementation. Externally, important changes in the nature of the SDI may be a consequence of the restructuring of other activities within government as a whole. The interaction between these two strands will govern the trajectory of SDI development.

6. SOME CHALLENGES FACING SDI IMPLEMENTATION

SDIs have attracted a lot of attention from governments all over the world over the last ten years. This raises the inevitable question as to how far they will be able to deliver the promised benefits over time. Bregt and Crompvoets (2004) have argued that some SDIs may have already raised unrealistic expectations and their benefits that are not proven. These are likely to attract few stakeholders and can be classified as 'hype'. In contrast, the successful 'hit' SDIs will be those that have developed in response to realistic expectations and can deliver proven benefits. Most or all the relevant stakeholders are likely to be involved in such SDIs.

In practice, as the discussion in section 3 of this article shows, SDI outcomes are likely to be much more complex in practice because of the nature of the implementation processes. The discussion of multi level governance in section 4 of this article suggests that successful SDI implementation will be heavily dependent upon the extent to which sub national agencies are actively involved. The challenges that arise at the sub national level are highlighted in the findings of the Advanced Regional SDIs workshop that was organised by the Joint Research Centre last year (Craglia and Campagna, 2009) and are the central focus of attention in the current ESDInet+ project that is funded by the European Commission (www.esdinetplus.eu).

The findings of these and other studies suggest that effective SDI implementation is often facilitated in countries such as Australia and Germany where many important administrative responsibilities are devolved to the state level and established institutions already exist at this level for policy making and implementation. However, it should also be noted that the information infrastructures that come into being at this level are often different in many respects from those at the national level (De Man, 2007) and also that these fall essentially into the type 1 governance category. Consequently, further challenges may have to be overcome in these situations in order to respond to the needs of type 2 governance agencies.

The discussion in the fifth section of this article considers another of the dilemmas facing those involved in SDI implementation: i.e. the challenges presented by both internal and external organisational changes during the course of SDI implementation. To be successful SDIs must be sustainable over long periods of time. The current INSPIRE road map, for example, covers the period up to May 2019 (www.inspire.jrc.ec. europa.eu/). As a result the SDIs that are most likely to succeed will be those that meet the three sets of conditions originally identified by Campbell and Masser (1995, pp. 45-48) for effective GIS implementation. These are 1) a consistent strategy that identifies the evolving needs of users and takes account of the changing resources and values of the participants, 2) a commitment to and participation in the implementation of the SDI by groups and individuals from both type 1 and 2 types of governance structures and 3) an ability to cope with all kinds of change.

REFERENCES

Bregt, A., and J. Crompvoets (2004). Spatial data infrastructures: hype or hit? *Proceedings of GSDI 8*, Cairo, at http://www.gsdi.org.

Butler, D. (2006). The web-wide world. *Nature,* 439(7078): 776-8.

Campbell, H. and I. Masser (1995). *GIS and organisations: how effective are GIS in practice?* London: Taylor and Francis.

Craglia, M. and M. Campagna (2009). *Advanced regional spatial data infrastructures in Europe*, Ispra: Joint Research Centre Institute for Environment and Sustainability.

Craglia, M., M. F. Goodchild, A. Annoni, G. Camara, M. Gould, W. Kuhn, D. Mark, I. Masser, D. Maguire, S. Liang and E. Parsons (2008). Next generation Digital Earth: a position paper from the Vespucci initiative for the advancement of geographic information science, *International Journal of Spatial Data Infrastructures Research,* 3: 146-167.

Crompvoets, J., A. Bregt, A. Rajabifard and I.P. Williamson (2004). Assessing worldwide developments of national spatial data clearinghouses. *International Journal of Geographic Information Science*, 18(1): 1-25.

De Man, E. (2007). Beyond spatial data infrastructures there are no SDIs – so what, *International Journal of Spatial Data Infrastructure Research,* 2: 1-23.

Department of Sustainability and Environment, (2008). Victorian Spatial Information Strategy 2008-2010, Melbourne: Department of Sustainability and Environment.

Executive Office of the President (1994). Coordinating geographic data acquisition and access, the National Spatial Data Infrastructure, Executive Order 12906, *Federal Register,* 59: 17671-17674.

Gore, A. (1998). The Digital Earth: understanding our planet in the 21st century, *Australian Surveyor,* 43(2): 89-91.

Hooghe, L., and G. Marks (2003). Unravelling the central state, but how? Types of multi level governance, *American Political Science Review*, 97(2): 233-243.

Maguire, D. (2005). GeoWeb 2.0: implications for ESDI, *Proceedings of the 12th EC-GI-GIS Workshop*, Innsbruck, Austria. (http://www.ec-gis.org/Workshops/12ec-gis/).

Masser, I. (1999). All shapes and sizes: the first generation of National Spatial Data Infrastructures, *International Journal of Geographical Information Science,* 13(1): 67-84.

Masser, I. (2005). *GIS Worlds: creating spatial data infrastructures,* Redlands: ESRI Press.

Masser, I. (2007). *Building European spatial data infrastructures*, Redlands: ESRI Press.

Masser, I., A. Rajabifard and I. Williamson (2008). Spatially enabling governments through SDI implementation, *International Journal of Geographical Information Science,* 22(1): 5-20.

McLaughlin, J.D. (1991). Towards national spatial data infrastructure, *Proceedings of the Canadian Conference on Geographic Information Systems*, Ottawa, pp. 1-5.

National Research Council (1993). *Toward a coordinated spatial data infrastructure for the nation*, Mapping Science Committee, Washington D.C: National Academy Press.

National Research Council (1999). *Distributed geolibraries: spatial information resources*, Mapping Science Committee, National Research Council, Washington D.C: National Academy Press.

O'Reilly, T. (2005). What is Web 2.0: design patterns and business models for the next generation of software, at: http://www.oreillynet.com.

Rajabifard, A., A. Binns, I. Masser and I. Williamson (2006). The role of sub-national government and the private sector in future spatial data infrastructures, *International Journal of Geographical Information Science*, 20(7): 727-741.

Rajabifard, A., M. E. Feeney, I. Williamson and I. Masser (2003). National spatial data infrastructures, in I. Williamson, A. Rajabifard and M.E. Feeney (eds.), *Development of Spatial Data Infrastructures: from Concept to Reality,* London: Taylor and Francis.

Spatial Applications Division, Catholic University of Leuven (SADL) (2003). Spatial data infrastructures in Europe: state of play spring 2003. Summary report, at: http://inspire.jrc.it/state_of_play.cfm.

Thompson, B., M. Warnest and C. Chipchase (2003). State SDI development, in I. Williamson, A. Rajabifard, M. E. Feeney (eds.), *Development of Spatial Data Infrastructures: from Concept to Reality,* London: Taylor and Francis.

Vandenbroucke, D., K. Janssen and J. van Orshoven (2008). "INSPIRE State of play: development of the NSDI in 32 European countries between 2002 and 2007", *Proceedings of GSDI 10*, Trinidad, at: http://www.gsdi.org.

Cooperation – a Key Factor for Sustainable Spatial Data Infrastructure

E. Olof Olsson

KTH, the Royal institute of Technology, Stockholm, and Lantmäteriet, Gävle, Sweden
olof.olsson@lm.se

Abstract
Questions of cooperation are essential in the process of constructing and operating spatial data infrastructures (SDIs). However, even if cooperation in this field is often regarded as an important factor it is seldom a prime target for studies and research. This is in contrast to other fields where breakthroughs have been made with cooperation in focus. In this article, the concept of cooperation is explored in relation to SDI. Based on a literature study as well as experiences in the Swedish SDI, different stages of cooperation are presented and linked to SDI development. The research shows that research on cooperation in the context of SDI is gaining momentum and is very relevant to SDI development. The resulting problem analysis presents directions for further research.

Keywords: Spatial data infrastructures (SDI), cooperation, collaboration, coordination.

1. INTRODUCTION

Cooperation is an intensely discussed 'method' used for production and further management of spatial data. Complex interrelationships can often occur in the field of spatial data. Not only are the cooperating parties concerned but there are also other important actors. One, governments - and organisations like the European Union - stand for power and decisions and set many rules affecting the handling and use of spatial data. At the other end is the market which should ultimately set the agenda. In between there is some sort of association of organisations for which good and working cooperation is a critical factor for success.

Within the field of SDI, cooperation has often been treated as an important factor and a number of studies have touched different aspects of cooperation (see Clausen et al., 2006; Harvey, 2001; Onsrud and Rushton, 1995). However, there are few studies were cooperation is the main target of research. This is somewhat surprising, because in many other areas such as economy, biology and different social and organisational subjects, cooperation is given a lot of attention and some breakthroughs have been made during the last 20 years. A better understanding of the mechanisms behind cooperation within spatial data administration and the creation of national spatial data infrastructures can improve efficiency and lead to greater benefits within this area.

It is perhaps a paradox that the cold war after the Second World War boosted discussions and research about cooperation. Cooperation in this respect is often handled as an important step on the road away from conflict and war. Areas of conflict research and conflict handling are not the main scope of this study. However, cooperation and conflict handling are intertwined and a better understanding of how conflicts emerge and should be handled can also be a starting point for studies on cooperation. There are many roads that lead to questions of cooperation.

Studies of cooperation as a subject in itself, however, lead directly to Robert Axelrod. He is known for his groundbreaking work in game theory and complexity theory and has also written books that broaden the understanding of cooperation for both scientists and others (see Axelrod, 1984). Cooperation is an interdisciplinary subject and greater understanding of this field goes back to basic works y Axelrod and approaches to game theory by him and others.

Another well known name is Elinor Ostrom. She has made major contributions to the analytical literature on institutions and to the understanding of human cooperation. She has shown that neither the state or the market has been successful in solving common-pool resource problems. She also shows that common-pool problems are sometimes solved by voluntary organisations rather than by a coercive state. In her opinion the evolution of institutions for collective action is a worldwide working model with deep historical roots which combine long-term economic profits with ecological sustainability and social justice (see Ostrom, 1990).

In this article, the concept of cooperation is addressed through Axelrod's approach. Then, this approach is applied to SDI. Finally, three steps of SDI development with respect to coordination are identified and directions for further research are provided.

2. COOPERATION

There is no accepted general definition of the term cooperation but in dictionaries it is explained as "joint operation or action". Different subject fields have their own definitions. For instance in sociology, cooperation can be defined as 'activity shared for mutual benefit'. In people's minds cooperation often means 'working together for a common goal'. Within the field of spatial data we take that explanation as a starting point.

Cooperation is often associated with positive feelings and values. It would, however, be better if it was regarded as a 'tool' that can be used for both positive and negative issues. It is perhaps not clear for everyone that cooperation also can be associated with illegal activities - cartel formation for example.

Even at a very basic level it appears that cooperation is quite complex. Several conditions have to be fulfilled for cooperation to be a realistic alternative to not cooperating (Axelrod, 1984):

– The benefit gained by cooperation must be considerably greater than the alternative benefit gained by attacking/competing;
– The other party must be assumed to have a positive attitude towards cooperation;
– Each of the cooperating parties can only have a share of the total benefit. This share must be assumed to be big enough to motivate cooperation;
– "Cooperation demands that the future is important" (Axelrod, 1984). It is unlikely that the benefits of cooperation are gained immediately. A time lapse must be accepted by all parties and it has been shown that giving up quick rewards in favour of future benefits is something typically human.

It is clear that the study of cooperation is a truly interdisciplinary one. Knowledge and experience stem from a variety of subjects for instance biology, psychology, political science, economy and studies about conflicts and peace. One field, game theory, stands out and during recent years has contributed considerably to understanding the mechanisms of cooperation. Game theory is regarded as a part of the science of economics, often with mathematical theories in focus. The study of certain games in particular like the prisoners dilemma has contributed to new insights about cooperation.

General knowledge about cooperation is gained from studies using game theory (see Axelrod 1984), which show that cooperation within organisations often spontaneously

appears at low levels and can seek its way upwards. The other way is much harder. Cooperation initiated at top level often has great difficulty in penetrating downwards in the organisation. In short one could say that cooperation and organisations that are very hierarchical oriented seldom match very well.

From a theoretical standpoint, cooperation can be explained, by three characteristics: (1) reciprocity, (2) stability, and (3) robustness. Reciprocity is perhaps the most important quality of cooperation. It means that the cooperating parties have a mutual understanding and act in similar if not exactly the same ways. There is also an element of equality between the parties. Stability is the capacity to resist attacks from outside aimed at the destruction of the cooperation. Robustness is the capacity to grow in an environment that alters from time to time.

3. SPATIAL DATA INFRASTRUCTURES

SDI represents the important parts of a framework for effective production, handling and dissemination of spatial data within a nation or a region. In this context there is a frequent need of referencing not only to this framework (SDI) but also to the total spatial system or network that is held together and served by the SDI in question. This feature is referred to in this article as the 'SDI-based network'.

An SDI-based network is a very complex construction. It consists of a number of cooperating organisations with different goals and cultures that to some extent at least produce spatial data of common interest to society. Based on agreements, or law, they form a network and try to follow a set of standards and rules to fulfil an overarching goal benefiting society. Fulfilling these standards and rules and participation in constructing them can be seen as acts of cooperation. Practical implementation in particular often calls for advanced mutual planning and coordination of activities at the base level of the organisations concerned. This can be problematic and a possible explanation why even established SDI principles are difficult to realise in practise.

Relating SDI to the basic concept of cooperation raises two questions in particular:

1. Is the 'SDI-based network' in question *robust* enough to guarantee sustainable data access on defined terms?
2. Is the 'SDI-based network' in question *flexible* enough to adapt to new conditions in the future in terms of, for instance, technical standards and pricing policies?

These questions are closely related to the fact that an 'SDI-based network' does not exist in a vacuum but in close connection to the rest of society. This can be simplified by saying that we have two very strong 'magnets' pulling the 'SDI-based network' and with the ultimate capacity to change or even destroy the 'SDI-based network'. One magnet is 'the users as a whole'. The users represent the reason why the SDI network was constructed and are the primary beneficiaries. The other magnet is government, including the European Union, which sets many rules and lays down conditions for operating the system.

The 'SDI-based network' has to live within the 'magnetic field' between these two strong magnetic poles (see Figure 1). So, the SDI-based network is not only a complex construction in itself, but is also constantly subject to developments and changes in the 'magnetic field' between users and the government.

Figure 1: The 'SDI-based network' within a 'magnetic field'.
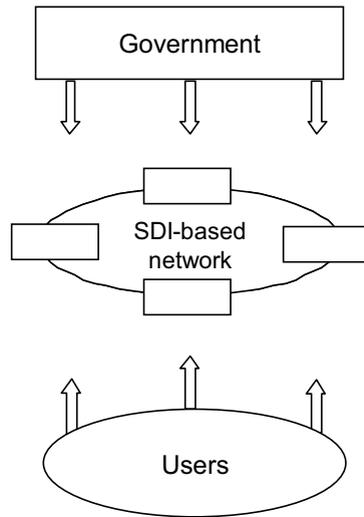
4. GENERAL AND SPECIFIC FINDINGS

Literature covering SDI as well as cooperation show quite a dispersed pattern but show interesting trends and developments over a period from about 1990 until today. Papers dealing with cooperation as the main subject are scarce but there are many papers that in some respect deal with issues of cooperation (see Campbell and Masser,1995; Huxhold and Levinsson, 1994; Craig, 2005). The way this has evolved is quite interesting. Issues of cooperation have often been treated as questions of data sharing (Carter, 1992; Onsrud and Rushton, 1995), information policies (van Loenen, 2006), interoperability (Cabinet Office, 2005), the role of governments and other actors, openness, harmonisation and benefits to society (Assimakopoulos, 2000; McDougall et al., 2005; van Loenen and Kok, 2004; van Loenen, 2006). It is evident that cooperation has been an important factor in the development of these aspects, but cooperation as such has often played second fiddle; an important factor but not a main target for discussion and research.

However, during the last decade one can notice a considerable change. There are still a few studies where cooperation is the main subject, but developments in the SDI context shifting from pure technical issues to questions of 'how to put the pieces together for maximum benefits' are evidently coming closer to questions of cooperation. We identified three distinct periods of SDI in which cooperation is addressed differently.

4.1 Findings since about 1990 – the first period

This first period focuses on questions about data sharing, technical standardisation, questions of project management, and simply how to use GIS and for what. Georgiadou (2006) characterised this period as "a period dominated by techniques and visionaries". This period approximately started around 1990 and lasted until 1997/98. Examples are Carter (1992), Campbell and Masser (1995), and Onsrud and Rushton (1995).

Many of the problems of a non-technical nature are typical only for a specific nation or a group of nations. This shows that historic background, traditions and nationally grounded cultures of handling problems in general have also had an impact.

4.2 The second period

The second period can be estimated from a few years before the millennium until today. This period is dominated by questions concerning:

− Networks, infrastructures, SDI concepts;
− Organisational issues;
− Information standards;
− The importance of words like awareness, trust, culture, benefits to society.

Georgiadou (2006) mentions periods that are socio-technical and multi-disciplinary. In this context we can, however, call this second period 'a period where information, knowledge and coordinated activities dominate the scene'. During this period many articles and reports have been published covering a broad area of subjects concerning SDI and even cooperation to some extent.

Questions of techniques, 'how to do' instructions and project reports have diminished in number and instead quite a broad variety of different new subjects has emerged. This happened parallel to tremendous technical developments – internet, new ways of organising databases, new media for storage and use and expanding international standardisation and policy declarations. The first period of technique euphoria and visionary expectations has turned into a more realistic working phase where many aspects and societal circumstances are taken into consideration – a growing understanding that new technical possibilities must have a lot of organisational, societal and other backup to work properly.

Questions of cooperation appear more frequently in the second period than in the first. Many authors find cooperation important in many contexts within this field (see Craig, 2005; De Bree and Rajabifard, 2005; Harvey, 2003). Some aspects are described and one can notice efforts to define cooperation (in contrast to collaboration and coordination) (see Clausen et al., 2006). However, it is difficult at this stage to find literature devoted to cooperation and penetrating this subject in depth. Late and interesting exceptions are Hörnemalm (2008) and Thellufsen (2008).

These findings together indicate that there is a need to study questions of cooperation specifically in the context of SDI and that this could be valuable in increasing understanding within this area. However, this should also be seen in the light of what might happen in the future. In other words, are we now leaving the second period and entering a third? We believe so and this will call for increased attention to questions of cooperation.

4.3 A third period?

There are indications that we are about to leave the second period and enter a third. One is that the movement from 'technical' questions to a variety of 'new', often complex questions described above is accelerating. Another is an underlying movement in society itself. A third is the ongoing general development of organisations studied in organisational research.

Questions of SDI are strongly integrated with development of society itself. So, the question of whether we now are entering a new period in the context of SDI depends on how society will evolve. Describing it in terms used by the analysts 'Paradigmmaklarna' (http://www.paradigmmaklarna.com, see Table 1) would mean the following. The dominating metaphor in society today is 'knowledge' but according to them it will soon change to 'understanding'. The metaphor for the earlier society (1850 and onwards) could be named 'energy'. Some key words for these different paradigms further clarify it.

Table 1: Paradigms and key words.

Energy	Knowledge	Understanding
Muscles	Brain	Communication
Technique	Education	Network
Resources	Patent	Concept
Stationary	Long termed	Short termed and iterative
Generalist	Specialist	Complex
Materialist	Functionalist	Ruled by values
Collective	Individual	Context

Although all predictions about the future are largely speculation they can give a valuable indication. If this prediction is correct, this change is clearly in favour of SDI and cooperation. The words under the heading 'Understanding' are familiar to those used in connection with SDI and cooperation. This can be an indication that cooperation which has received more attention during the second period than in the first will be even more important as a concept in the future.

However, many of the 'new' non-technical questions are discussed and explored solely within the 'SDI community' and to a lesser extent in contact with other fields of research and development. For instance, the SDI community does not necessarily have the best experts on organisational issues and should, therefore, seek help in promoting cooperation. There is at least one interesting study in this context. Koerten (2007) analysed literature about GIS and SDI using organising metaphors. He concludes that the literature in this field is dispersed and lacks clear focus. Koerten relates to paradigms and metaphors in organising theory. He sees the need for a shift from the dominating 'functionalist paradigm' to the 'interpretationist paradigm' which represents revolutionary science and a shift from the objective realm to a subjective realm. He welcomes encounters with the interpretationist paradigm. He states that what is considered as a revolutionary paradigm for GIS and SDI is already orthodox in organisational science. He concludes that we need more knowledge about organising, interorganisational networks and cooperation strategies. This, he states, calls for unorthodox researchers, willing to use theories that focus on what goes on in people's minds rather than focusing on organisational structures.

Koertens study seems to support and strengthen the indication above that cooperation will be of more importance in the field of SDI. It also gives a clear hint that new ways of doing research are essential and that the 'SDI community' should cooperate better with other fields, for instance organisational research.

4.4 Important aspects

As cooperation is a vast and somewhat undefined subject it is important to focus on decisive aspects. From a theoretical standpoint some guidance can be found in Brunsson (2002). Brunsson is known for his way of simplifying processes in organisations to only three stages: Talk, Decisions and Action (see Figure 2).

In the SDI world information policies, national strategies, and plans should be categorised as 'Talk' according to Brunsson. 'Talk' should then be followed by 'Decisions' then concrete 'Action'. This is what is supposed to happen in large projects and political processes. According to Brunsson there is no straight line from talk to action. Instead these three stages tend to live there own 'life'. Things circulate and happen within each stage but there is little communication between them. There is of course some communication top-down, but not so much as is generally believed and there is certainly not much communication bottom-up. Communication between these layers in the process is generally rare.

Figure 2: Talk, decisions and action are generally not compatible but worlds of their own (Brunsson, 2002).

Talk

Decisions

Action
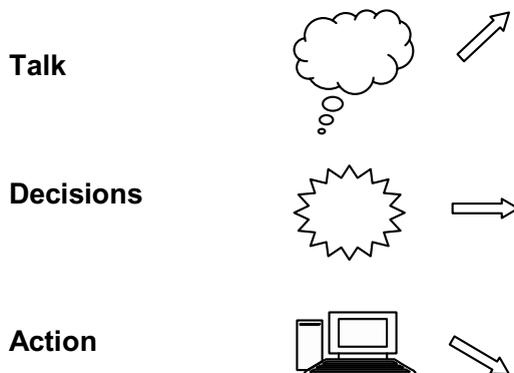
For SDI in particular cooperation should take place in all three layers and also exist between the layers (see Figure 3). In the 'Talk' layer there are already many people working with strategies, plans and policies and doing valuable research where even aspects of cooperation are handled. In the 'decision' layer there is a lot to gain from organisational research and experience. In the 'action' layer, however, it seems that the actual process when people meet (what is happening in their minds and what is the outcome) is not particularly well studied and evaluated in terms of cooperation. It would also be interesting to see if cooperation or at least communication between the layers 'decision' and 'action' can be improved.

Referring to the periods 1, 2 and (perhaps) 3 a trend where cooperation seems to be a concept affecting organisations and people engaged in questions of SDI in more 'depth' can be seen. In addition, problems of 'data sharing' and general agreements between the managements of different organisations have changed to practical daily cooperation in the running of national SDIs. Using Brunssons terminology this is a development from 'Talk' to 'Action'.

This leads to the conclusion that within the area of GIS and SDI can be especially interesting to study:

– Cooperation within the 'Action' layer, and
– Cooperation between the layers 'Decisions' and 'Actions'.

Figure 3: Prime focus on actions between people in different organisations – not on the prerequisites.
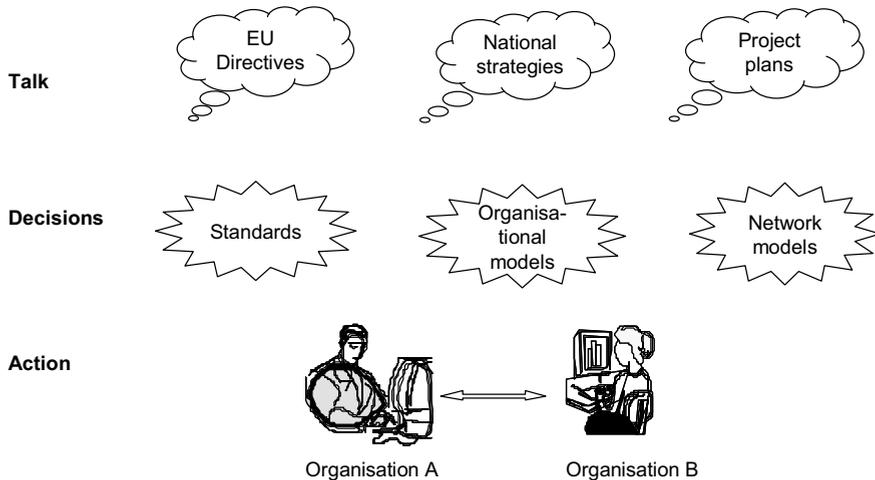
5. SWEDISH EXPERIENCE

In Sweden the handling of, for instance, cadastral information and information about physical plans and regulations depends on cooperation between central governmental authorities, local (municipal) authorities and private firms. Large projects have been carried out during the last decade presupposing working cooperation, for example the new national cadastral index map.

Although projects in this field have been at least partly successful, some severe problems have occurred. These related to time schedules, costs and daily communication between the participating organisations. This happened despite good planning and the required prerequisites in place (strategies, written agreements, project plans, technical solutions, financing, key personnel). It has led to the suspicion that existing problems are complex and relate problems of cooperation.

From a Swedish perspective it is important to have better knowledge of how to deal with problems of this nature. National strategies point to large future projects based on SDI and new types of services where cooperation between growing numbers of organisations can be a decisive factor. Projects already carried out provide a good base for case studies using the concepts of cooperation broadly outlined in this article.

6. CONCLUSION

Results so far indicate that cooperation as a concept is generally not fully understood within the SDI context. In many cases cooperation can be a decisive factor for realising efficient 'SDI-based networks'. There are indications that effective cooperation at 'base

level' within and between organisations can be an important factor for success. The following conclusions have been reached:

1. Cooperation as a general concept has been studied quite intensely during the past decades and a number of interesting results have been published.
2. In the field of SDI and adjoining areas cooperation has gained much attention but it is seldom the prime target for research and development.
3. Experience from many projects and activities connected to SDI points to a lack of understanding what cooperation is really about and how cooperation should be implemented in daily practice.
4. Since about 1990 different periods related to handling of issues of cooperation can be distinguished. In the period which we are probably about to enter, questions of cooperation seem to be especially decisive.
5. For the establishment and running of 'SDI-based networks' there are indications that it is important to study questions of cooperation at 'base level' (in daily work) within and between organisations.

With this in mind three questions related to relevant trends and experience have been formulated, which are addressed in the following steps of this research.

1. 'Sustainability' is an important quality of an SDI-based network. What is the relation between this quality and cooperation?
2. Sharing and supplying data in a distributed network is discussed more and more. What is required in terms of cooperation of a distributed network?
3. Cadastral information is considered to be key information in many business activities with reference to spatial data and in Sweden a lot of relevant experience stems from handling cadastral information. Can experience from this field be of value in a broader perspective?

Previous discussion and formulation of problems can also be summarised by stating that the following aspects of cooperation are of special interest for further study:

- Cooperation within organisations participating in the 'SDI-based network' (network of organisations);
- The importance of cooperation in the process of introducing new members in the 'SDI-based network', and
- The attitudes and capabilities, in terms of cooperation, of forerunners.

With the general background about cooperation presented in this article the Swedish SDI development will be analysed with the framework and concepts from the 'talk-decision-action' concept in particular.

REFERENCES

Alsén, L. and T. Troedsson (2002). "Don´t Panic" Paradigmmäklarna I Sverige AB (www.paradigmmaklarna.com), Sweden.

Assimakopoulos, D. (2000). Social network analysis as a tool for understanding the diffusion of GIS innovations: the Greek GIS community. *Environment and Planning B: Planning and Design,* 27 (4): 627-640.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books, Inc.

Brunsson, N. (2006). *The organization of Hypocrisy. Talk, Decision and Actions in Organizations*. Copenhagen Business School Press.

Cabinet Office, e-Government unit, Great Britain (2005). Geographic information: An analysis of Interoperability and information Sharing in the United Kingdom. Cabinet Office.

Campbell, H. and I. Masser (1995). *GIS and organizations. How effective are GIS in practise?* London: CRC/ Taylor & Francis.

Carter, J.R. (1992). Perspectives on Sharing Data in Geographic Information system. *Photogrammetric Engineering & Remote Sensing*, 58(11): 1557-1560.

Clausen, C., Rajabifard A., Enemark S., Williamsson I. (2006). "Awareness as a foundation for developing Effective Spatial Data infrastructures". *XXIII FIG Congress*, Munich, Germany.

Craig, W. (2005). White knights of Spatial Data Infrastructure: The role and motivation of key individuals. *URISA Journal,* 16(2): 5-13.

De Bree, F., and Rajabifard A. (2005). "Involving users in the process of Using and Sharing Geoinformation within the context of SDI-initiatives", *FIG Working Week 2005 and GSDI-8* Cairo, Egypt April 16-21, 2005.

Georgiadou, Y. (2006). SDI ontology and implications for research in the developing world. *International Journal of Spatial Data Infrastructures*, 1: 51-64.

Harvey, F. (2001). Constructing GIS: Actor Networks of Collaboration. *URISA Journal,* 13(1): 29-37.

Harvey, F. (2003). Developing geographic information infrastructures for local government: the role of trust. *The Canadian geographer,* 47(1): 28-37.

Huxhold, W. and A. Levinsson (1994). Managing geographic Information System projects. Oxford University Press.

Hörnemalm, J. (2008). "Samverkan är ett magiskt ord" – Motstridiga ambitioner och ideal I nätverksorganisering (PhD) Luleå Technical University, Sweden.

Koerten, H. (2007). "Blazing the trail or follow the Yellow brick Road? On geoinformation and organizing theory". In F Probst & C Kessler (Eds.), *GI-Days 2007-Young researchers forum*, Munster: IFGI, pp. 85-104.

McDougall K., Rajabifard A., Williamsson I. (2005). Understanding the motivations and capacity for SDI-development from the local level. *FIG working week 2005 and GSDI-8* Cairo, Egypt April 16-21, 2005.

Onsrud H.J. and G. Rushton (eds.) (1995). *Sharing geographic information,* New Brunswick, New Jersey: Center for Urban Policy Research.

Ostrom E. (1990). *Governing the commons. The evolution of Institutions for Collective Action*. Cambridge University Press.

Thellufsen C. (2008). Awareness – a tool for investigating inter-organizational collaboration in land administration systems? (PhD) Aalborg University, Denmark.

van Loenen, B. (2006). Developing geographic information infrastructures. The role of information policies (PhD). Delft: DUP Science.

van Loenen, B. and B. C. Kok (eds.) (2004). *Spatial Data infrastructure and policy development in Europe and the United States*. Delft: DUP Science.

Seamless SDI Model – Bridging the Gap between Land and Marine Environments

Sheelan S. Vaez, Abbas Rajabifard and Ian Williamson

Centre for Spatial Data Infrastructures and Land Administration
Department of Geomatics, University of Melbourne, Australia
s.sheikheslamivaez@pgrad.unimelb.edu.au

Abstract

With climate change, rising sea levels pressing harder year on year and the need to manage our resources more carefully in this dynamic environment, the inability to integrate land and marine base information is an increasing problem in many countries. The absence of a seamless spatial data framework prevents the execution of standard practice of locating and referencing spatial data across the land-marine interface where so much pressure and development is taking place. There is a growing and urgent need to create a seamless SDI model that bridges the gap between the terrestrial and marine environments, creating a spatially enabled land-sea interface to more effectively meet sustainable development objectives. This article discusses drivers for integrating land and marine environments and proposes a seamless SDI model as an abstract level SDI and its associated components. This is followed by issues and challenges that must be overcome in developing an overarching architecture for a seamless SDI that allows access to and interoperability of data from marine, coastal and terrestrial environments.

Keywords: Spatial Data Infrastructure (SDI), Seamless SDI, Interoperability, Land-sea interface.

1. INTRODUCTION

Due to the high economic value of coastal and marine activities, and to the social value of coastal zones for quality of life, managing the coastal zone is a key component of the socio-economic framework in most nations with coastlines. In recent times several natural disasters hit some part of the coastal areas around the world in particular small islands and archipelagic countries causing hundreds thousands of people lost their lives, while those who survived had lost their properties. Moreover anthropogenic global warming will undoubtedly cause substantial sea-level rise and shoreline movement during this century and beyond. Current climatic models predict a global rise in sea level of six meters or more due to climate change. This will affect the rights, restrictions and responsibilities of both governments and individuals who own or manage land along the coastal strip. This is especially problematic in some pacific island states that may be wholly inundated with a six meter rise.

Learning from such kind of devastating disasters, it is important to have accurate, complete and up-to-dated spatial data resources and services of coastal area for better development planning and timely disaster management. Spatial data infrastructure (SDI) as an enabling platform can facilitate access and integration of different datasets from different disciplines. In addition, sustainable development also requires integrated and comprehensive spatial data throughout the country both land and marine area, that can easily be found and accessed by the public.

However, current SDI design focuses mainly on access to and use of land related datasets or marine related datasets. Most SDI initiatives stop at the land-ward or marine-ward boundary of the coastline, institutionally and/or spatially. Consequently, there is a lack of harmonised and universal access to seamless datasets from marine, coastal and land-based spatial data providers. This leads to inconsistencies in spatial data policies, data creation, data access, and data integration across the coastal zone. The extension of a national SDI covering the land and marine environments on a seamless platform would facilitate greater access to more interoperable spatial data across the land-marine interface enabling a more integrated and holistic approach to management of the coastal zone.

This article aims to discuss the drivers for integrating land and marine environments and the potential for adding a coastal dimension to an SDI to facilitate coastal zone management. It also proposes a seamless SDI model as an enabling platform to increase the efficiency and effectiveness of management across regions and disciplines followed by an introduction to its associated components and specifications. Finally it highlights the issues and challenges that must be overcome in developing an overarching architecture for a seamless SDI that allows access to and interoperability of data from marine, coastal and terrestrial environments.

2. DRIVERS FOR INTEGRATING LAND AND MARINE ENVIRONMENTS

The integration of land and marine base information is an increasing problem in many countries. This especially applies to archipelagos where seawater is the 'bridge' connecting islands. While most of the countries are aware the problem of disconnected land and marine information, few have committed to resolving the problem (Murray, 2007). This is partly due to complexity as it requires two or more organisations and users to identify and address the key issues. The ability to access and integrate data has been identified as a problem by people involved in coastal zone management. Also the development of Integrated Coastal Management (ICM) initiatives has encountered similar problems (Strain et al., 2006). However, the primary drivers for land and marine integration can be categorised into societal and commercial and technological drivers.

2.1 Societal Drivers

The coast as the interface between the land and marine is a unique geologic, ecological and biological domain of vital importance to a vast array of terrestrial and aquatic life forms-including humankind. The importance and value of the coastal zone can not be underestimated. Since early settlement days the coastline has been used in many ways. Largely for transportation reasons, major industrial and commercial centres developed around port cities. Some two-thirds of the planet's population lives in a narrow 400-kilometre coastal band. Demographic trends suggest that coastal areas around the world are undergoing serious population growth pressures. Population growth is the driver behind many, if not most, coastal problems (Brower et al., 2002). This puts more pressure on the land-marine environment through greater demand for development and the resulting increase in effluent and pollution. These problems can no longer be viewed in isolation. There is a need for connectivity and replacing a fragmentation with a collaborative, integrated approach (Toth, 2007).

Society is now using resources and producing wastes at rates that are not sustainable. Oceans and the coastal zone have been used as dumping grounds for many years. For instance population increases along Australia's shorelines and the corresponding industrial development has resulted in a rapid increase in sewage outflow into rivers, es-

tuaries and oceans (Plunkett, 2001). Land-based sources of marine pollution account for around 80% of contamination in the marine environment (SOEAC, 1996). Environmental problems have to be addressed globally.

The Intergovernmental Panel on Climate Change (IPCC) has calculated that by the end of this century, sea levels could rise by up to 89 centimetres and temperatures could rise by between 1.4° C and 5.8°C (http://www.ipcc.ch/). Consequently climate change and global warming are a serious threat to coastal areas requiring greater attention to coastal protection and change management. Other drivers are cost and time efficiencies, public expectations coupled with greater awareness and focus on temporal issues and policy drivers such as the European Union Water Framework Directive (2000/60/EC) or other legislation concerned with limiting adverse impact of natural forces and processes.

2.2 Commercial Drivers

The coastal zone is one of the most productive areas accessible to people. However, there are increasingly serious signs that economic uses of our coast are undermining their long term sustainability. For example, overfishing is exhausting and depleting fisheries around the world. In Australia, according to the Bureau of Rural Sciences (BRS, 2002), 11 target species in Commonwealth fisheries were classified as overfished, 11 as fully fished and a further 35 classified as 'uncertain', despite the highly regulated and regarded best-managed fisheries in the world. This overfishing came about partly due to lack of knowledge of the distribution, abundance and biology of the stocks, but also due to inadequate management arrangements resulting in unsustainable catches (NOO, 2002). Additionally, production of offshore oil and gas is declining due to depleting resources. The protection of marine ecosystems and fishery resources can not be tackled by individual eco-systems. There is an economic and social need to manage, explore and exploit the nation's ocean territories in a way that will maximise benefit, while protecting the ocean environment.

2.3 Technological Drivers

Seamless discovery and seamless use are two main user aspirations. The user would like to be able to search widely, at different levels and access all that exists. This entails the needs for agreements in terms of data descriptions, metadata definitions, protocols, data access and sharing policy. Also the user would like to identify easily the data available and to find easily what fits the purpose and download it directly to their analysis software. Figure 1 illustrates major marine and coastal management issues and challenges and some of their potential impacts.

As the interface between marine and terrestrial environments, coasts have diverse and ever increasing conflicting pressures and demands requiring effective administration and management. In spite of this, current marine and coastal zone management systems are neither effective nor sustainable (Thia-Eng, 2003; Neely et al., 1998). There is a need to make the land and marine infrastructures interoperable so that planning, management and solutions can be identified in a seamless and holistic way.

3. SEAMLESS SDI

An essential requirement for the consistent and effective management of the oceans is reliable, comprehensive and accurate spatial data. The notion that considerable benefits accrue to a society by 'freeing up' access to spatially referenced data has provided

Figure 1: Marine and coastal management issues.
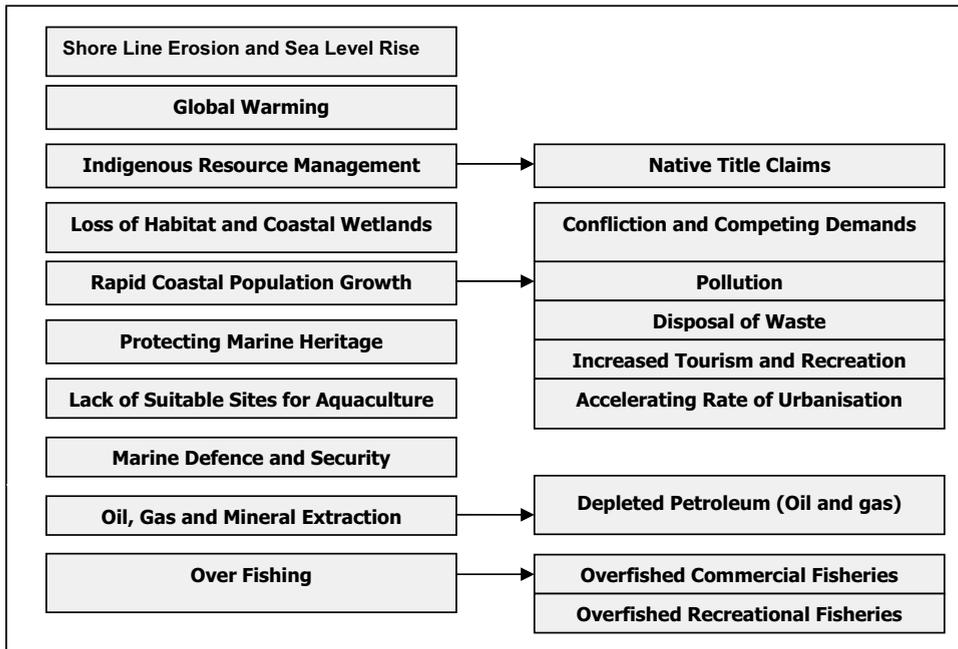
impetus for the construction of local, national, and global spatial data infrastructures (SDIs) (Rajabifard et al., 1999; Rhind, 2001). SDIs theoretically comprise networked, spatially-enabled databases or datasets that are accessible for downloading or manipulation using contemporary technologies, usually according to explicit institutional arrangements and are supported by policies, standards, and human capital (Rajabifard and Williamson, 2001; Nebert, 2004). However, the development of SDIs are confined to the landward or seaward sides of the coastal zone, with little or no thought given to the interaction between these two environments. The reality is that the need for access and coordination of spatial data does not stop at the coastline. Many coastal management issues could be overcome if a spatial data platform that enables a holistic, integrated and coordinated approach to spatial data for decision-making existed.

The complex physical and institutional relationships existing within the coastal zone make it impossible to develop a marine SDI in isolation from land based initiatives. Furthermore a seamless infrastructure aids in facilitating more integrated and effective approaches to coastal zone management, dealing with problems such as marine pollution from land based sources (Williamson et al., 2004). A seamless infrastructure was endorsed by the UN as part of the International Workshop on Administering the Marine Environment (see Rajabifard et al., 2005). It was recommended that a marine cadastre act as a management tool within a Marine SDI (MSDI) as an extension to National SDIs across Asia-Pacific (Figure 2).

In November 2005, the International Hydrographic Organisation (IHO) has organised and conducted a seminar on "The Role of Hydrographic Services with regard to Geospatial Data and Planning Infrastructure". This seminar recognised formally an option for Hydrographic Offices to become responsible or partner in national MSDI and the possible connection of Marine SDI to the National SDI (IHO, 2005). A resolution of the

Figure 2: Seamless administration system (UN-PCGIAP, 2004).
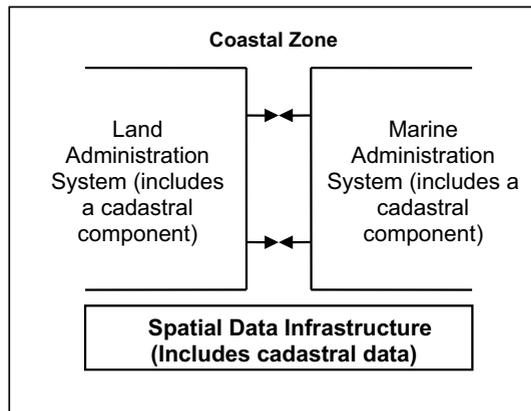
17th United Nations Regional Cartographic Conference for Asia and the Pacific (UNRCC-AP) further supported the inclusion and development of a marine administration component as part of a seamless SDI to "ensure a continuum across the coastal zone" (UNRCC-AP, 2006).

On land, issues and challenges such as data interoperability and data integratability have been identified as major issues. However, there are more issues facing marine environment as it is highly dynamic with 4D boundaries. Thus natural resources or features are more likely to move with time which leads to poor accuracy, precision, consistency and completeness of marine spatial data. These difficulties compound in the coastal zone, as it is both the on- and offshore environments combined and interrelated. The Port Philippe Bay (PPB) case study identified land, marine and coastal management issues. Several interviews with organisations involved in the management of PPB were interviewed. Figure 3 shows the conceptual demonstration of issues and challenges of the land, coast, and marine environments. It implies the need for an overarching spatial data framework to facilitate the management of the whole environment.

To improve management of the coastal zone, there needs to be access to and interoperability of both marine and terrestrial spatial data. A more integrated and holistic approach to management of coastal and marine environments would be facilitated by the extension of the SDI on a seamless platform. This would promote data sharing and communication between organisations and facilitate better decision-making. Based on the spatial hierarchical reasoning and object oriented modelling method the seamless SDI model can be postulated as one abstract class SDI at the higher level (parent level) and can be used as a super-class for marine SDI and land SDI classes that extend the abstract class while both land and marine SDI class would inherit seamless SDI properties, they continue to have their specific characteristics and components at the same time (Figure 4).

Just like abstraction is closely related to generalisation, the inheritance is closely related to specialisation. The specialisation and generalisation relationships are both reciprocal and hierarchical. Seamless SDI generalises what is common between land and marine SDI, and they specialise seamless SDI to their own specific subtypes. Figure 5 illustrates a conceptual view of seamless platform architecture.

Figure 3: Issues and challenges of the land, coast and marine environments.

Coast

Issues
- Complex physical and institutional relationship
- One of the more hazardous region
- Conflicting uses, activity and interests
- Contribution to the social and economic development
- Integration of oceanic and land-based databases
- Inherent interrelationships between marine/coastal data and data covering in-land regions
- Data gaps over the coastal zone

Challenges
- Harmonised and universal access to oceanic, coastal and land-based spatial data
- Capacity building, funding
- Security and privacy issues
- Encouraging cooperation and creating a culture for spatial data sharing
- Develop the national Coastal SDI as a subset of national SDI

Land

Issues
- Data interoperability
- Immature institutional arrangements
- Data integratability

Challenges
- Copyright, ownership privacy and licensing
- Pricing and cost recovery

Marine

Issues
- Highly dynamic with 4D boundaries
- Lack of framework for accessing and sharing marine spatial data
- Poor accuracy, precision, consistency and completeness
- No spatial descriptions for legislation and various boundaries
- Lack of Metadata
- Wireless data transfer
- Complex spatial and temporal interactions
- Immature institutional arrangements

Challenges
- Building partnerships
- Privacy and sensitivity

Figure 4: Seamless SDI model.

Seamless SDI

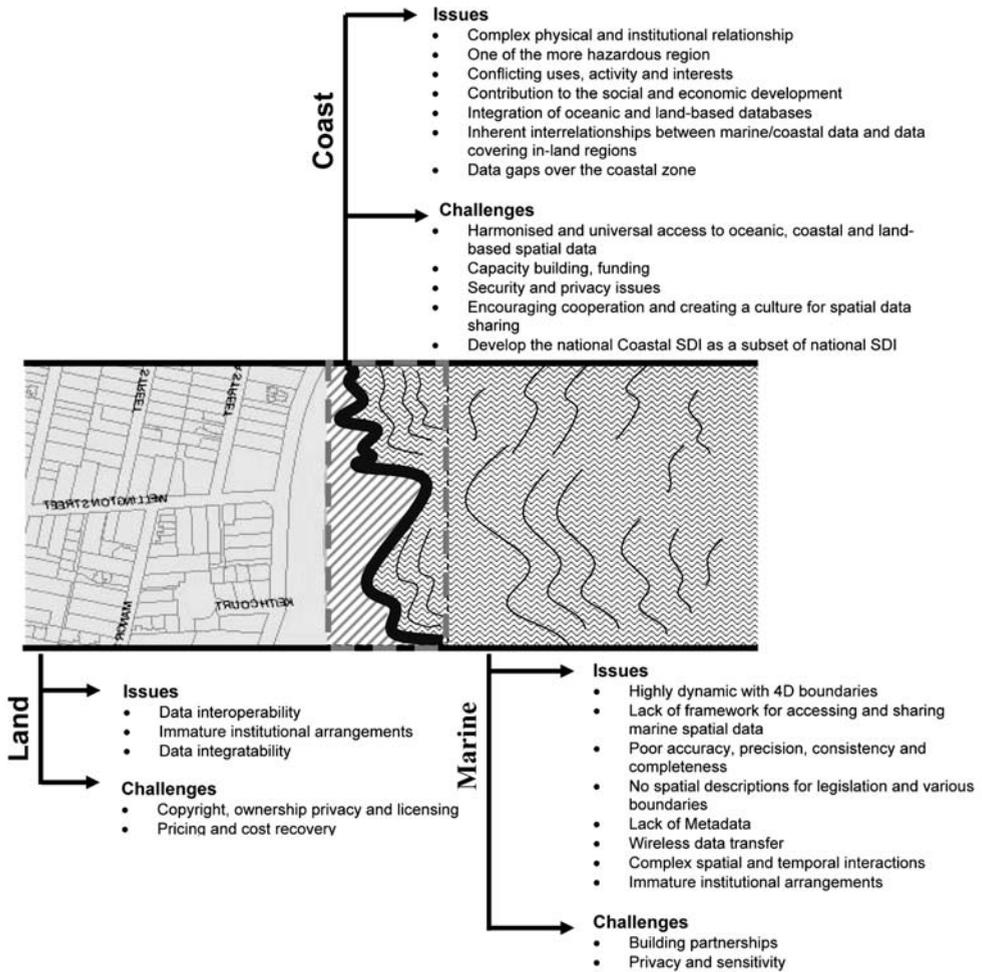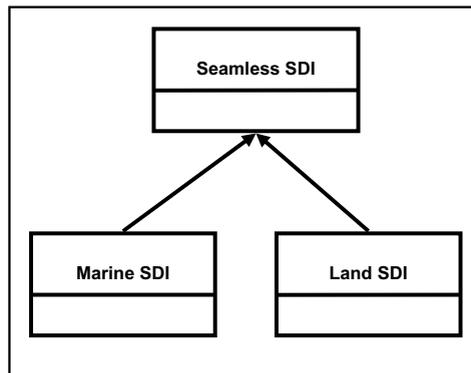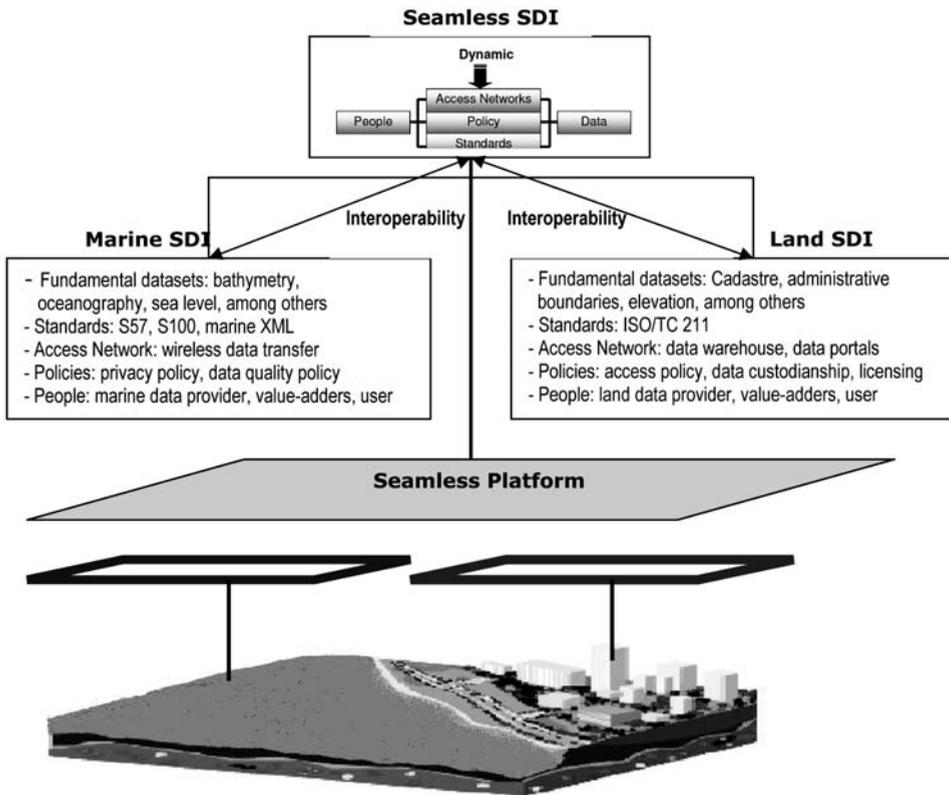Marine SDI   Land SDI

Figure 5: Seamless platform.

A seamless SDI should have the following characteristics:

- Seamless: the digital spatial data is stored continuously throughout and across juris-dictions;
- Multi-purpose: the same data can be used for different purposes;
- Multi-users: the same data can be accessed by different users concurrently, and
- Interoperable: the data stored in the database can be accessed using different GIS software and applications.

A seamless SDI platform would enable the utilisation of common boundaries across the coastal zone to ensure no ambiguity exists and no areas are unaccounted for over the coastal interface. This infrastructure will become a powerful information resource for managers in fields as varied as fisheries habitat management, pollution monitoring and control, shoreline erosion, weather forecasting and tourism development. The informa-tion derived from such a fully integrated information infrastructure will facilitate im-proved decision making at all levels.

4. SEAMLESS SDI COMPONENTS

The SDI concept has been used to describe land related spatial data and recent initia-tives such as marine SDI, marine cadastre and marine spatial planning have all

emerged in response to a global realisation of the need to improve management and administration of the marine environment. Figure 6 shows the components of SDI that link people to data: the standards, policies and access networks.

Figure 6: Components of SDI (Rajabifard, 2002).
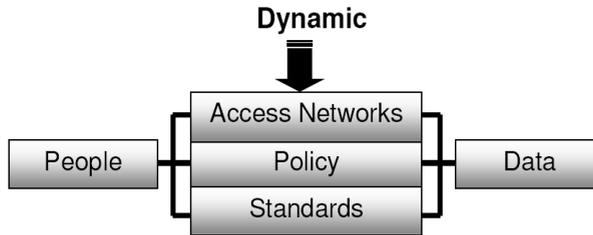
This section examines each component of SDI (data, standards, policies, access networks and people) and discusses its applicability to seamless SDI. It is important to note that the concept is dynamic, in that it provides an ability to be updated with changing technology or human attitudes or with the need for including new environments.

4.1 Fundamental Datasets

The lack of accurate information seamlessly crossing the land-sea interface creates a serious obstacle for coastal zone managers. These managers need precise, accurate, and timely data and products that are easily accessible and usable for a wide variety of applications. In the land environment an SDI includes 'fundamental datasets', those that will be needed to support most business processes, with a designated custodian responsible for managing them. The seamless SDI model as an infrastructure at the higher level needs to cover all the fundamental datasets from land, marine and coastal environments.

This aligns with the INSPIRE Directive consisting of 34 spatial data themes required to successfully build environmental information systems. The integration of land and marine data is applicable to a number of themes in Annex I-III across the land and marine environments such as the elevation, hydrography/hydrology, transport networks, protected sides, buildings, land use, oceanographic geographical features, utility information, addresses, and geology. Other relevant themes are: environmental monitoring facilities, area management, natural risk zones, sea regions, bio-geographical regions, habitats and biotopes, species distribution and energy resources.

Furthermore, the IHO Marine SDI Working Group (MSDIWG) defined marine SDI as the component of national SDI that encompasses marine geographic and business information in its broadest sense covering sea areas, inland navigable and non-navigable waters. This would typically include seabed topography, geology, marine infrastructure (e.g., bathymetry, wrecks, offshore installations, pipelines and cables); administrative and legal boundaries, areas of conservation, marine habitats and oceanography.

In some countries like USA, national SDI bathymetry is a sub layer of the elevation fundamental dataset. Also INSPIRE Annex III elevation dataset includes bathymetry and shoreline. This may be possible for other datasets. However, it is likely there will be datasets that are fundamental only for the marine environment (i.e. salinity, waves, and water quality).

4.2 Standards

SDI must be based on interoperability (seamless databases and systems). Standards are used to ensure interoperability and integratability of different datasets (Strain et al., 2006).The implementation of spatial standards at national level will assure that every institution and organisation creates spatial data in the same manner and it will ease spatial data sharing and exchange. These must be developed using international procedures and practises to cover not only the national needs, but also cooperation at an international level. In this respect the IHO has an important role to play in developing the appropriate standards needed for its hydrographic and cartographic applications, in close cooperation with appropriate organisations responsible for standardisation, such as ISO.

The development of S-100 (the next edition of S-57) has been a great step toward creating a seamless SDI. S-57 standard, although limited in scope and implementation, provides important compatibility for data sharing in the hydrographic information community.The next edition of S-57 standard will not be a standard just for hydrography, but will have manageable flexibility that can accommodate change and facilitate interoperability with other GIS standards. It will also allow hydrographic offices to use other sources of spatial data. S-100 is being based on the ISO/TC211 base standard and will make provision for imagery and gridded data in addition to the existing vector data, defined in the present version. This will facilitate the development of additional products and services other than for navigation purposes (Maratos, 2007). It also plays a key role for IHO and hydrographic offices in any marine SDI development. Therefore common standards and well documented metadata are essential for data discovery, management and compatibility within an SDI.

4.3 Policies

Other issues also need to be considered, including the need for harmonised data access policies and exploitation rights for spatial data, data custodianship, conformity, quality, content, industry engagement, avoidance of duplication and sensitivity. These policies for terrestrial spatial data and marine and coastal spatial data are likely to differ in terms of data quality, data access and privacy. Privacy over spatial data in the marine environment is a concern with many countries reluctant to share spatial data relating to their marine jurisdictions. As such there may be a need to maintain the different privacy policies for offshore data (Bartlett et al., 2004).

Appropriate policy and governance models could assist SDI development in several ways: by stimulating more rapid evolution of SDIs, by addressing current deficiencies in the application of standards, and by helping to achieve an increase in public penetration of SDI related technology and services through more tightly integrating a userperspective in both SDI design and operational management. Therefore there is a need for an appropriate policy model to create a seamless infrastructure across jurisdictions.

4.4 Access Networks

Decisions affecting coastal environment need to be timely and based on a strategic interpretation of all available data, presented in an easy and accessible format. Access networks usually comprise data warehouse, data portals, one-stop shops, on-line atlases or similar. For the access network to support interoperable and coordinated data they must comply to SDI standards and policies.

The OGC/TC 211 implementation specifications have deficiencies particularly in relation to manipulating marine data types which typically have 3 or 4 dimensional components (e.g. latitude, longitude, depth, and/or time). For instance, based on the Australian marine SDI activities, it was difficult to deal with the time dimension in OGC Web Map Services (Finney, 2007). A lack of reference implementation for combinations of specific standards is problematic for communities that need to implement these international standards.

4.5 People

Developing an agreed interoperable framework requires organisational collaboration and a clear use case and applications addressing interoperability cross borders and cross sectors (land-marine) scenarios. An overarching framework is supporting data policies, data access, data specifications (datum, feature catalogue) and standard implementation.

 An international workshop for land and marine integration in March 2007 identified the need for a single body to support land-marine integration for the region to keep the land and marine communities working together was noted (http://www.eurosdr.net). However, many issues and challenges could be overcome through better coordination arrangements and existence of a single management authority or forum for collaborative planning, and deficient legislation. More information about seamless SDI is required to have a better understanding and knowledge about SDI among different institutions and organisations and there should be proper regulation to enforce that all spatial data providers should involve in and contribute to the development of a seamless SDI.

5. CHALLENGES IN CREATING A SEAMLESS SDI MODEL

In order to create a seamless SDI across terrestrial and marine environments and jurisdictions, it is important to recognise and accept that building and maintaining an SDI is not easy, even for well-developed states. It is a dynamic and complex process at different levels of government and requires research and collaboration with academia and private industry.

Sustainable development requires an integrated spatial data system which provides built and natural environmental datasets that are available to the public. The integration of spatial data at national level encounters several problems either of technical, institutional or policy nature.

5.1 Technical Issues

Spatial data may come from various sources or data providers. Each data provider has its policies and methods of managing spatial data. Often, land and marine data products are incompatible in terms of scale, projection, datum and format (Gillespie et al., 2000). Several technical issues that should be taken into consideration when integrating spatial data from various data sources are: differences in spatial reference system (horizontal datum, vertical datum, and coordinate system), storage format, scale of data source, feature or object definition (feature catalogue), spatial data quality due to the differences of resolution or data acquisition method and finally differences in spatial data modelling (geometry, features name, attributes, field type, topology) (Syafi'l, 2006). In the MOTIIVE project these problems were also recognised by coastal managers in Europe regarding data. They added the lack of metadata and correspondingly difficulties to discover data (see http://www.motiive.net).

One more concern linked to the establishment of seamless SDI is the issue of a national shoreline. As the fundamental boundary for so many applications and studies, the lack of a consistently defined shoreline has frustrated coastal zone managers, planners, and scientists for many years. Different representations of the coastline in marine and land datasets leads to data overlaps while most of the applications require a single seamless layer with no duplication of common features. Table 1 shows an example of the differences on several aspects of two main data sources (Topographic Map and Nautical Chart) of Australia that should be considered when integrating land and marine spatial data.

The lack of standardisation and guidance for data and metadata and associated publishing protocols is the main problem of the above differences. Each organisation creates spatial data for its own purposes using their own technical specification without considering that the data may be shared or distributed to other communities.

Table 1: Different aspects of land and marine spatial data integration.

Item	Topographic Map	Nautical Chart
Coastline	- Mean Sea Level (MSL) which is determined by modelling the topography	- Local Astronomic Tide (LAT)
Horizontal Datum	- GDA94 - WGS84	- GDA94 - WGS84 - AGD66
Vertical Datum	- AHD (Australian Height Datum or Mean Sea Level) for land elevations. - no depth information	- Mean Sea Level (MSL) for land elevations - Chart Datum for depth - information: LAT, ISLW
Projection system	- Universal Transverse Mercator (UTM).	- Mercator
Digital Storage Format	- Various format (DWG, ARC, SHP, Hardcopy)	- Digital Nautical Charts: Raster(TIFF, ECW) - Electronic Navigation Chart: DIGITAL - S-57 Version 3.1 - Nautical Chart: Digital and Non digital - Raster HCRF V2 / GEOTIFF V1 (not to be used for navigation), Hardcopy Printed Charts - Bathymetric Map: Digital and Non digital-ASCII, Hardcopy - Printed maps
Scale	- Systematically (1 to 10K, 25K, 50K, 100K, 250K)	- Not Systematically (range from large scale to small scale)

5.2 Institutional Issues

There are several non-technical issues that should be overcome to develop a seamless SDI. The coastal zone is difficult to manage due a complex array of legislative and institutional arrangements varying from local to global levels. Furthermore, there is currently confusion about the management of the land-sea interface. This shows, for example, in Australia where local governments manage land to High Water Mark (HWM), and state governments manage the marine environment from the Low Water Mark (LWM). This means that there are no overlapping arrangements in place to enable efficient coastal zone management. There is also a strip of land between the two boundaries which is not within a management jurisdiction at all (Binns and Williamson, 2003).

Results from European Spatial Data Research (EuroSDR) questionnaire sent out to all European mapping agencies and hydrographic offices and geological organisations in late 2006 showed that only in a small number of cases the land and marine data is managed by a single organisation. In others collaboration across two or more organisations is required (typically national mapping agency, hydrographic office and sometimes the geological organisation) (Murray, 2007). Institutional integration increases the efficiencies and effectiveness of the management in any jurisdiction with land and marine environments. If national mapping and hydrographic charting agencies are separate, they need to work under the same banner and their policy should align with each other and the national policy to create a seamless SDI.

National mapping agencies and hydrographic offices use different coordinate systems, projections, horizontal and vertical datums and contents. Therefore users can not reference any object consistently across the coastal zone. A common framework will support interoperable coordinate systems and datums, interoperable objects along agreed boundary and interoperable feature catalogues. This agreed interoperable framework will contribute to the seamless SDI.

Immature institutional arrangements result in organisations working in the same jurisdiction or in the same discipline collecting similar data in different ways, engage in much duplication of effort, suffer from insufficient or inappropriate standards, are insufficiently aware of methods that should be used, or of the availability of existing data.

5.3 Policy Issues

The population and development pressures that coastal areas experience generate several critical problems and policy issues and raise serious and difficult challenges for coastal planners. A coastal state may be a party to many international conventions (i.e. RAMSAR, MARPOL, and London Convention) in addition to developing its own national, and even state or local regulations. Activities and resources are usually managed in a sectoral and ad-hoc approach with legislations or policies created when the need arises and specific to only one area of interest (Strain et al., 2006).

In many parts of the world, access to detailed information about the coast is considered a very sensitive issue, primarily due to concerns over national security. These restrictive policies lead to coastal data being withheld from stakeholders and the general public. Accordingly this complex, fragmented regulating framework for marine and coastal management causes the inability to adequately handle the pressure of different activities and stakeholders within the coastal zone.

The development of a framework such as a seamless SDI would aim to aid in facilitating decision making to respond to these technical, institutional and policy issues, to facilitate more effective management of the land sea interface.

6. CONCLUSION

In the terrestrial domain, the need to share and integrate spatial data for more efficient resource information management has been recognised for over a decade. There is now increasing recognition by the public at large of the need to support sustainable development of both the coastal and marine environments. The practical implementation of a marine SDI is mainly occurring separate from the terrestrial SDI, using the same components but adapting them to suit the different environment. However the multidisciplinary interactions in the land–sea interface require sophisticated information infra-

structures that not only do not yet exist, but which will not appear if disciplines continue to develop their SDIs in isolation from one another. Research now needs to focus on combining these initiatives and developing a seamless SDI as one abstract class SDI at the higher level. The development of a seamless SDI will ensure this data is interoperable and thus improve decision-making and administration in the coastal and marine environments. However, the differences in the marine and terrestrial environments in fundamental datasets, data collection and technology used in these environments will make interoperability and integratability between marine and terrestrial spatial data a big challenge.

ACKNOWLEDGEMENTS

REFERENCES

Bartlett, D., Longhorn R. and Garriga, M. (2004). "Marine and Coastal Data Infrastructures: a missing piece in the SDI puzzle", *7th Global Spatial Data Infrastructure conference*, *Bangalore, India.*

Binns A. and Williamson I. (2003). "Building a national marine initiative through the development of a marine cadastre for Australia", *International Conference on the Sustainable Development of the Seas of East Asia, December 8-12 2003, Putrajaya, Malaysia.*

Brower, D.J., Schwab, A.K. and Beatley, T. (2002). *An Introduction to Coastal Zone Management*, 2nd edn, Washington D.C.: Island Press.

BRS (Bureau of Rural Sciences) (2002). *Fisheries Status Report,* Commonwealth of Australia, Canberra.

Finney, K. (2007). A Bottom Up Governance Framework for Developing Australia's Marine Spatial Data Infrastructure (SDI), *Data Science Journal,* 6(July): 64-90.

Gillespie, R., Butler, M., Anderson, A., Kucera, H. and LeBlanc, C. (2000). MGDI: An Information Infrastructure to Support Integrated Coastal Zone Management in Canada, *GeoCoast,* 1(1): 15-24.

IHO (2005). The Role of Hydrographic Services with regard to Geospatial Data and Planning Infrastructure, at: http://www.iho.int/COMMITTEES/CHRIS/CHRIS/ IHO _SDI_Seminar/SDI_Seminar.htm.

Maratos, A. (2007). "The IHO Responding to MSDI Requirements", *Proceedings of 5th International Congress Geomatica 2007, Workshop on Marine/Hydrographic Spatial Data Infrastructure, February 12 2007, Havana, Cuba.*

Murray, K. (2007). "Land-Sea Information Base-Information Interoperability: Results of EuroSDR Questionnaire", *Proceedings Joint EuroSDR and IHO Land and Marine Information Integration Workshop, March 21-23 2007, Dublin. Ireland.*

Nebert , D.D. (ed.) (2004). Developing Spatial Data Infrastructures: The SDI Cookbook, GSDI Association, at: http://www.gsdi.org/docs2004/Cookbook/cook-bookV2.0. pdf.

Neely, R.M., Treml, E., LaVoi, T. and Fowler, C. (1998). Facilitating Integrated Regional Ocean Management Using a Web-based Geographic Information System, Coastal Services Centre, National Oceans and Atmospheric Administration, at: http://www.csc.noaa.gov/opis/html/occ_98.htm.

NOO (National Oceans Office) (2002). *Resources-using the ocean, The south-east Regional Marine Plan,* National oceans office, Hobart, Tasmania.

Plunkett, G. (2001). A History of Sea Dumping off Australia and its Territories, at: http://www.ea.gov.au/coasts/pollution/dumping/history.html.

Rajabifard, A., Chan, T.O., and Williamson, I.P. (1999). "The Nature of Regional Spatial Data Infrastructures", *Proceedings AURISA 99 - The 27th Annual Conference of AURISA, Fairmont Resort, Blue Mts, Australia.*

Rajabifard, A. and Williamson I.P. (2001). "Spatial Data Infrastructures: Concept SDI Heirarchy and Future Directions", *GEOMATICS80, Tehran, Iran.*

Rajabifard, A. (2002). Diffusion of Regional Spatial Data Infrastructures: with particular reference to Asia and the Pacific, PhD Thesis, University of Melbourne, Australia.

Rajabifard A., Binns A. and Williamson I. (2005). Administering the Marine Environment – The Spatial Dimension, *Journal of Spatial Science*, 50(2): 69-78.

Rhind, D. (2001). Lessons learned from local, National and Global Spatial Data Infrastructures, *GISdevelopment.net*, at: http://www.gisdevelopment.net/policy/gii/gii 0006.htm.

SOEAC (State of the Environment Advisory Council) (1996). *State of the Environment Australia,* CoA, Canberra.

Strain, L., Rajaibard, A. and Williamson, I.P. (2006). Spatial Data Infrastructure and Marine Administration, *Marine Policy*, 30(4): 431-441.

Syafi'I, A. (2006). "The Integration of Land and Marine Spatial Data Set, As Part of Indonesian Spatial Data Infrastructure Development", *17th UNRCC-AP*, *September 18-22 2006, Bangkok.*

Thia-Eng, C., Bernad, S. and San, M. (2003). Coastal and Ocean Governance of the Seas of East Asia: Towards an Era of New Regional Cooperation and Partnerships, *Tropical Coasts*, 10(1): 46-55.

Tóth, K. (2007). "Role of Integrated Land and Marine Information in INSPIRE", *Proceedings Joint EuroSDR and IHO Land and Marine Information Integration Workshop, March 21-23 2007, Dublin. Ireland.*

UN-PCGIAP (2004). Working Group 3. Report on the international workshop on administrating the marine environment - the spatial dimension, May 4-7 2004, Kuala Lumpur, Malaysia.

UNRCC-AP (2006). Resolution 3: Marine Administration-the spatial dimension, *17th UNRCC-AP*, *September 18-22 2006, Bangkok,* at http://www.gsi.go.jp/PCGIAP.

Williamson I.P., Rajabifard A. and Binns A. (2004). "Issues in Developing Marine SDI", *International Workshop on Administering the Marine Environment – The Spatial Dimensions*, *May 4-7 2004, Kuala Lumpur, Malaysia*, 11p.

The RRR Toolbox: a Conceptual Model for Improving Spatial Data Management in SDIs

Rohan Bennett and Abbas Rajabifard

Department of Geomatics, The University of Melbourne, Australia
{rohanb, abbas.r}@unimelb.edu.au

Abstract

Spatial Data Infrastructures (SDI) aim to link people with spatial services and data. Increasingly, SDI initiatives are focusing on a particular type of data: large-scale people relevant data. Examples include the ownership parcel layer and built environment information. To improve the management of these essential SDI layers, consideration of land parcels and their administration is needed. In particular, the complex array of rights, restrictions and responsibilities (RRRs) applying to land needs to be understood. Moreover, the contemporary models of RRRs management must also be understood. To this end, this article introduces the RRR Toolbox, a holistic framework for understanding, creating and managing land interests. The nature and design of RRRs are discussed along with their problematic management. This leads to a description of the development and components of the RRR Toolbox. The dualism of RRRs, being both spatial and land based, makes the RRR Toolbox highly applicable to SDI. Indeed, seven of the model's eight components are found to be relevant in the SDI context.

Keywords: RRRs (Property rights, restrictions and responsibilities), SDI (Spatial Data Infrastructure), Land Administration.

1. INTRODUCTION

Early SDI activities and research often suffered from a 'means' over 'ends' mentality: too much focus was placed on technologies, not enough was given to the activities and datasets they underpinned. As the technological problems subsided, SDI science broadened into a multi-disciplinary field: theories and concepts from numbers of areas came together to enable the design, construction and management SDIs. One area offering useful insights was Land Administration, the discipline of cadastral management and property rights organisation. Land Administration was essentially a subset of the larger SDI problem; its long and well-documented history provided many lessons for SDI practitioners. Conversely, SDI has also fed back into the discipline of Land Administration through its call for data sharing and institutional collaboration.

This article intends to continue the cross-pollination between the disciplines. It introduces the RRR Toolbox and aims to demonstrate it utility to other data sets within SDIs. The model focuses on reorganising the management of the hundreds of new RRRs placed on land by governments. RRRs are created for a myriad of social, economic and environmental reasons and impact greatly on the tenure, value and potential use of land. Consequently, information about RRRs is keenly sort after by citizens and government. Unfortunately, RRR creation, management and access mechanisms are usually poorly implemented. The RRR Toolbox, the focus of this article, emerged from applied scientific research and consists of eight principles for improving RRR management. Subsequently, the tools are also useful to the realm of SDI and it's broader aim of spatial data integration.

The article begins by outlining the close relationship between SDI and Land Administration: this serves to justify the article's central argument. An overview of RRRs is then provided: their nature, design and poor management is discussed. Responses to the problem are reviewed leading to the introduction of the RRR toolbox. The applicability of the toolbox to SDI is then discussed by considering the RRR Toolbox components and their relevance to SDI science. The usefulness of the toolbox is left open for debate, however, the article concludes by calling for more testing of the Toolbox through its application in the realm of SDI.

2. THE RELATIONSHIP BETWEEN SDIs, LAND ADMINISTRATION AND RRRs

SDI and Land Administration have shared backgrounds. In many countries SDI initiatives initially emerged out of Land Administration operations. Additionally, a large component of Land Administration research concentrates on organising and sharing land related spatial data: a specific case of SDI's larger brief. This suggests there is much opportunity for further cross-pollination between SDIs and RRR management.

Land Administrators have long acknowledged that SDIs will be crucial to achieving their long-term objectives, particularly, the achievement of sustainable development. Williamson (2001) included SDIs in his Land Administration toolbox; SDIs were seen as an important component in the integrated management of land:

"Spatial data infrastructures (SDI) are a key component of any land administration infrastructure (Mooney and Grant, 1997; Groot and McLaughlin, 2000). An understanding of the role and potential of SDIs in supporting land administration systems greatly assists any land administration reform process. In particular the generic principles concerned with the development of an "infrastructure", as distinct from "business systems" which rely on the infrastructure, are very useful (Chan and Williamson, 1999). Also an understanding of the role and maintenance of the cadastral or land parcel layer in an SDI is important (Williamson et al., 1998). At the same time an understanding of key SDI principles, such as the hierarchy of SDIs in a jurisdiction and the dynamic nature of SDIs, are useful (Rajabifard et al., 2000)."

In more recent times, Bennett (2007) and Ting (2002) point out that the basic components of an SDI as defined by Rajabifard et al. (2002a) will be an important component of any framework aimed at improving the management of RRRs (Bennett, 2007). A physical access network, overarching policy statement mandated by an empowered leadership agency and operational standards will be essential elements. Already a number of jurisdictions are using the SDIs initiated in the early 2000s to assist in the management of land information. Western Australia's SLIP model and accompanying Register of Interests provides a good example (Searle and Britton, 2005).

Therefore, the relationship between SDI and Land Administration is well established: there has been an ongoing role for SDI in the management of land and land interests. However, the statement should also be considered in reverse: Can the new lessons from Land Administration, particularly the management of RRRs, be applied to SDI initiatives? Before exploring potential this relationship further the nature, design and management of RRRs needs to be explained.

3. THE NATURE, DESIGN AND MANAGEMENT OF RRRs

In the latter stages of the 20th century, sustainable development principles began guiding government decision-making processes. The principles demanded that growth oc-

curring in the present must not compromise the ability of future generations to meet their own needs (UN, 1987). To achieve sustainable development, governments increasingly turned to legislating new RRRs over land. These laws gave powers to governments, individuals and other mandated bodies, and were aimed at controlling the community's behaviour in relation to land.

The new RRRs increased in number and complexity for several decades to a point where most land related activities were subjected to some form of legislative control. Examples included the alienation of land for use as national parks, the creation of water and timber rights on private land, and the reallocation of land rights to indigenous peoples. In addition to these highly visible interests many other lower profile interests were created, including entry powers for agents of the state and the allocation of private parking spaces.

The volume of legislation involved was enormous: a 2002 study found that, in the Australian state of Queensland, almost two hundred individual statutes created some type of control over land (Lyons et al., 2002), and the number of interests was continuing to increase. An investigation into Australia's regulatory environment by the federal government's Regulation Taskforce found that the Australian Parliament had passed more legislation since 1990 than in the previous ninety years of federation (Regulation Taskforce, 2006). Similar statistics could be found in other jurisdictions and countries. Additionally, the RRRs were often created in isolation to one another and administered using a complex range of government bodies and information systems.

The field most likely to bring together the disparate array of legislation and accompanying information was Land Administration. Traditionally, Land Administration had focused on managing one type of RRR: privately held ownership rights. These interests are central to modern economies: they are responsible for generating much of the wealth in developed countries (De Soto, 2000). Unlike many of the newer RRRs, these traditional interests were well understood and respected by citizens. They were backed by theoretical, legislative and institutional frameworks that evolved over hundreds of years. However, the majority of the new RRRs were not created within these traditional Land Administration frameworks.

During the 1990s Land Administrators broadened their traditional focus: rather than dealing solely with ownership rights, there was more attention given to understanding and managing emerging RRRs on land (Ting, 2002). The early literature (Williamson et al., 2005; Enemark and Williamson, 2004; Van der Molen, 2003; Lyons et al., 2002 and 2004; Ting, 2002; Ting et al., 1999; Ting and Williamson, 1998; FIG, 1998) revealed a number of problems. Many were poorly designed, many were poorly administered and some interests did not exist where they ought to. For example, in the Australian state of Victoria, there were minimal controls preventing people from building on contaminated land.

The problems with RRRs impeded the achievement of Land Administration's greater goal: sustainable development. Indeed, they made it virtually impossible: sustainability could not be achieved without integrated management of RRRs (UN-FIG, 1999). The traditional Land Administration systems needed reform and integration using a holistic design framework, one that encompassed most RRRs from outright ownership down to simple access rights. While components of this framework already existed, there was not yet a complete coherent understanding. Contributions to knowledge in the realm were urgently required.

4. INITIAL RESPONSES TO THE RRRs CHALLENGE

Responses to the RRRs problem could be found as far back as the 1970s when researchers began to recognise the potential of Land Administration systems to assist in the management of new types of laws and information. Authors such as Peter Dale, John Mclaughlin and Ian Williamson were the first to recognise the potential (McLaughlin, 1975; Dale and McLaughlin, 1988; Williamson, 1985 and 1993). They saw that cadastres could be used for more than just fiscal and juridical management: they could assist in the management of natural resource information.

During the mid 1990s the discipline of 'Land Administration' formally emerged through a series of statements: The FIG statement on the cadastre (FIG, 1995), The Bogor declaration (FIG, 1996), Cadastre 2014 (Kaufmann and Steudler, 1998) and The Bathurst Declaration (UN-FIG, 1999). These documents suggested that the role of the cadastral system was to disclose the complete legal situation of land, including all "public rights and restrictions" and introduced the concept of the legal land object. These were significant statements and visions. They, along with numerous government initiatives, highlighted the RRRs problem, gave it international prominence and provided a high level vision.

Another emerging body of literature dealt with more practical implementations. The literature could be divided into three categories: government restructures, technological solutions and legislative strategies. In relation to government restructures, the most organised recommendations came from the work of land administrators Dale and Baldwin (1999) and Dale (2000) who suggested the use of markets to improve the management of land interests. Lyons et al. (2002; 2004), built upon this work and identified that a land market can be unbundled into separate resource sub-markets. To build these markets a complete overhaul of all existing land administration functions would be required. The costing of the drastic changes was never undertaken. Other commentators in Australia suggested scaling up the pre-existing Torrens systems for the management of new RRRs (Young and McColl, 2002; ACIL Tasman, 2004). The Torrens system proposal would enable holistic management; however, it risked cluttering up the land ownership management system. Discussions about the merits of government reorganisation continue, however, by 2008 none of the top-down solutions being offered were financially sound enough to be implemented.

Meanwhile, smaller scale technology driven approaches took precedence in most countries. Technology removed the need to restructure government by enabling the creation of virtual links between departments and their information. Cadastre 2014 promoted the use of technology in its vision for future cadastres. In 2006 members of Commission 7 of the FIG completed a core land administration domain model that could be used to manage RRRs collectively (Van Oosterom et al., 2006). The model provides a standard for all agencies managing and storing information about land interests. The model challenges existing cadastres, which are based around the ownership parcel. It was released in 2006 but has not yet been implemented in any jurisdiction.

Australian governments were also utilising technological options, with the majority of work being undertaken at a state level. At the Expert Group Meeting on Sustainability and Land Administration held at the University of Melbourne in 2005, all the States represented were undertaking projects to improve land information management (Williamson et al., 2005). There was a particular focus upon the utilisation of newly available spatial technologies and concepts, including spatial data infrastructures (SDI), spatial

databases and web mapping services. These tools allow for complex legislative and administrative systems to be integrated without reorganising government institutions. They also assist in the distribution of land information to citizens. Western Australia's Shared Land Information Platform (SLIP) and accompanying Register of Interests (ROI), a whole-of-government web mapping service infrastructure, provide very good examples of the tools in action (Searle and Britton, 2005).

With regard to legislative strategies, Western European countries tended to lead. They were introducing new laws and codes to improve information management. The Netherlands passed a law on the Registration of Public Encumbrances 2005 that obliged all municipalities to establish and maintain a publicly available register of the land interests that they imposed upon real estate (Van der Molen, 2005; Zevenbergen and De Jong, 2002). Additionally, the European Union introduced requirements for the publication of land information documents like - 'EuroStat, 2000 Statistical Requirements Compendium'- a 10-year agricultural survey (Statistical Office of the European Communities, 2007). The administrative practicalities of these new laws are still being resolved; however, legislative burdens on government appeared to be an important component of any solution.

The above approaches illustrate the different tools that could be applied to address the RRRs problem, however, focusing on technology, legislation or institutions alone would only result only in short term success: no single tool had provided a sustained long-term solution. A more holistic approach was required.

5. DEVELOPING THE RRR TOOLBOX

In response to the need for more holistic approaches to manage RRRs, Bennett (2007) proposed the RRR Toolbox. The toolbox is based upon previous Land Administration research that promotes holistic approaches to managing land (MacLauchlan and McLaughlin, 1998, Williamson, 2001; Williamson, 2004). Williamson's models suggested all Land Administration systems must incorporate eight broad principles to be sustained. However, to achieve each principle a range of tools are available, the selection of which is based on a country's circumstances. Bennett (2007) suggested that Williamson's toolbox needed to be further extended to include appropriate tools for the creation and administration of 'all' RRRs.

Research was undertaken to determine what new principles and tools Williamson's model required. Particular focus was given to the need to meet sustainable development objectives. The framework would focus on a range of technical and non-technical aspects including: policy, legal, tenure, institutional, cadastral, registration, technical and human resource aspects. Each broad element was seen as essential to every Land Administration system. The framework would primarily be designed for use in developed countries; however, parts would be applicable to all country typologies.

The research was conducted as follows. A number of specific questions were generated in order to discover how Williamson's (2001) toolbox might be extended. These questions guided the research activities. The activities were built upon around a mixed methodology framework (Frechtling and Westat, 1997; Tashakkori and Teddlie, 1998 and 2003; Creswell, 2003; Johnson and Onwuegbuzie, 2003) involving both qualitative and quantitative case studies. Additionally, because the research problem focused on the requirements of government and citizens two perspectives were required: top-down (or government – Australia, State of Victoria, Moreland City Council) and bottom-up (or parcel level – 4 properties: urban, rural, agricultural, suburban). Each perspective in-

cluded a qualitative and quantitative study. Quantitative studies were used to answer the 'how many' research questions and the qualitative studies were used to answer the 'how should' research questions. Thus, the research design could be considered a two-by-two matrix; incorporating case studies from government and citizens' perspectives, each with a quantitative and qualitative component. See Bennett (2007) and Bennett et al. (2008) for further details.

All four studies undertaken were considered equal in weight. Additionally, the two different perspectives acted as a check (or test) of the results obtained and hence the final framework was considered robust and justified. Together, the results from these equally weighted case studies were used to generate components of the updated Land Administration toolbox, or as it is referred to hence forth, The RRR Toolbox.

6. AN INTRODUCTION TO THE RRR TOOLBOX

The RRR Toolbox is a framework for managing RRRs that is understandable and applicable to individuals, institutions and the wider society (Figure 1). If a jurisdiction wishes to manage coherently all it's RRRs, then each of the eight components needs to be addressed and acted upon. A major strength is its cross-disciplinary nature: rather than dividing RRR management into disparate components, the toolbox provides a simplistic overarching framework that encourages practitioners from a range of disciplines to understand their role in the larger administrative process. This holistic view has been lacking in previous models.

From an overarching perspective, the toolbox is organised into the categories similar to those in Williamson's (2001) Land Administration toolbox (Figure 1). However, a number of alterations are made: Legal principles and HR (Human Resource) capacity building principles are included. These tools did not appear in the original toolbox, however, they were added in later versions and therefore appear here also. Cadastral principles now include registration principles. Numerous registration options are available for dealing with land interests that do not equate to full ownership; these are included in this component. SDI and technology principles are merged, reflecting the convergence of spatial technologies and ICT that occurred after the original toolbox was developed. A new component, emerging principles, is also included. This component groups the emerging concepts and theories discovered throughout the research and that are highly applicable to land interest management.

While the 8 principles are included here, the individual tools enabling achievement of each of the principles are not. These are examined in depth in Bennett (2007) and Bennett et al. (2008). What is of interest is how the broad RRR Toolbox principles might assist the implementation and management of SDI initiatives.

7. APPLYING THE RRR TOOLBOX TO SDI

As discussed earlier, RRRs are a subset of spatial information layers that ultimately need to be integrated into SDIs. It therefore follows that most of the principles that guide RRRs management could also be applicable to SDIs. Preliminary analysis, based on the five core components of SDIs (Rajabifard et al., 2002b) and the future challenges facing SDIs (Williamson et al., 2006), suggests that seven of the eight RRR Toolbox principles hold relevance to SDI (Figure 2). Moreover, more recent SDI literature (Masser et al., 2007; Craglia et al., 2008) points to human resource/capacity building, technological issues and governance/institutional issues being the future challenges facing SDIs. Based on these examinations, three of the eight principles: Legal

Principles, Institutional Principles and Spatial and Technology Principles; are considered to be very relevant to SDI. The model appears to have utility in the field of SDI, however, further analysis is required to test this hypothesis.

Figure 1: The RRR Toolbox – a framework for holistically managing the majority of land interests (Bennett, 2007).
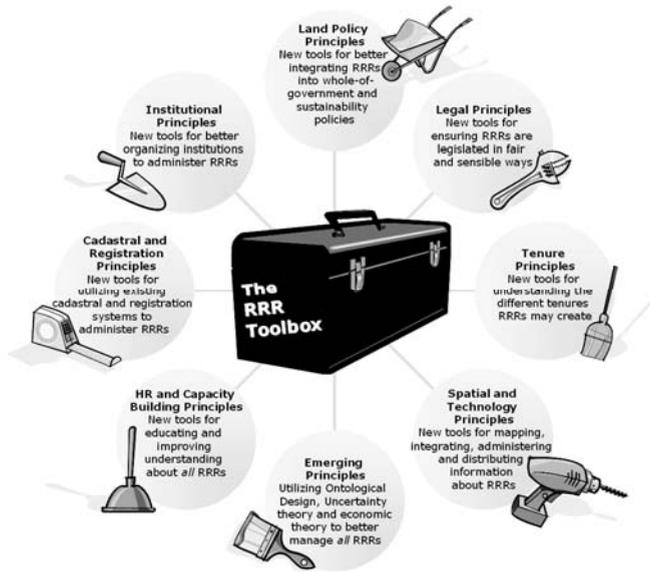
Figure 2: Predicted relevance of RRR Toolbox to SDI.
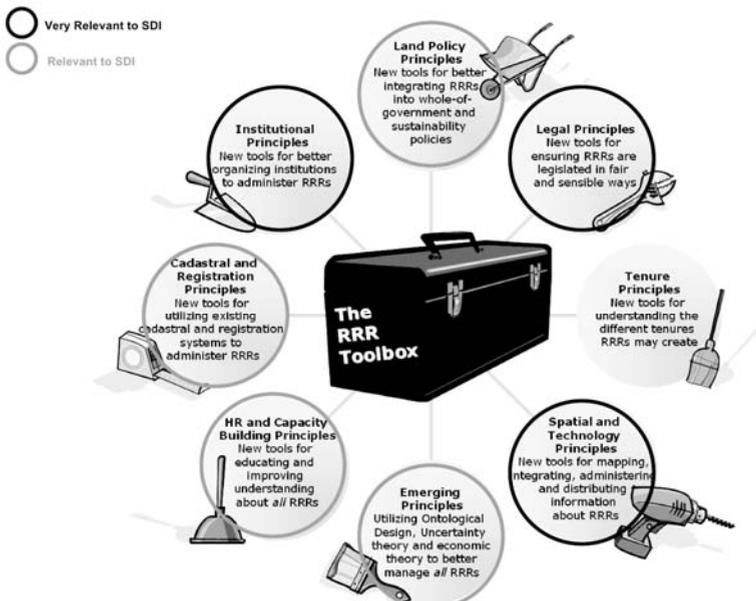
As an example the 'Spatial and Technology Principles' (Figure 3) is now considered. Each of the seven concepts listed appears to have much utility in the realm of SDI: data acquisition, information attributes, data sourcing, information access, infrastructure, interfaces and standards are all integral elements of SDI. SDI researchers and practitioners should examine the requirements of Land Administration in terms of these principles, if only to improve their knowledge and understanding of the RRR based datasets within SDI. Moreover, the principles revealed in The RRR Toolbox appear to be generic enough to apply to many SDI initiatives.

Figure 3: Spatial and ICT principles from the RRRs Toolbox.

Principles, tools and explanations
Acquisition: **lack of datasets and lack of integration** All land interests have a spatial extent. The studies showed that a large majority of interests had no formal mapping or spatial identification. Those that had been mapped were often not integrated with other key datasets e.g. cadastre, roads. Those interests that require spatial enablement need identification. Better programs for integrating spatial datasets using SDI concepts are required.
Information: **spatial extents, duration and people impacted must be recorded** Location, time and place attributes should be defined and recorded in uniform fashion by government agencies. This will enable better ordering, integration and searching of core land interest information. In Australia's case this would be through PSMA.
Source: **identify best available information** In some cases multiple agencies and organizations hold information relating to an interest. The most authoritative source of information relating to a source needs to be identified and indicated in some way.
Access: **land interest information and transactions should be online and affordable** Limited land transactions are available to citizens online. Many transactions are still paper based: only printable forms are provided online. Governments should strive to include the processes of creation, alteration and removal online. Generic standards should apply to information types. The most important information should be provided for cost of provision using web services and guaranteed. Any damage suffered because of incorrect information should be recoverable. Less generically important information should be provided for cost of collection and provision.
Infrastructure: **SDI overcomes the need to reorganize government** SDI removes the need to reengineer governments. Standard infrastructure platforms enable the integration of government information.
Interface: **web services need to be designed around land activities not datasets** Existing government web sites tend to allow citizens to view different land datasets, however, sites should be designed around core activities and transactions.
Standards: **uniform spatial identifiers, units and access need to be developed** In the past different agencies used different spatial identifiers. For example, addressing is still unreasonably complex in Victoria. In urban areas when numbers increase, odd numbers are on your left and even numbers are on your right. When even numbers are on your left, numbers are decreasing. In case of rural addressing, the number multiplied by 10 indicates the distance in metres from the start of the road. Integration and efficiency demands that uniform units and identifiers be adopted.

8. CONCLUSION

SDI is fundamentally a concept about coordinating the sharing of spatial data, services and other resources between stakeholders from different political/administrative levels.

260

The commonalities between SDIs and the objectives Land Administration systems provide strong grounds for the derivation of shared evaluation and performance indicators. This article has sought to show how one such model, The RRR Toolbox, could be applied in the SDI context.

In essence, RRRs are a challenge for both SDIs and Land Administration: they exhibit a form of dualism being both spatial and land related entities. The lessons learnt from research into RRRs offer utility to SDI practitioners. It can improve the management of both RRRs and SDI through the use of the broad range of principles and tools. Preliminary studies show that these principles are highly relevant to SDI initiatives; however, further investigation is needed to test the Toolbox's full utility within the SDI discipline.

ACKNOWLEDGEMENTS

REFERENCES

ACIL Tasman in association with Freehills (2004), *An Effective System of Defining Water Property Titles,* Research Report, Land and Water Canberra, Australia.

Bennett, R. (2007). Property rights, restrictions and responsibilities: their nature, design and management*, PhD Thesis, Department of Geomatics, The Univeristy of Melbourne, Austraila.

Bennett, R., Wallace, J., Williamson, I.P. (2008). A framework for mapping and managing land interests, *Survey Review*, 40(307): 43-53.

Chan, T. O. and Williamson, I. P. (1999). The different identities of GIS and GIS diffusion. *International Journal of Geographical Information Science*, 13(3): 267-281.

Craglia, M., Goodchild, M.F., Annoni, A., Camara, G., Gould, M., Kuhn, W., Mark, D., Masser, D., Maguire, D., Liang, S., and Parsons, E., (2008), Next Generation Digital Earth – A Position Paper from the Vespucci Initiative for the Advancement of Geographic Information Science, *International Journal of Spatial Data Infrastructures Research*, 3: 146-167.

Creswell, J.W. (2003). *Research design: qualitative, quantitative, and mixed method approaches*, 2nd edition, Sage Publications, Thousand Oaks, CA, United States.

Dale, P. F. (2000). The Importance of Land Administration in the Development of Land Markets: A Global Perspective, *Proceedings of UDMS 2000,* Delft University of Technology, the Netherlands, 11-15 September 2000, pp. 31-41.

Dale, P.F. and Baldwin, R. (1999). Emerging land markets in Central and Eastern Europe, *Proc of the Second World Bank/FAO Workshop on Lessons for EU Accession*, Warsaw, Poland, June 27-29, (ed) C.Csaki and Z.Lerman, World Bank Technical Paper No. 465, pp. 81-109.

Dale, P., and McLaughlin, J.D. (1988). *Land Information Management*, Oxford, United Kingdom: Clarendon Press.

De Soto, H. (2000). *The Mystery of Capital*, Basic Books, 1st edition.

Enemark, S., and Williamson, I.P. (2004). Capacity Building in Land administration - A Conceptual Approach, *Survey Review*, 37(294): 639-650.

European Communities (Statistical Office of). (2006). *Eurostat Homepage*, at: http:// epp.eurstat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33076576 &_dad=portal&_schema=PORTAL.

FIG (1995). *Statement on the Cadastre,* Report prepared for the International Federation of Surveyors by Commission 7 (Cadastre and Land Management), FIG Publication No. 11, at: http://www.sli.unimelb.edu.au/fig7/cadastre/statement_on_ cadastre.html.

FIG (1996). *The Bogar Declaration*, United National Interregional Meeting of Experts on the Cadastre, Bogor, Indonesia, March 18-22, at:http://www.sli.unimelb.edu. au/fig7/Bogor/BogorDeclaration.html.

FIG (1998). Kaufmann, J. and Steudler, D. (ed.), *Cadastre 2014: A Vision for a Future Cadastral System*, International Federation of Surveyors, Switzerland.

Frechtling, J., and Westat, L.S. (eds.) (1997). *User-Friendly Handbook for Mixed Method Evaluations*, Division of Research, Evaluation and Communication, National Science Foundation, at: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/ START.HTM.

Groot, R. and McLaughlin, J.D. (2000). *Geospatial Data Infrastructure: concepts, cases and good practice*, New York, USA: Oxford University Press.

Johnson, R.B., and Onwuegbuzie, A.J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come, *Educational Researcher*, 33(7): 14-26.

Kaufmann, J. and Steudler, D. (1998). *Cadastre 2014: A Vision for a Future Cadastral System*, International Federation of Surveyors, Switzerland.

Lyons, D.,K., Cottrell, E.C., and Davies, K.J. (2002). *On the Efficiency of Property Rights Administration in Queensland*, Research Report, Queensland Government, Department of Natural Resources and Mines, Brisbane. Australia.

Lyons, D., K., Cottrell, E.C., Davies, K.J. (2004). A Summary of the Paper "On the efficiency of property rights administration", Research Report, Queensland Government, Department of Natural Resources and Mines, Brisbane, at: www.anzlic. org.au/pubinfo/2393092707.

MacLauchlan, W., and McLaughlin**,** J.D. (1998). Environmental democracy and stewardship through land researouce and information: The case of the Land Gazette, *Proc of International Conference on Land Tenure in the Developing World*, Ed. M. Barry , Cape Town, South Africa, November 1997.

Masser, I., Rajabifard, A., and Williamson, I.P. (2007). Spatially Enabling Governments through SDI implementation*, International Journal of Geographical Information Science*, 22(1): 5-20.

McLaughlin, J.D. (1975). *The Nature, Design and Development of Multi-Purpose Cadastres*, PhD Thesis, University of Wisconsin-Madison, Wisconsin, United States.

Mooney, J. D. and Grant, D. M. (1997). "The Australian Spatial Data Infrastructure", in Rhind D. (ed.). *Framework of the World*, Cambridge: GeoInformation International, pp. 187-201.

Rajabifard, A, Williamson, I P, Holland, P, and Johnstone, G. (2000). "From Local to Global SDI initiatives: a pyramid building blocks", *Proceedings of the 4th GSDI Conference*, Cape Town, South Africa.

Rajabifard, A., Feeney, M.A., and Williamson, I.P.W. (2002a). "Spatial Data Infrastructures: Concept, Nature and SDI Hierachy", in I.P. Williamson, A. Rajabifard, M.-E.F. Feeney (eds.). *Developing SDIs: from Concept to Reality*, New York, United States: Taylor and Francis.

Rajabifard, A., Feeney, M.A., and Williamson I.P. (2002b). Future Directions for the Development of Spatial Data Infrastructure*, International Journal of Applied Earth Observation and Geoinformation*, 4(1): 11-22.

Regulation Taskforce. (2006). Rethinking Regulation: Report of the Taskforce on Reducing Regulatory Burdens on Business, Report to the Prime Minister and the treasurer, Canberra, Australia, January, at: http://www.regulationtaskforce.gov.au.

Searle, G., and Britton, D. (2005). The Western Australian Shared Land Information Platform and Modern Land Information Systems, *Proceedings of the. Expert Group Meeting on Sustainability and Land Administration*, Melbourne, Australia, November 9-11, pp. 117-132.

Tashakkori, A. and Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches*, Vol. 46, Applied social research methods series, Thousand Oaks, United States: Sage Publications.

Tashakkori, A., and Teddlie, C. (eds.) (2003). *Handbook of mixed methods in social and behavioral research*, Thousand Oaks, United States: Sage Publications.

Ting, L. (2002). Principles for an Integrated Land Administration System to Support Sustainable Development, PhD Thesis, Department of Geomatics, The University of Melbourne, Australia.

Ting, L., and Williamson, I.P. (1998). Cadastral Trends: A Synthesis. *Australian Surveyor*, 4(1): 46-54.

Ting, L., Williamson, I.P., Parker, I., and Grant, D. (1999). Understanding the Evolution of Western Societies' Land Administration Systems: A basis for cadastral reform, *Survey Review*, 35(272): 83-102.

UN. (1987). *Our Common Future: From One Earth to One World*, (World Commission on Environment and Development / Brundtland), Oxford: Oxford University Press.

UN-FIG (1999). *The Bathurst Declaration*, UN-FIG International Workshop on Land Tenure and Cadastral Infrastructures in Support of Sustainable Development.

Van der Molen, P. (2003). PS1.3 The Future Cadastres – Cadastres after 2014, *FIG Working Week 2003, PS1 Cadastre*, Paris, France, April 13-17, at: http://www.eurocadastre.org/pdf/vandermolen2.pdf.

Van der Molen, P. (2005). Incorporating Sustainable Development Objectives into ICT enabled Land Administration Systems in the Netherlands. *Proceedings of the. Expert Group Meeting on Sustainability and Land Administration*, Melbourne, Australia, November 9-11, pp. 83-96.

Van Oosterom, P.J.M., Lemmen, C.H.J., Ingvarsson, T., van der Molen, P., Ploeger, H., Quak, W., Stoter, J., and Zevenbergen, J. (2006). The core cadastral domain model. *Computers, Environment and Urban Systems*, 30(5): 627-660.

Williamson, I.P. (1985). Cadastres and Land Information Systems in Common Law Jurisdictions, *Survey Review*, 28(217): 114-119.

Williamson, I.P., Chan, T.O., and Effenberg, W. W. (1998). Development of spatial data infrastructures - lessons learned from the Australian digital cadastral databases. *Geomatica*, 52(2): 177-187.

Williamson, I.P. (2001). Land Administration "Best Practice" Providing the Infrastructure for Land Policy Implementation, *Journal of Land Use Policy*, 18(4): 297-307.

Williamson, I.P. (2004). *Land Administration 'best practice' and the 'toolbox' concept*, Lecture 11 - 451-418/607 Land Administration, 1 April, at: http://www.geom.unimelb.edu.au/subjects/451/418/418_2004/table.htm.

Williamson, I.P. (1993). The Role of Land And Geographic Information Systems in Economic Development and Environmental Management - A Developing Country Perspective, *International Conference on Land Information and Land Management with Modern Cadastre as the Centerpiece,* Olsztyn, Poland, at: http://www.geom.unimelb.edu.au/research/publications/IPW/CadastrePoland.htm.

Williamson, I.P., Enemark, S., and Wallace, J. (eds.) (2005). *Proceedings of the Expert Group Meeting on Incorporating Sustainable Development Objectives into ICT Enabled Land Administration Systems*, 9-11 November 2005, Centre for Spatial Data Infrastructures and Land Administration, The University of Melbourne, Australia.

Williamson, I.P., Rajabifard, A., and Binns, A. (2006). Challenges and issues for SDI Development, *International Journal of Spatial Data Infrastructures Research*, 1: 124-35.

Young, M.D., and McColl, J.C. (2002). *Robust Separation: A search for a generic framework to simplify registration and trading of interests in natural resources*, CSIRO Land and Water, Policy and Economic Research Unit, Folio No: S/02/1578, September.

Zevenbergen, J. and De Jong, J. (2002). Public Law Information Regarding Land: Dutch proposal for registration, *FIG XXII International Congress*, Washington, D.C., USA, April 19-26.

Building SDI Bridges for Catchment Management

Dev Raj Paudyal, Kevin McDougall and Armando Apan

Faculty of Engineering and Surveying, Australian Centre for Sustainable Catchments
University of Southern Queensland, Australia
{paudyal, mcdougak, apana}@usq.edu.au

Abstract

This research paper discusses the importance of spatial data and Spatial Data Infrastructure (SDI) for catchment management. It reviews four SDI theories including hierarchical spatial theory, diffusion theory, evolution theory and principal-agent (P-A) theory and discusses their characteristics and potential utilisation for catchment management. As catchment management issues are characterised by multi-level stakeholder participation in SDI implementation, the theory of hierarchy and the P-A theory may assist in exploring in greater depth the context of building SDI at the catchment level. Based upon existing SDI theory, it explores a conceptual framework and its implications for more effective development of catchment-based SDI. The framework which is based upon hierarchical theory, investigates the community-government interaction between various catchment and administrative/political levels for developing SDI. Such a framework is complex and potentially has many levels. Additionally, the cross-jurisdictional linkages required to implement this framework within the existing administrative/political SDI framework also need to be carefully examined. The framework is explored through a case study of the Murray-Darling Basin in Australia, one of the world's largest catchments. The challenges for developing an SDI which effectively supports the decision making within and across this catchment will be discussed and the potential strengths and weakness of the proposed framework identified in the context of this case study.

Keywords: spatial data infrastructure (SDI), catchment management, spatial data, natural resource management, SDI Theory.

1. INTRODUCTION

Spatial data plays an important role in many social, environmental, economic and political decisions (McDougall and Rajabifard, 2007) and is increasingly acknowledged as a national resource essential for sustainable development (Warnest, 2005). Accurate, up-to-date, relevant and accessible spatial data is essential in addressing various global issues such as climate change, urban change, land use change, poverty reduction, environmental protection and sustainable development. One of the potential areas where spatial data can make a positive impact is for improved decision making to support catchment management. Reliable spatial data infrastructure (SDI) is needed to record the environmental, social and economic dimensions of natural resource management and to support appropriate decision making and conflict resolution. However, the integration of spatial data in such environments has been problematic as the available spatial data often have different scale, content and formats. By building an appropriate SDI, disparate spatial data can be accessed and utilised to facilitate the exchange and sharing of spatial data between stakeholders across catchment communities.

SDI is a dynamic, hierarchic and multi disciplinary concept that encompasses policies, organisational remits, data, technologies, standards, delivery mechanisms and human resource dimensions (Rajabifard, 2007). SDI can also be viewed as a portal where

each stakeholder can access, use, and exchange spatial data for social, economic and environmental well-being (Feeney et al., 2001; McDougall, 2006). In many countries, SDI is regarded as a necessary component of the basic infrastructure required to efficiently support the operations and economic development of the nation. SDIs have been developed to manage and better utilise our spatial data assets by considering the needs and information flows from the local level, up through state, national and regional levels and finally to the global (GSDI) level. This has resulted in the emergence of varying forms of SDI at, and between, these levels (Hjelmager et al., 2008).

Although Australia is recognised internationally as a leader in SDI development and spatial information management, current SDI initiatives are more heavily dominated by national mapping agencies and state government organisations (Warnest, 2005). Currently, this SDI hierarchy is focused on SDI development at different political-administrative levels, ranging from local to state/provincial, national, regional and global levels (Chan and Williamson, 2001; Rajabifard and Williamson, 2001). However, catchment management issues cut across political-administrative boundaries and do not follow the rules of political-administrative hierarchies. Catchments have their own socio-spatial extents and coverage and could be considered to be community centric in nature and therefore more closely aligned to local government. Although local government has extensive local knowledge and experience in catchment management, an integrated management approach with a greater emphasis on community involvement is required to achieve sustainable catchment outcomes. Government organisations, business and community groups are the main stakeholders in catchment management. Unfortunately, the sharing of spatial data among these groups is generally characterised by a one way flow of spatial data. The majority of catchment data is government managed and there is limited spatial capacity within the many catchment groups. Therefore, to successfully address catchment management objectives, SDI frameworks must carefully consider the institutional arrangements and the needs of the various stakeholders across these catchment environments (Paudyal and McDougall, 2008).

The aim of this research paper is to explore a conceptual or theoretical framework for building or developing SDI from a catchment management perspective. The framework is then explored through a case study of the Murray-Darling Basin in Australia, one of the world's largest catchments. The challenges for developing an SDI to effectively support the decision making within and across this catchment will be discussed and the potential strengths and weakness of the proposed framework are identified.

2. CATCHMENT MANAGEMENT AND SDI

A catchment can be defined as a natural collection area where all rainfall and run-off water eventually flows to a creek, river, lake, ocean or into the groundwater system. Natural and human systems such as rivers, bush land, farms, dams, buildings, infrastructures, plants, animals and people co-exist in a catchment (Sydney Catchment Authority, 2008). Catchment management involves consideration of land use and land use change in relation to the land and water resources and the consequential effects on runoff and groundwater, as well as the effects of changes on land use (Laut and Taplin, 1989).

Catchment management is not readily amenable to systems analysis in a precise fashion, partly because of the complexity of the land, water and environment relationships and the lack of management tools capable of handling this in a spatial context. There are two main schools of thought in the catchment management doctrine namely: the total catchment management (TCM) and the integrated catchment management (ICM)

approaches. TCM is a holistic approach that seeks to integrate water and land management activities and the community and government involvement associated with these activities in a catchment. Total catchment management involves the co-ordinated use and management of land, water, vegetation, and other physical resources and activities within a catchment to ensure minimal degradation of the environment (Cunningham, 1986). The boundary of a catchment in the context of TCM is (at least in theory) the entire catchment, including all biophysical processes active within that catchment.

On the other hand, ICM has a philosophy for achieving the long-term sustainable use of land, water and related biological resources. It aims to coordinate the activities of landholders, community groups, industry groups and all spheres of government within the river catchment (CCMA, 2001). ICM mostly considers issues and problems which are known and whose affects are being felt by those within the catchment and is the management philosophy more commonly adopted by most jurisdictions in Australia.

Spatial data underpins decision-making for many disciplines (Clinton, 1994; Gore, 1998; Longley et al., 1999; Rajabifard et al., 2003a) including catchment management. It necessitates the integration of spatial data from different sources with varying scales, quality and currency to facilitate these catchment management decisions. However, the institutional arrangements for catchment management do not easily align with the SDI development perspectives as multiple stakeholders work to achieve multiple goals with government organisations, often guiding many catchment decisions.

SDI can facilitate access to the spatial data and services through improving the existing complex and multi-stakeholder decision-making process (Feeney, 2003; McDougall and Rajabifard, 2007). Moreover, it can facilitate (and coordinate) the exchange and sharing of spatial data between stakeholders within the spatial data community. A preliminary step toward achieving decision-making for catchment management has been the increasing recognition of the role of SDI to generate knowledge, identify problems, propose alternatives and define future courses of action (Paudyal and McDougall, 2008). In recent years, many countries have spent considerable resources on developing their own National Spatial Data Infrastructure (NSDI) to manage and utilise their spatial data assets more efficiently, reduce the costs of data production and eliminate duplication of data acquisition efforts (Masser, 2005; Rajabifard et al., 2003a).

Various researchers (Rajabifard et al., 2000; Rajabifard et al., 2002; Rajabifard and Williamson, 2001) argue that a model of SDI hierarchy that includes SDIs developed at different political-administrative levels is an effective tool for the better management and utilisation of spatial data assets. This SDI hierarchy is made up of inter-connected SDIs at corporate, local, state/provincial, national, regional (multi-national) and global levels. The relationship among different levels of SDIs is complex due to the dynamic, inter- and intra-jurisdictional nature of SDIs (Rajabifard et al., 2003a). However, this perspective, although useful, does not encompass the many complex relationships that operate between jurisdictions nor does it recognise the varying institutional objectives. The hierarchical model for SDI development therefore needs to be re-examined for the purpose of catchment management as catchment issues cut across jurisdictional and administrative/political boundaries.

Many countries are developing SDI at different levels ranging from corporate, local, state, national and regional to a global level, to better manage and utilise spatial data assets. Each SDI, at the local level or above, is primarily formed by the integration of spatial datasets originally developed for use in corporations operating at that level and

below (Rajabifard et al., 2003a). However, the catchment hierarchy is somewhat differ-ent to this administrative hierarchy. In catchment environments, the hierarchy begins from farm level and extends to the sub-catchment, catchment up to the basin level (see Figure 1).

Figure 1: Interrelation between administrative hierarchy and catchment hierarchy.
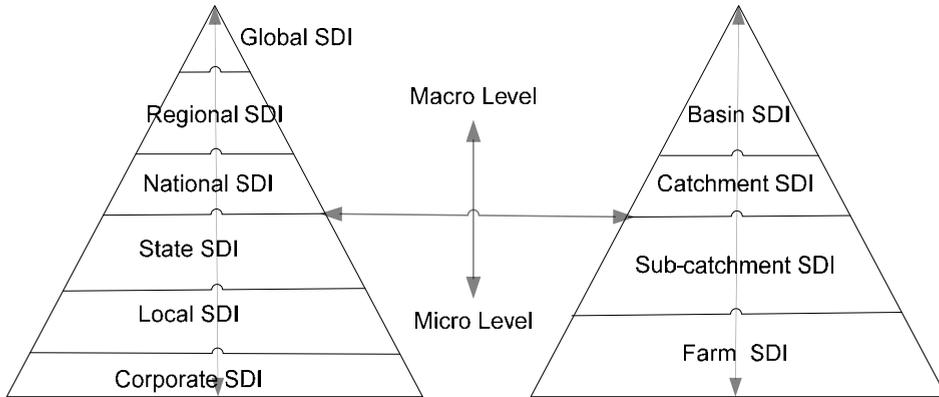
The existing SDI hierarchy for SDI development does not readily fit neatly with catch-ment management as their issues extend beyond the jurisdiction of administra-tive/political boundaries and can often cross the territorial boundaries of several coun-tries. Therefore, it is important to explore the extent to which hierarchical government environments contribute to the various components of SDI development and which SDI framework might be suitable for achieving catchment management objectives.

3. SDI THEORETICAL FOUNDATION

Many countries are developing SDI from the local to global level to better manage and utilise their spatial data for promoting economic development, to support better gov-ernment and to foster environmental sustainability (Masser, 1998). SDI development is supported by various theoretical backgrounds. The following section describes some of the important theories relevant to the development of SDI for catchment management.

3.1 Hierarchical Spatial Theory and SDI Hierarchy

In the past much research has been conducted toward maximising the efficiency of computational processes by using hierarchies to break complex tasks into smaller, simpler tasks (Car et al., 2001; Timpf and Frank, 1997). Examples of hierarchical appli-cations include classification of road networks (Car et al., 2000), development of politi-cal subdivisions and land-use classification (Timpf et al., 1992). The complexity of the spatial field, as highlighted by Timpf and Frank (1997), is primarily due to the space being continuous and viewed from an infinite number of perspectives at a range of scales.

Rajabifard et al. (2000) demonstrated that the principles and properties of hierarchical spatial reasoning could be applied to SDI research to better understand their complex nature and to assist modelling of SDI relationships. The hierarchical nature of SDI is well established in describing relationships between the administrative/political levels (Rajabifard et al., 2000). They support two views which represent the nature of the SDI

hierarchy namely; the umbrella view - in which SDI at the higher level encompasses all SDIs at a lower level, and the building block view - where a level of SDI such as at the state level, supports the SDI levels above (i.e. national, regional) with their spatial data needs. Rajabifard (2002) made use of hierarchical reasoning in his work on SDI structures in which a SDI hierarchy is made up of inter-connected SDIs at corporate, local, state/provincial, national, regional (multi-national) and global levels. In the model, a corporate SDI is deemed to be an SDI at the corporate level - the base level of the hierarchy. Each SDI, at the local level or above, is primarily formed by the integration of spatial datasets originally developed for use in corporations operating at that level and below. Hierarchical government environments have the potential to contribute to different components of SDI development and hence are important from a catchment management perspective.

3.2 Diffusion Theory and SDI Diffusion

Diffusion can be referred to as the process of communicating an innovation to and among the population of potential users who might choose to adopt or reject it (Zaltman et al., 1973) as cited by Pinto and Onsrud (1993). Gattiker (1990) views diffusion as 'the degree to which an innovation has become integrated into an economy'. He emphasises the relation between innovation and an economy. Spence (1994) describes diffusion as "the spread of a new idea from its source to the ultimate users". Diffusion can be viewed as 'the process by which an innovation is communicated through certain channels over time among the members of a social system' (Rogers, 1983). This definition gives rise to four elements of diffusion namely the innovation, the communication channel, time and the social system, which has constituted the foci of research activities in the past decades. Further, Rogers explains that it is a special type of communication in which the messages are about new ideas. The newness, in this case as highlighted by Chan and Williamson (2001) means that some degree of uncertainly is involved in diffusion.

The theory of diffusion as an innovation model (Rogers, 1995) is appropriate for the study of SDI diffusion, though the diffusion of innovations model has been criticised for its pro-innovation bias. This can be seen in the statements that are made in connection with SDI development which constantly stress its positive impacts in terms of promoting economic growth, better government and improved environmental sustainability (Masser, 1998). More than half the world's countries claim that they are involved in some form of SDI development (Crompvoets, 2006), but most of these initiatives can better be described as 'SDI like or SDI supporting initiatives'. Only a few countries can be described as having operational SDIs. The diffusion of SDI came from a tradition of SDI like thinking or national GI systems before SDI itself formally came into being.

Cultural factors are also likely to influence SDI adoption. De Man (2006) used a four dimensional model developed by Hofstede and Hofstede (2005) to assess the cultural influences on SDI development. They found that national cultures varied with respect to four main variables: power distance (from small to large), uncertainty avoidance (from weak to strong), masculinity versus femininity, and collectivism versus individualism. In a SDI environment, De Man argues that cultures where there are large power distances are likely to use SDI to reinforce the influence of management, whereas those with small power distances will be more receptive to data sharing and accountability. Both diffusion and innovation theory are potentially important to understanding the adoption of SDI within catchment management environments.

3.3 Evolution Theory and SDI Evolution

The creation of SDIs is a long term task that may take years or even decades in some cases before they are fully operational. This process is likely to be an evolving one that will also reflect the extent to which the organisations that are involved re-invent themselves over time (Masser, 2006). Rogers (1995) defines reinvention as 'the degree to which an innovation is changed or modified by a user, in the process of its adoption and implementation. The concept of SDI first emerged in the mid 1980s around the need for cooperation and sharing of spatially-related information across countries and organisations. In Australia, national land-related information initiatives commenced with a government conference in 1984 which eventually led to the formation of a committee responsible for national SDI development. Likewise, in USA discussion about the national SDI initiatives started around 1989, primarily in the academic community (National Research Council, 1999) and progressed rapidly after the executive order from the President's Office was issued in 1994 (Gore, 1998)

This national SDI development has been coined the first generation of national SDI initiatives and the motivations were in reducing duplication, using resources more effectively, and creating a base from which to expand industry productivity and the spatial market. It was a "product based" approach and the coordinators of SDI developments were dominated by National Mapping Agencies. The second generation of national SDI initiatives started around 2000 when some of the leading nations on SDI development changed their development strategies and updated their SDI conceptual models (Rajabifard et al., 2003b). This approach is "process based" and includes people as a component of SDI and the interoperability of data and resources. The concept of more independent organisational committees or partnership groups representative of different stakeholders is now tending to dominate SDI development.

3.4 Principal Agent Theory and Partnerships and Collaboration

According to neo-institution economics (NEI), the Principal-Agent (P-A) Theory which focuses on authority and sharing responsibilities (North, 1990) provides another relevant perspective for SDI development. In P-A relationships there are three aspects that are considered. The first aspect is the definition of who has authority/responsibility (principal) and who is carrying out work on the behalf of an authority (agent). The second aspect describes the extent to which a principal can control or check the agent, and the third considers the extent to which an agent can take on authority/responsibility. P-A theory may be useful in defining SDI partnerships or collaborations as there is often multi-level stakeholder participation in SDI implementation, particularly for catchment management.

Effective data sharing among participants is needed for SDIs to become fully operational in practice. Continuous and sustainable data sharing is likely to require considerable changes in the organisational cultures of the participants. To facilitate sharing, the GIS research and user communities must deal with both the technical and institutional aspects of collecting, structuring, analysing, presenting, disseminating, integrating and maintaining spatial data. For this reason there is a pressing need for more research on the nature of data sharing in multi level SDI environments. The studies that have been carried out by Nedovic-Budic and Pinto (1999) and Nedovic-Budic et al. (2004) in the USA provide a useful starting point for work in other parts of the world. Similarly, the findings of Harvey and Tulloch (2004) during their survey of local governments in Kentucky demonstrate the complexity of the networks involved in collaborative environments of this kind. Wehn de Montalvo's (2003) study of spatial data sharing percep-

tions and practices in South Africa from a social psychological perspective also highlights the issues associated with the sharing of data. This study which utilised the theory of planned behaviour found that the personal and organisational willingness to share data depends on attitudes to data sharing, social pressures to engage or not engage and perceived control over data sharing activities of key individuals within organisations. Likewise, McDougall (2006) reported on critical factors that impact on the success of partnerships for spatial data sharing including policy, governance, funding, leadership and vision.

As catchment management issues is characteristics by multilevel stakeholder participation in SDI implementation, the theories of hierarchies and P-A may assist in exploring in greater depth the context of building SDI at catchment scale. Table 1 summarises the various SDI theory, main contributors of that theory in spatial science domain, their characteristics, strengths, limitations, and value for catchment governance.

Table 1: Summary of SDI theoretical foundation and their contribution to catchment SDI development.

SDI Theory/ Citation	Contributors in Spatial Science Domain	Characteristics	Strength	Limitations	Value for Catchment Governance
Hierarchical Spatial Theory (Car, 1997)	(Car et al., 2000; Chan and Williamson, 2001; Rajabifard, 2002; Rajabifard et al., 2002; Timpf and Frank, 1997; Rajabifard et al.,2003)	Describes the vertical (inter) and horizontal (intra) relationships between different levels of SDIs.	Assist modelling of SDI relationships in structured environments	Horizontal relationships between different levels is not well addressed	Horizontal (intra) relationships between different levels of SDIs is useful
Diffusion Theory (Rogers, 1971; Zaltman et al., 1973)	(Campbell and Masser,1995; Chan and Williamson, 2001; Gattiker, 1990; Pinto and Onsrud, 1993; Rajabifard, 2002; Spence, 1994; Rajabifard et al., 2003)	Process of innovation of a new idea from its source to the ultimate users	Special type of communication in which the messages are about new idea	Innovation bias and some degree of uncertainty involved	Diffusion and adoption of innovation through the catchment community is important
Evolution Theory (Rogers, 1995)	(Rajabifard et al., 2003)	An innovation is changed over time or modified by a user	User centric and dynamic	May be less important over multiple organisations	Process based SDI model or new model is appropriate
P-A Theory (North, 1990)	(Harvey, 2001; McDougall, 2006; Nedovic-Budic and Pinto, 1999; Wehn de Montalvo, 2003)	Determine who has authority/responsibility and who is carrying on the behalf of authority	Useful for SDI partnership and collaboration	Does not cope with the theory of planned behaviour as organisational willingness is important for data sharing	Useful for data sharing and partnerships across catchments

Hierarchical spatial theory describes the vertical (inter) and horizontal (intra) relationships between different levels of SDIs. It assists the modelling and understanding of SDI relationships. The horizontal or intra-jurisdictional relationship between different hierarchies may not be easily accommodated by these theories. These relationships are particularly important for catchment governance. The diffusion theory describes the

spread of a new idea from its source to the ultimate users. The concept of SDI has emerged from developed economies and spread all over the world. Now, the developing countries are also initiating various forms of SDIs to improve the utilisation of their spatial data assets for economic and social well-being. The limitation of diffusion theory is that it has an innovation bias and a degree of uncertainty involved in it. Diffusion theory is also applicable for catchment management as new ideas are spread through the community and stakeholders via diffusion. The evolution theory (Rogers, 1995) describes the dynamic nature of SDI as an innovation that is changed over time or modified according to users' requirement. The first generation of SDIs (product based) evolved into second generation (process based) and included people as a component of SDI and the interoperability of data and resources. Now, the third generation of SDIs are evolving where users play a vital role for information management (Budhathoki et al., 2008; Goodchild, 2008). The advent of spatial technology and web services provides the way for more inclusive and open models of spatial services where grass-root citizens and community groups with no prior experience in spatial technologies can participate. Google Earth, OpenStreetMap (www.openstreetmap.org) and Wikimapia (www.wikimapia.org) are a few examples where the custodianship of spatial data is no longer in the hands of mapping agencies but the vast majority of society who are utilising these products. The application of SDI for catchment governance and management may well utilise a new conceptual model of SDI within this environment. The Principal-Agent theory is useful for gaining a better understanding of the relationships in sharing spatial data and partnership/collaboration. The first and most important task is identification of stakeholders and determining the interests, importance and influence. This could be determined by an interest power matrix (De Vries, 2003). This then enables strategies to be developed for community led stakeholders participation to support catchment governance and management.

4. CONCEPTUAL FRAMEWORK DERIVED FROM SDI THEORETICAL FOUNDATION

From our understanding of the various theories which relate to SDI development a conceptual framework can be explored for catchment SDI. It is assumed that there are basically two broad groups of stakeholders in catchment management namely, government and the community. Activities undertaken by land care groups or property owners at the grass root level will impact on large scale issues such as climate change, land use change and ecological system change. As Figure 2 illustrates, there are four management hierarchies in catchment governance including farms, sub-catchments, catchments and basin. The landcare groups, indigenous community members and individual land owners are the main stakeholders at the farm level which have horizontal relationships with local government and can share property-related spatial data in the form of a farm level SDI.

The sub-catchment authorities and other community groups share water, land and nature data with local government and sometimes other levels of government build sub-catchment SDI. Catchment authorities work towards the ecological sustainability of catchments. They share catchment data to state government and other levels of government. They work for the broad vision of natural resource management building catchment level SDI. The Basin SDI is the highest level of SDI hierarchy within the catchment management framework. The Basin SDI could be a part of Global Spatial Data Infrastructure (GSDI) or Regional SDI. In countries like Australia, Basin SDI covers the whole country or part of the country. For example, the Murray-Darling Catchment which stretches across four states and one territory is an example of Basin SDI. In some countries, it may cross the international boundaries.

Figure 2: Conceptual framework: Applying hierarchical spatial theory.
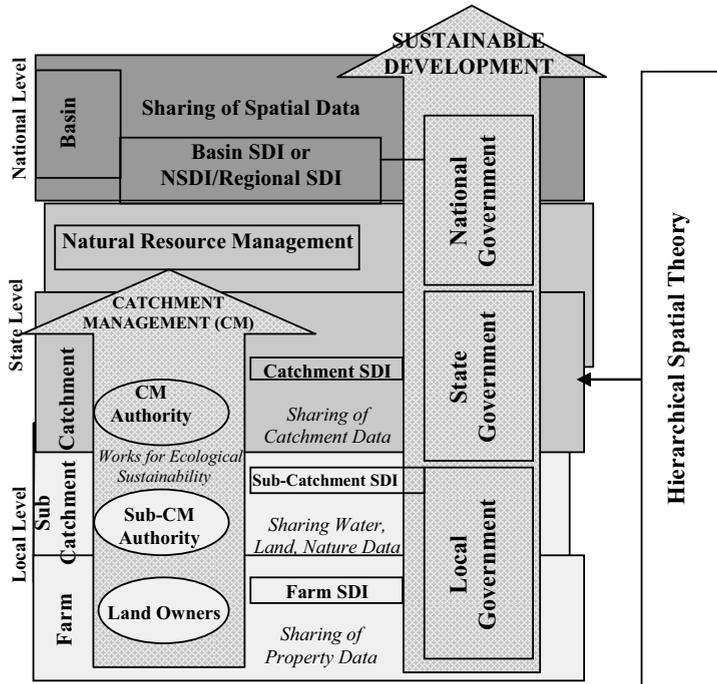
The emergence of catchment management authorities to facilitate improved local and regional outcomes for natural resource management now also introduces a multi-jurisdictional level of activity involving many stakeholders. Australia, like the USA, is a federation of states and understands the complexities of sharing and managing spatial data across three tiers of government. SDI development in Australia has been significantly constrained by these traditional jurisdictional structures which continue to slow our progress. Therefore, to support initiatives such as catchment management, it is important that new frameworks be examined which may facilitate improved SDI development at the catchment level.

The proposed framework modelled on the hierarchical spatial theory has a number of strengths and limitations. Firstly, if we examine the strengths of the proposed framework, we already know and understand the many formal and informal hierarchical structures and processes exist within a catchment environment. These structures and processes enable the modelling of responsibilities and hence potential data flows. For example, hierarchies of catchment SDI already fit nicely with existing management groups such as land care, farming groups and catchment management authorities. Secondly, stakeholders interact in a hierarchical fashion in many instances in line with existing institutional arrangements. Finally, the catchment authority's goals are often aligned to government priorities/goals and therefore a hierarchical framework is perhaps appropriate.

However, the framework also has a number of potential weaknesses. Perhaps the most obvious of these is the complex and large number of levels and cross-jurisdictional linkages which have the potential to dilute information flows and create

institutional complexities. This is particularly evident where the hierarchy in catchment SDI and administrative/political SDI do not align.

5. CASE STUDY

The purpose of this case study is to examine the proposed conceptual framework in the context of an operational catchment environment. The case study to be examined is the Murray-Darling Basin (MDB) which is an area of national significance for social, cultural, economic and environmental reasons in Australia. Administratively, the MDB falls under the four state government jurisdictions, namely Queensland, Victoria, New South Wales, South Australia and one territory, the Australian Capital Territory as shown in Figure 3. It includes the catchment of Australia's three longest rivers, the Darling (2,740 km), Murray (2,530 km) and Murrumbidgee (1,690 km) and their many tributaries (Australian Bureau of Statistics 2008). Both the MDB community and governments are partners in protecting the health and productivity of the MDB.

Figure 3: Case study area (Murray-Darling Basin Authority (MDBA)).

In the Murray Darling Basin, there are 22 Catchment Management Authorities (CMA) which work at local level forming catchment authorities and sub-catchment authorities for integrated catchment management. In addition, there are various volunteer groups (like landcare, bushcare, coastcare) and indigenous communities which also work at the grass-root level to achieve the integrated catchment management goals (Australian Bureau of Statistics, 2008). The three tier government structure (commonwealth, state and local) also exists to manage and utilise the resources of the basin in a way that is economically sustainable. Among the 22 CMAs, 4 are in Queensland, 9 are in New South Wales, 5 are in Victoria, 3 are in South Australia and 1 is in Australian Capital Territory. There are many overlaps and gaps between catchment boundaries and the administrative boundaries in Murray-Darling Basin. Figure 4 highlights the management hierarchies in catchment governance in the MDB.

Figure 4: Catchment management hierarchies in MDB.
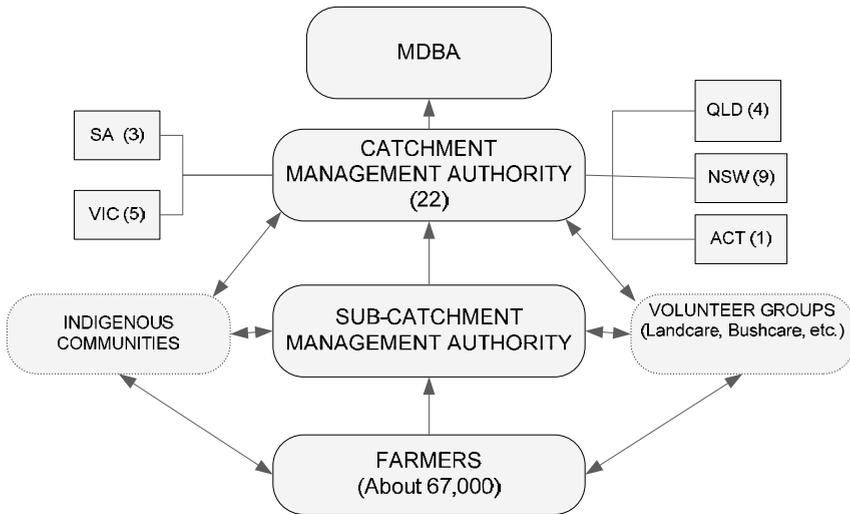
6. DISCUSSION

The spatial data obtained from MDBC and Australian Bureau of Statistics has been used to analyse the spatial interaction across the existing local governments and the catchments. Using spatial analysis tools, it can be shown that many catchments overlap a number of local and state government boundaries. Table 2 shows the number of local government boundaries which within individual catchment boundaries. It is interesting to note that a large number of local authorities (more than 60%) straddle catchment boundaries, although the catchments are often larger than the local government authorities.

Table 2 illustrates the institutional complexities for building SDI for catchment management. The proposed conceptual framework in section 4 has been examined using the case study of Murray-Darling Basin. As described in the conceptual framework, the main players are government organisations and community groups for catchment governance in MDB. The hierarchies of catchment management fit nicely with existing management groups such as land care, farming groups, indigenous communities and catchment management authorities as shown in Figure 4. There are good practices where stakeholders interact in a hierarchical fashion for better environmental outcomes

with existing institutional arrangements. Therefore, the proposed framework modelled on the hierarchical spatial theory is considered appropriate for building SDI at catchment level. However, the hierarchies in catchment and administrative/political SDI do not align so effective cross-jurisdictional linkages will be required to improve the efficiency of information flows and institutional arrangements. The large number of local government authorities and the disparity of spatial extents and boundaries require new and innovative approaches to manage spatial data across these environments.

Table 2: Local authorities status with catchment boundaries.

STATE (Name)	CMA (Number)	LOCAL GOVERNMENT AUTHORITIES (LGAs)			
		Number of LGAs that fall within catchment boundary	Number of LGAs that straddle catchment boundary (number)	Total	Proportion of LGAs that straddle catchment boundary in each state
QLD	4	9	29	38	76%
NSW	9	30	48	78	62%
VIC	5	10	24	34	71%
SA	3	4	15	19	79%
ACT	1	1	0	1	0%
Total	**22**	**54**	**116**	**170**	

7. CONCLUSIONS

Spatial data and the development of SDI offer great potential for catchment managers and decision makers. Current SDI initiatives are generally dominated by national mapping agencies and state government organisations and modelled on the existing administrative/political hierarchies. However, catchment management issues do not follow the rules of these hierarchies and are community centric in nature. Therefore, there is a need to re-examine SDI development approaches to accommodate the needs of catchment governance and management.

Hierarchy theory holds some promise for building the community-government interaction required for SDIs at various catchment levels such as farm, sub-catchment, catchment and basin level. This framework is complex, having potentially many levels and linkages. The case study of the Murray-Darling Basin illustrates the complexity of the catchment management environment with a large number of local government authorities and a disparity of spatial extents and boundaries. There is no doubt that SDI holds some promise in solving these complex data management problems and can contribute the final goal of delivering improved catchment management outcomes.

ACKNOWLEDGEMENTS

REFERENCES

Australian Bureau of Statistics (2008). *Water and the Murray-Darling Basin - A Statistical Profile, 2000-01 to 2005-06* Murray-Darling Basin uses the majority of Australia's water. ABS (Media Release).

Budhathoki, N., B. Bruce and Z. Nedovic-Budic. (2008). Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal,* 72(3): 149-60.

Car, A., G. Taylor and C. Brunsdon (2001). An analysis of the performance of a hierarchical way finding computational model using synthetic graphs. *Computers, Environment and Urban Systems* 25(1): 69-88.

CCMA (2001). *Condamine Catchment Strategic Plan* Condamine Catchment Management Association Inc., Toowoomba QLD.

Chan, T.O. and I.P. Williamson. (2001). The different identities of GIS and GIS diffusion. *International Journal of Geographical Information Science,* 13(3): 267-81.

Clinton, W.J. (1994). Coordinating geographic data acquisition and access to the National Spatial Data Infrastructure. Executive Order 12096. Washington, DC, US, at: http://govinfo.library.unt.edu/npr/library/direct/orders/20fa.html.

Crompvoets, J. (2006). *National Spatial Data Clearinghouses: Worldwide Development and Impact*. Wageningen University.

Cunningham, G.M. (1986). Total Catchment Management, Resource Management for the Future. *Journal on Soil Conservation* N.S.W. Special Edition: Total Catchment Management, 42(1): 4-5.

De Man, W.H.E. (2006). Understanding SDI: complexity and institutionalization. *International Journal of Geographical Information Science,* 20(3): 329-43.

De Vries, W.T. (2003). "Tools to Map and Identify Stakeholders" in PD Lecture Notes. ITC, Enschede, The Netherlands.

Feeney, M.-E.F. (2003). "SDIs and Decision Support" in I.P. Williamson, A. Rajabifard and M.-E.F. Feeney (Eds.). *Developing Spatial Data Infrastructures: from Concept to Reality,* London, UK: Taylor & Francis, pp. 195-210.

Feeney, M.-E.F., A. Rajabifard and I.P. Williamson (2001). Spatial Data Infrastructure Frameworks to Support Decision - Making for Sustainable Development. Cartagena de Indias, Columbia, 21-25 May, 2001.

Gattiker, U.E. (1990). *Technology Management in Organizations.* Newbury Park: Sage Publications.

Goodchild, M. (2008). Commentary: whither VGI? *GeoJournal,* 72(3): 239-44.

Gore, A. (1998). The Digital Earth: understanding our planet in the 21st century. Digital Earth, Wikipedia, at: http://en.wikipedia.org/wiki/Digital_Earth.

Harvey, F. and D. Tulloch (2004). "How do governments share and coordinate geographic information? issues in the United States", *Proceedings of 10th EC & GIS workshop*, Warsaw, Poland.

Hjelmager, J., H. Moellering, A. Cooper, T. Delgado, A. Rajabifard, P. Rapant, D. Danko, M. Huet, D. Laurent, H. Aalders, A. Iwaniak, P. Abad, U. Dren and A. Martynenko (2008). An initial formal model for spatial data infrastructures, *International Journal of Geographical Information Science,* 22(11&12): 1295 - 309.

Hofstede, G. and G.J. Hofstede (2005). *Cultures and organizations: software of the mind,* New York: McGraw-Hill.

Laut, P. and B.J. Taplin (1989). *Catchment Management in Australia in the 1980s*, Institute of Natural Resources and Environment, Canberra ACT 2601 Australia.

Longley, P.A., M.F. Goodchild, M. D.J. and D.W.E. Rhind (eds.) (1999). *Geographic Information Systems*. Second Edition, New York: John Wiley & Sons.

Masser, I. (1998). The first generation of national geographic information strategies. Canberra, Australia, 17-19 November.

Masser, I. (2006). What's Special about SDI Related Research? *International Journal of Spatial Data Infrastructures Research,* 1: 14-23.

Masser, I. (2005). *GIS Worlds, Creating Spatial Data Infrastructures*. Redlands, California: ESRI Press.

McDougall, K. (2006). *A Local-State Government Spatial Data Sharing Partnership Model to Facilitate SDI Development*. The University of Melbourne.

McDougall, K. and A. Rajabifard (2007). "A Mixed-Method Approach for Evaluating Spatial Data Sharing Partnerships for Spatial Data Infrastructure Development", in H.J. Onsrud (ed.). *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, Redlands, California, USA: ESRI Press, at: http://gsdidocs.org/gsdiconf/GSDI-9/papers/TS22.3paper.pdf

National Research Council (1999). *Distributed Geolibraries: Spatial Information Resources*, Washington, D.C.: National Academy Press.

Nedovic-Budic, Z. and J.K. Pinto (1999). Understanding interorganizational GIS activities: A conceptual Framework. *Journal of Urban and Regional Information Systems Association,* 11(1): 53-64.

Nedovic-Budic, Z., M.-E. Feeny, A. Rajabifard and I. Williamson (2004). Are SDIs serving the needs of local planning? Case study of Victoria, Australia and Illinois, USA. *Computers, Environment and Urban Systems,* 28(2): 329-51.

North, D.C. (1990). *Institutions, institutional change and economic behaviour.* Cambridge University Press.

Paudyal, D.R. and K. McDougall (2008). "Building Spatial Data Infrastructure to Support Sustainable Catchment Management", *Proceedings of Queensland Spatial Conference 2008*, Gold Coast, Queensland, 17-19 July, 2008.

Pinto, J.K. and H.J. Onsrud (1993). "Correlating adoption factors and adopter characteristics with successful use of geographic information systems", in I. Masser and H.J. Onsrud (Eds.). *Diffusion and Use of Geographic Information Technologies*. Dordrecht, Netherlands: Kluwer Academic publishers.

Rajabifard, A. (2002). *Diffusion of Regional Spatial Data Infrastructures: with particular reference to Asia and the Pacific*. The University of Melbourne, Melbourne, Australia.

Rajabifard, A. and I. Williamson (2001). "Spatial data infrastructures: concept, SDI hierarchy and future directions", *Proceedings of GEOMATICS'80 Conference*, Tehran, Iran.

Rajabifard, A., F. Escobar and I.P. Williamson (2000). Hierarchical Spatial Reasoning Applied to Spatial Data Infrastructures, *Cartography Journal Australia,* 29(2): 41-50.

Rajabifard, A., M.-E.F. Feeney and I.P. Williamson (2002). Directions for the Future of SDI Development, *International Journal of Applied Earth Observation and Geoinformation,* 4(1): 11-22.

Rajabifard, A., M.-E.F. Feeney and I. Williamson (2003a). "Spatial Data Infrastructures: Concept, Nature and SDI Hierarchy", in I. Williamson, A. Rajabifard and M.-E.F. Feeney (Eds.), *Developing Spatial Data Infrastructures: from Concept to Reality,* London, UK: Taylor & Francis, pp. 17-40.

Rajabifard, A., M.-E.F. Feeney, I. Williamson and I. Masser (2003b). "National SDI Initiatives", in I. Williamson, A. Rajabifard and M.-E.F. Feeney (Eds.). *Developing Spatial Data Infrastructures: from Concept to Reality,* London, UK: Taylor & Francis, pp. 94-110.

Rajabifard, A. (2007). "SDI Design to Facilitate Spatially Enabled Society", in A. Rajabifard (ed.). *Towards a Spatially Enabled Society*, Melbourne: The University of Melbourne, pp. 219-232.

Rogers, E.M. (1983). *Diffusion of Innovations,* New York: The Free Press.

Rogers, E.M. (1995). *Diffusion of Innovations,* New York: The Free Press.

Spence, W.R. (1994). *Innovation: The Communication of Change in Ideas, Practices and Products,* London, UK: Chapman & Hall.

Sydney Catchment Authority (2008). What is a catchment? at: http://www.sca.nsw.gov.au/the-catchments.

Timpf, S. and A.U. Frank (1997). "Using hierarchical spatial data structures for hierarchical spatial reasoning", *Lecture Notes in Computer Science (*Proceedings of International Conference COSIT'97), 1329: 69-83.

Timpf, S., Volta, G.S., Pollock, D.W., and Egenhofer, M.J. (1992). A Conceptual Model of Wayfinding Using Multiple Levels of Abstractions*, Lecture Notes in Computer Science* (Proceedings of Theories and Methods of Spatio-Temporal Reasoning in Geographic Space)*,* 639: 348-367.

Warnest, M. (2005). *A collaboration model for national spatial data infrastructure in federated countries*. The University of Melbourne, Australia.

Wehn de Montalvo, U. (2003). *Mapping the determinants of spatial data sharing.* Aldershot: Ashgate.

Zaltman, G., R. Duncan and J. Holbeck (1973). *Innovations and Organisations.* New York: John Wiley & Son.