SPATIAL ENABLEMENT

IN SUPPORT OF ECONOMIC DEVELOPMENT AND POVERTY REDUCTION

RESEARCH, DEVELOPMENT AND EDUCATION PERSPECTIVES

EDITORS HARLAN ONSRUD & ABBAS RAJABIFARD

GSDI Association Press

Spatial Enablement in Support of Economic Development and Poverty Reduction

Research, Development and Education Perspectives

SPATIAL ENABLEMENT IN SUPPORT OF ECONOMIC DEVELOPMENT AND POVERTY REDUCTION

RESEARCH, DEVELOPMENT AND EDUCATION PERSPECTIVES

EDITED BY
HARLAN ONSRUD and ABBAS RAJABIFARD

GSDI ASSOCIATION PRESS

Spatial Enablement in Support of Economic Development and Poverty Reduction

Harlan Onsrud and Abbas Rajabifard (Editors)

ISBN 978-0-9852444-2-2

Cover design by Matthew Hamilton

Table of Content

Foreword

This book is the result of a collaborative initiative between the Global Spatial Data Infrastructure Association (GSDI), the School of Computing and Information Science at the University of Maine, and the Centre for SDIs and Land Administration (CSDILA) in the Department of Infrastructure Engineering at the University of Melbourne. The articles featured in this peer-reviewed book were mostly the result of the traditional Call for Papers for the GSDI 14 Global Geospatial Conference "Spatial Enablement in Support of Economic Development and Poverty Reduction", but also contains contributions of full articles which were solicited for publication in this book.

The authors and reviewers were advised of the theme in advance and, in most cases, they addressed this theme in their papers. Even in cases where the theme was not directly referenced, the article reflected the impact and application of spatial data infrastructures that are now being developed worldwide. The peer-review process resulted in 15 chapters that when considered together, reflect how SDIs are enabling us all today, particularly in meeting the global challenges of poverty and sustainable economic development.

We thank the authors of the chapters and the members of the Peer Review Board. We are grateful to the GSDI Association Press for its willingness to publish this work under a Creative Common Attribution 3.0 License. It allows all to use the experiences and research presented in this book to their own best advantage. We would like to thank Dr Hamed Olfat, Ms Serene Ho and Ms Pamela Chew for their editorial assistance in preparation of this publication, as well as Mr Matthew Hamilton for the design of the cover.

Harlan Onsrud and Abbas Rajabifard (Editors)
GSDI Association

Peer Review Board

About Editors

Harlan Onsrud is a Professor of Spatial Information Science and Engineering in the School of Computing and Information Science at the University of Maine and a research scientist with the National Center for Geographic Information and Analysis (NCGIA). His research and teaching focuses on the analysis of legal, ethical, and institutional issues affecting the creation and use of digital databases and the assessment of the social and societal impacts of spatial and tracking technologies. He is past president and current Executive Director of the Global Spatial Data Infrastructure Association (GSDI), past-president of the University Consortium for Geographic Information Science (UCGIS), and past Chair of the U.S. National Committee (USNC) on Data for Science and Technology (CODATA) of the National Research Council. He has participated in several U.S. National Research Council studies related to spatial data and services and has been funded as a Fulbright Specialist in Law with assignments in Australia and Germany.

Abbas Rajabifard is a Professor of Geomatics Engineering, Head of Department of Infrastructure Engineering and Director of the Centre for SDIs and Land Administration at the University of Melbourne. He is Past President of the Global Spatial Data Infrastructure Association (GSDI) and a member of Victorian Spatial Council (VSC). He was Vice Chair, Spatially Enabled Government Working Group of the UN supported Permanent Committee on GIS Infrastructure for Asia and the Pacific (PCGIAP). He has authored/co-authored over 240 publications including 8 books. He has been also consulted widely on spatial data management, SDI, land administration and spatial enablement, to many national government agencies and ministries.

PART ONE

FUNDAMENTAL FUNCTIONALITIES AND FRAMEWORKS IN SDIS

CHAPTER 1

Spatial Enablement, Poverty Reduction and Economic Development

Harlan Onsrud[1] and Abbas Rajabifard[2]

[1] Spatial Informatics Faculty, School of Computing and Information Science, University of Maine, USA
[2] Department of Infrastructure Engineering, The University of Melbourne, Australia

harlan.onsrud@maine.edu; abbas.r@unimelb.edu.au

1. Introduction

Since the early 1990s, the concept of an SDI has progressively entered into the lexicon of governments all around the world and gained an increasingly prominent profile as an enabling infrastructure, critical to development by linking information to location. The development of jurisdictional, national, regional and global SDI initiatives has become a matter of priority for many governments, with associated concepts such as spatially enabled governments and spatially enabled societies becoming a focal point. These numerous initiatives have benefitted from various global coordinating bodies, including the Global Spatial Data Infrastructure (GSDI) Association, and most recently, the United Nations Initiative on Global Geospatial Information Management (UN-GGIM). As location is fast being recognized as the fourth driver in decision-making, spatial data and SDIs are being leveraged to address some of the most pressing challenges facing the world today, including poverty and economic development, which is the focus of the theme of this book.

Governments from around the world, through their adoption of the United Nations Millennium Declaration in September 2000, have already committed to working towards the reduction of extreme poverty and achieving pro-poor development. However, in some countries, the gap between rich and poor is in fact even widening, compounded by a range of economic issues stemming from a litany of inequalities – gender, urban/rural divide, health and employment, among others. In today's society, where market structures are increasingly interconnected at a global level, there is less of a buffer for vulnerable countries in terms of impacts from global and regional events, including the recent series of catastrophic natural disasters that have occurred.

As a community, spatial professionals, now more than ever, have an important role to play in bringing their skills and knowledge to the fore to support public, private and non-governmental sectors, academia, and local communities around the world. Through spatial enablement of information, we can facilitate efforts to ensure decision-making is informed by evidence; that resources are delivered to the communities that are most vulnerable; that knowledge can flow from the grassroots level (from the local 'experts') to integrate with more formal, bureaucratic sources of information to foster participation and community awareness to build resilience to environmental or economic events.

It is therefore timely that the theme of the GSDI 14 World Conference, Addis Ababa 2013, is "Spatial Enablement in Support of Economic Development and Poverty Reduction". The collection of articles in this book provides a contribution from our profession to demonstrate the continual impact and new opportunities that spatial information and technologies present in the design of solutions and strategies to empower communities at all levels. In doing so, we hope to contribute in some way to this global push towards mitigating the extreme poverty that continues to be a reality for many around the world, and improving economic outcomes for future generations.

2. Spatial Data, Poverty Reduction and Economic Development

Many now acknowledge the importance of information and communication technologies (ICT) in underpinning economic development. It is therefore no surprise that increasing emphasis is being placed on the role of such 'soft' infrastructure (Button, 2002) on economic development, and hence its impact on both poverty and inequality (e.g. UN-Habitat 2011). Its importance is underscored by the specific reference the Millennium Development Goals make to such infrastructure types.

The impact that ICT has had on poverty reduction and economic development is well illustrated in the growing use of mobile phones in developing countries, particularly in rural areas (see Table 1 below). Although constrained in both connectivity and functionality – most still use asynchronous connections and rely only on voice and SMS services due to high data costs – mobile phone usage has undoubtedly stimulated a grassroots revolution in terms of access to services. For example, the humble SMS service has enabled many farmers in rural areas in improving their bargaining position by being better informed about price fluctuations and market activities (Kochi, 2012). UNICEF's RapidSMS initiative (a scalable SMS-based open source framework) has been utilized for many applications including health and logistics, but more importantly has been crucial in empowering local communities by collecting data on grassroots issues through SMS messages (UNICEF Innovation, 2013).

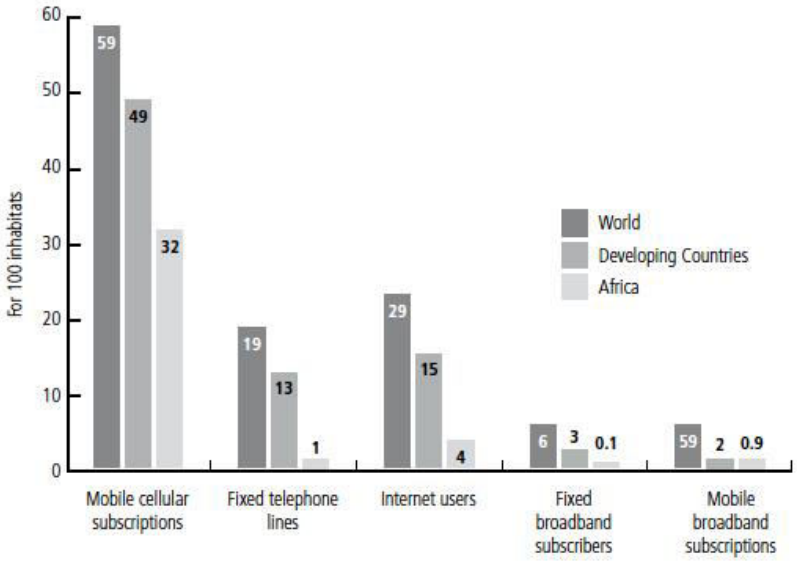

Table 1. ICT Penetration in Africa in 2009 (International Telecommunication Union, 2009 in UN-Habitat 2011: 35)

Against this backdrop, and closely aligned with the notion of ICT being transformational technologies (Coleman and McLaughlin, 1998), the geospatial community, who from as far back as the 1960s had long championed the benefits of integrated spatial information for improved analysis (e.g. Tomlinson, 1967), started applying the ideals of improved data sharing and information infrastructures to spatial data, resulting in the conceptualization of SDIs.

SDIs first appeared in the mid-1980s, and these 'first generation' versions were designed to promote economic development through supporting the objectives of governments and supporting environmental sustainability (Masser, 1998). These early initiatives which focused on data and data accessibility, soon led to the development of a product-based approach to SDI development that was driven by national governments around the world (Rajabifard *et al*., 2003). Throughout the 2000s, this conceptualization of SDIs began to shift towards a more user-oriented approach, where the focus moved towards the management of data, or a more process-based approach (Rajabifard *et al*., 2006).

Today, due to differing levels of maturity among countries, both first and second generation SDIs are still fairly common. As such, there is still a considerable level of research effort focused on SDI framework and development. However, there is a trend towards a broadening of the SDI research agenda, as seen in Figure 1 below, which was based on an analysis of over 2,000 research articles. Increasingly, SDI research is also beginning to explore applied areas such as SDI for disaster management, smart cities and e-Government. There is also greater emphasis on the services SDIs can deliver

through improving the technical aspects, such as data management and use of technology, as well as evaluating the impact of SDIs on the economy and wider society.

Figure 1. Broadening of the SDI research agenda

There is no doubt that spatial data and SDIs are now being used in many different capacities – particularly in the coordination, analysis and use of large-scale, people-relevant data. In the area of poverty management, spatial data has become fundamental to this discipline by revealing poverty distribution patterns otherwise hidden by national aggregated information, and providing a way to connect poverty with a range of social, economic and environmental factors through location, enabling the identification of key poverty variables. Common spatial datasets about the physical environment include soil information, topography, rainfall and vegetation (Hyman *et al*., 2005); other datasets regarding social or economic indicators include distance to market and transport costs (Van de Walle, 2002; Jacoby, 2000). The ability to drill down to detailed local-level analysis is essential to the design and delivery of poverty reduction programs (Baker and Grosh, 1994; Bigman and Fofack, 2000; Elbers *et al*., 2004). SDIs therefore continue to play a key role in facilitating these information needs, and in developing countries with limited data and resources, SDIs can provide a more cost-effective approach in data production by reducing data duplication.

The issues that contribute to poverty and economic development are also not typically confined to the constraints of administrative boundaries. Civil unrest, natural disasters, food insecurity and other environmental events all result in large-scale issues such as population displacement, unemployment, lack of secure housing and limited food and water supply. To deal with these issues effectively requires input from many

governmental, non-governmental and aid agencies across national boundaries, as well as a coordinated and collaborative approach across organizations with differing information needs: by providing an enabling platform, SDIs can play a crucial role in facilitating these activities. It is therefore also not surprising that the use of SDIs are becoming more widespread in the related activity of disaster management, given the central role spatial data plays in all aspects of disaster management, including the design of mitigation and preparedness strategies.

The applications and utility of SDIs are continually evolving, as the technical constructs continue to develop, leveraging the latest ICT developments to better support its users and deliver improved outcomes. Some of the articles in this book demonstrate how new technologies such as cloud computing and data mining can be used in the context of SDIs. However, intrinsically, the success of SDIs and any spatial enablement continues to rest not on the technical aspect, but on the ability of people to continue to reach across organizational and administrative divides to nurture an environment that supports information sharing to achieve better outcomes for the community. Therefore, despite advances in technology, some of the strategic considerations that underpin SDI development continue to remain consistent and there is still much work that can be done to support and advance knowledge in these areas.

3. Book Outline

This book is a compilation of articles as book chapters each focusing on different aspects the application of spatial data or spatial technologies as common infrastructure to facilitate poverty reduction or economic development.

The chapters presented in this book have gone through a full peer review process as part of the joint and fully integrated GSDI 14 World Conference and AfricaGIS Conference in 2013. The chapters represent a range of views that have been categorized as: (1) Fundamental Functionalities and Frameworks in SDIs, (2) Empowering Communities – Participatory Applications for SDIs, and (3) Spatial Enablement in Support of Development.

Part 1: Fundamental Functionalities and Frameworks in SDIs

In this first section, a collection of six articles illustrate the theoretical and strategic considerations relevant to the fundamental functionalities and frameworks in SDIs. In chapter two, *Towards Modeling the SDI Supply Chain in South Africa: the Case of Land Administration Data*, Edward Kurwakumire, Serena Coetzee and Peter Schmitz apply the concept of supply-chain models to provide a business process-oriented perspective of an SDI. Given the SDI's central role in facilitating access and discovery of spatial information for a jurisdiction, it is imperative that an SDI functions as efficiently as possible in the delivery of data and services. The application of a supply chain model to analyze the provision of land administration data in this chapter identifies the key

stakeholders and their interactions. An understanding of this environment provides a mechanism for tracking errors and improving the quality of spatial data.

Continuing on the theme of spatial data quality, metadata is fundamental to the ability of an SDI to connect users to data effectively and efficiently. In the current climate where the amount of spatial data that is being created continues to grow, many SDI initiatives are struggling to manage metadata creation, let alone any efforts to update and improve metadata content. In chapter three, *Design and Development of a Spatial Metadata Automation Framework Applied in Australia,* Hamed Olfat, Abbas Rajabifard, Mohsen Kalantari and Chris Pettit present the outcomes of a research project undertaken at the University of Melbourne. A case study approach is used to identify current requirements in metadata management and automation. An assessment of metadata management tools commonly used in the geospatial community is also presented. These provide the basis for the development of a framework that leverages a new GML-based integrated data model for storing and bundling spatial data and metadata. Such an approach not only facilitates and automates spatial metadata creation, but also its enrichment, based on end users' participation. The practical benefits of this framework are then showcased through an industry-based implementation for the Australian Urban Research Infrastructure Network (AURIN).

Not every SDI initiative is successful and there are often a variety of lessons that can be learnt to be fed back into the next undertaking, or to evince more generic learnings. In chapter four, *A Description of SDI Stakeholders in Ghana Using the ICA Model,* Wiafe Owusu-Banahene, Foster Mensah, Serena Coetzee, Antony K Cooper, Victoria Rautenbach, Kisco Sinvula, Emma Nangolo and Martin Hippondoka have set out to apply the categorisation of SDI stakeholders developed by the International Cartographic Association to the now defunct National Framework for Geospatial Information Management (NAFGIM) in Ghana. Their research is important on several fronts: many of the stakeholders in NAFGIM have reprised their roles in the current SDI development underway in Ghana. The mapping of stakeholders and their relationships provides summative feedback into the current initiative; as well, it can provide the basis of a pro-active strategy towards facilitating collaboration and participation. The use of the ICA model provides a platform for future comparative work on other jurisdictions, important both in terms of benchmarking as well as a way to further improve the ICA model.

It is widely recognized that the effectiveness of an SDI depends on the uptake of spatial data use and sharing by organizations in support of their processes. In chapter five, *Analyzing Organizational Levers of Spatial Enablement*, Ezra Dessers, Joep Crompvoets and Geert Van Hootegem demonstrate through a case study approach, how the level of spatial enablement can be correlated to the level of embeddedness between spatial data and organizational activities in the form of an integrated process. The authors set out to identify potential organizational levers of spatial enablement in the public sector in the region of Flanders by looking at the relationship between task division and spatial enablement and the relationship between the allocation of the spatial data function and the level of spatial enablement. In the first relationship, the

research suggests that the presence of an integrated process could be related to a higher level of spatial enablement, while research pertaining to the second relationship suggests a relation between spatial data function and spatial enablement. These findings are instructive for organizations involved in spatial data sharing in SDI initiatives.

In the final chapter of Part 1, the importance of assessing SDI activities underpins Garfield Giff and John Jackson's work, *Towards an Online Self-Assessment Methodology for SDIs.* The assessment of SDI activities remains an important consideration for the continual improvement and development of SDI initiatives and the authors contribute to the existing body of work by developing an efficient and cost-effective comprehensive integrated enterprise GIS/SDI assessment model to facilitate the assessment of SDIs, primarily from the perspective of its stakeholders. Such a perspective is important, as the function of an SDI is contingent on the participation of multiple stakeholders on several levels. The authors present an online self-assessment tool as a potential method for gathering feedback and evaluation from stakeholders. This tool, reflecting the hierarchical structure of SDIs, is therefore based on an assessment of the geospatial performance of lower-tier stakeholders as well as an assessment of the different levels of the SDI.

Part 2: Empowering Communities – Participatory Applications for SDIs

From the functioning of SDIs, this next section proceeds to consider the applications of SDIs to spatially enable communities of users. Empowering communities and facilitating greater grassroots participation is key to building resilience to weathering the impact of development. The articles in this section consider a range of issues, but are mainly focused on leveraging local knowledge sources. The ability to integrate crowd-sourced data with more formal sources of spatial data continues to provide rich and relevant opportunities for research in the SDI domain.

The first chapter in this section, *Public Participatory GIS, Spatial Data Infrastructure, and Citizen-Inclusive Collaborative Governance* by Michael Sutherland, Titus Tienaah, Amit Seeram, Bheshem Ramlal and Susan Nichols follows the proposition that views of the local community, in terms of Volunteered Geographic Information (VGI) inputs to Public Participatory GIS (PPGIS), can augment, complement or verify decision-making processes in a collaborative governance model. The authors put forward a prototype system, based on open-source software, which is capable of combing both empirical data and VGI while adhering to SDI standards. VGI inputs in the form of spatial objects are integrated within the PPGIS with empirical data from authoritative sources and positions communities as valid SDI data contributors. Legitimacy in the system is maintained through restricted ability to update content as well as the use of a moderator. The authors also recognize that open source tools reduce barriers to participation by providing communities and local governments with free rights to use, modify and redistribute copies for various projects. However, use of such platforms is contingent on Internet connection, which is often unreliable or slow in developing countries, where communities stand to benefit the most from such initiatives.

Dev Raj Paudyal, Kevin McDougall and Armando Apan use a case study on natural resource management (NRM) organizations to highlight motivations and barriers in spatial information sharing within a use community in their chapter, *Developing Spatial Information Sharing Strategies across Natural Resource Management Communities*. Through a broad questionnaire and follow-up interviews, the authors show that despite the existence of formal agreements of inter-organizational collaboration and recognition of the importance of knowledge transfer, most NRM organizations maintained a silo approach to spatial information management – a real barrier to spatial information sharing. On the basis of key factors which were shown to influence data sharing, spatial information sharing strategies were developed along the themes of governance, policy, economic, legal, cultural and technical aspects. While specific to the NRM community, these strategies can potentially provide insight for other SDI initiatives and suggest potential paths towards facilitating greater spatial data sharing.

The environmental consequences of industrial activities in pursuit of economic development pose a real threat to the viability and sustainability of communities, particularly those in remote areas. The ability to articulate an index of vulnerability can go a long way towards informing policies and increasing awareness at the local level. In chapter nine, *Application Of Multi-Criteria Decision Analysis and GIS Techniques in Vulnerability Assessment of Coastal Inhabitants in Nigeria to Crude Oil Production and Transportation Activities,* Omoleomo Olutoyin Omo-Irabor and Samuel Bamidele Olobaniyi present the use of multi-criteria decision analysis (MCDA) in conjunction with remote sensing and GIS techniques to undertake vulnerability assessment of coastal communities that may be impacted by crude oil production and transportation activities. Their assessment framework included environmental, social and economic criteria in impact assessment activities. Community resilience was articulated through scoring of adaptive capacity (based on eight socio-economic indicators) and human vulnerability. The study showed that most communities had poor to moderate adaptive capacities and required greater capacity building for the inhabitants to be equipped in dealing with threats posed by oil pollution. More broadly, the study shows the significance of this joint application of GIS and MCDA in assessing the impact of man's activities on the environment.

In the final chapter of this section*, Resource-Constrained Agriculture in Developing Countries and Where Geo-ICT Can Help,* Clarisse Kagoyire and Rolf de By set out to identify the range of geospatial information that can best be used to spatially enable sustainable agricultural activities in an ICT limited environment. Agricultural activities are often affected by factors along the supply chain that are commonly location-specific and through a case study, the authors argue that sustainable agriculture is contingent on the exchange of relevant geoinformation supported by geo-ICT. Specifically, the authors tap into local farming knowledge to facilitate more active participation in the production of location-specific information to facilitate decision-making. This information is used alongside the outcomes of spatially analyzed geo-referenced constraint-based farm-plot profiles, which generate new insights into the impact of various constraints on coffee farming.

Part 3: Spatial Enablement in Support of Development

The final section of this book focuses on spatial enablement in support of development. The various articles focus on the use of technology in support of spatial enablement by improving discovery, management and use of spatial data.

The first chapter in this section considers a perennial issue in development – how best to use available land resources sustainably. In *GIS-based Land Suitability Assessment for Optimum Allocation of Land to Foster Sustainable Development: the Case of the Special Zone of Oromia Regional State around Addis Ababa City, Ethiopia,* Dessalegn Gurmessa and Sileshi Nemomissa use GIS and remote-sensing technologies to undertake analysis land use suitability in the special zone of Oromia Regional State. Despite only having a small percentage of land classified as suitable for both crop and livestock production, more than half the study area is currently used for these purposes. The authors' analysis demonstrates the value of spatial technologies in improving understanding of land use suitability for decision-making. This would result in more appropriate use and allocation of land resources, which would facilitate sustainability of land productivity in regions whose economies are dependent on crop and livestock production.

The prolific use of the Internet has led to huge quantities of data being produced, leading often to information overload and reduced accessibility and discoverability. Web ontologies that define the semantic attributes of data are providing a way forward in data-mining activities where semantics are used as a way to link disparate search results and terms to create an awareness of users' activities and preferences. In this chapter, *Method of Context-Aware Recommender System Based on Ontologies,* Guillermo González Suárez, Tatiana Delgado Fernández, José Luis Capote Fernández and Rafael Cruz Iglesias have developed a recommender system specifically for an SDI environment by relying on spatial, semantic and collaboration filters to analyze data and preferences from mobile users to suggest more personalized search results. Such context-aware and targeting data mining provides users with more effective and directed access to spatial information. The authors demonstrate that such a system can improve the analytical capabilities of an SDI.

Like many developing countries, the use of spatial data to support national development is well-recognized in Rwanda. However, the government faces a challenge that is common to all SDI initiatives – overcoming data sharing barriers that exist between government departments due to a range of factors. In *A Discovery Geospatial Portal for Promoting Geo-ICT Use in Rwanda,* Felicia O. Akinyemi, Bernard Hakizimana and Jean Damascene Mazimpaka develop a geoportal that leverages web-mapping services to facilitate data sharing amongst government organizations while enabling them to retain ownership of data. This is proposed as a potential way to facilitate and improve data sharing, access, use and dissemination. Like many other developing countries where cost of implementation can be a limiting factor, the authors' prototype demonstrates that web-mapping services can be implemented at a minimal cost by using both commercial and free open-source software.

The importance of creating liveable and sustainable urban environments is widely recognized, and many international city ranking and benchmarking systems exist. However, at the neighborhood level, the data required for planners to enact local change and support decision making remain isolated within different local and state government departments. In chapter fourteen, *Spatially Enabling Information to Support Livability: A Case Study from the North Melbourne Metropolitan Region Australia,* Serryn Eagleson and Abbas Rajabifard describe a project carried out through the Centre for SDIs and Land Administration at the University of Melbourne which included the development of an open-source platform in the context of a Spatial Data Infrastructure (SDI) to facilitate the access and distribution of a series of integrated spatial datasets pertinent to the design of liveable neighborhoods. The value of integrated spatial data is demonstrated through the development of four Web-based tools to support decision making through scenario-testing. The project has integrated over 100 datasets from disparate sources and now provides them to support researchers from across Australia.

In this final chapter, *Cloud GIS in Geothermal Resource Data Management: A Case Study of the Kenya Electricity Generating Company,* Daniel Waweru Mwaura and Hunja Waithaka utilize cloud computing to propose a solution in response to the need of the Kenya Electricity Generating Company for harmonizing and sharing centrally stored geothermal data across a number of departments. The learnings from this case study have broader applications in terms of addressing common SDI network problems in terms of data sharing, management and retrieval. Recent developments in cloud computing and hosting services are leveraged in this prototype system. The case-study based implementation of the prototype system proves that cloud-based GIS is viable alternative for distributed spatial data use and management, as well as parallel processing of large datasets, which is a significant issue for many organizations who continue to rely on desktop-based data processing. The authors conclude with potential challenges for cloud-based GIS including a lack of existing legislation specific to cloud computing which offers protection against data security as well as a 'buyer-beware' position adopted by cloud hosting facilities indemnifying them against data losses or leaks.

References

Baker, J. and Grosh, M., (1994), Poverty Reduction through Geographic Targeting: How Well Does it Work? *World Development*, 22(7): 983-995.

Bigman, D., Fofack, H., (2000), Geographic Targeting for Poverty Alleviation: Methodology and Applications. World Bank Regional and Sectoral Studies. Washington DC.

Button, K., (2002), Effective Infrastructure Policies to Foster Integrated Economic Development. Paper Presented at the Third African Development Forum, Addis Ababa, March.

Coleman, D.J. and McLaughlin, J.D., (1998), Defining Global Geospatial Data Infrastructure (GGDI): Components, stakeholders and interfaces. *Geomatica*, 52(2): 129-143.

Elbers, C., Lanjouw, J., Lanjouw, P., (2003), Micro-level Estimation of Poverty and Inequality. *Econometrica,* 71(1): 355-364.

Federal Geographic Data Committee (1994), Executive Order 12906: Coordinating Geographic Data Acquisition And Access: The National Spatial Data Infrastructure**,** Federal Register, 59(71): 17671-17674.

Hyman, G., Larrea, C. and Farrow, A., 2005. Methods, Results and Policy Implications of Poverty and Food Security Mapping Assessments. *Food Policy*, 30(5-6): 453-60.

Jacoby, H., (2000), Access to Markets and the Benefits of Rural Roads. *Economic Journal,* 110: 713-737.

Kochi, E., (2012), How the Future of Mobile Lies in the Developing World, T*echCrunch*. Available from http://techcrunch.com/2012/05/27/mobile-developing-world/ [Accessed 26 Sept 2013].

Masser, I., (1998), The First Generation of National Geographic Information Strategies. In *Proceedings of Selected Conference Papers of the Third Global Spatial Data Infrastructure Conference*, Canberra.

Rajabifard, A., Feeney, M. F., Williamson, I. and Masser, I., (2003), National Spatial Data Infrastructures. In I. Williamson, A. Rajabifard, and M.-E. F. Feeney (eds.), *Developing Spatial Data Infrastructures: From Concept To Reality*, London: Taylor & Francis.

Rajabifard, A., Binns, A., Masser, I. and Williamson, I., (2006), The Role of Subnational Government and the Private Sector in Future Spatial Data Infrastructures. *International Journal of GIS,* 20(7): 727-41.

Tomlinson, R.F., (1967), An Introduction to the Geo-information System of the Canada Land Inventory, Department of Forestry and Rural Development, Government of Canada. Ottawa, Canada.

UN-Habitat (2011), Infrastructure for Economic Development and Poverty Reduction in Africa, *The Global Urban Economic Dialogue Series*, Nairobi: UN-Habitat.

Van De Walle, D., (2002), Choosing Rural Road Investments to Help Reduce Poverty. *World Development*, 30(4): 575-589.

CHAPTER 2

Towards Modeling the SDI Supply Chain in South Africa: The Case of Land Administration Data

Edward Kurwakumire, Serena Coetzee and Peter Schmitz

University of Pretoria, South Africa

kurwakumir18860@itc.nl; serena.coetzee@up.ac.za; pschmitz@csir.co.za

Abstract

A spatial data infrastructure (SDI) is a complex integrated network of spatial data producers, distributors and consumers that can be viewed as an extended geographic information enterprise. A way of gaining a better understanding of an SDI is to break it down into its constituent components so that individual SDI entities and their interactions can be analyzed. Prior SDI literature suggests analyzing SDIs as complex adaptive systems as they are dynamic rather than static in terms of behavioral aspects. The supply-chain model has been used to map, model and analyze complex business processes. This study views an SDI from the supply chain perspective by describing the business processes towards the creation of spatial data sets and the participation of different actors in this value-addition process. We discuss the benefits of applying the supply chain model not only in modeling the processes but also in managing the SDI as a whole. We model the supply chain for land administration data and discuss its relevance in analyzing the SDI. SDIs are crucial to sustainable development and thus, it is of importance that they operate effectively. Supply chain management has a history of better managing, monitoring and improving the efficiency of individual organizations within the manufacturing industry. This chapter details the applicability of supply chain management in improving its operational efficiency and effectiveness of the SDI but with focus on land administration data.

KEYWORDS: Spatial data infrastructure, SDI, geographic information, supply chain

1. Introduction

Virtually all public, private and non-governmental organizations, including the general public, use spatial data for various applications (Genovese *et al.*, 2008; Genovese *et al.*,

2009; Akinyemi, 2011). Access to spatial data is crucial for the sustainable and economic development of a nation (Campagna, 2006; Welle Donker, 2009; Makanga and Smit, 2010, Welle Donker *et al*., 2010). A spatial data infrastructure (SDI) supports the hosting, discovery, publishing and access of geographically referenced (spatial) information (Makanga and Smit, 2010). It is of importance for the SDI to be as efficient as possible in delivering spatial data and services as this ensures availability of data, which is one of the core objectives of the SDI. To understand how efficiency in availability and accessibility of data and services can be achieved, there is a need to understand the processes that occur from the initial production to the delivery of the final product.

A supply chain is the integrated process wherein a number of entities (i.e. suppliers, manufacturers, distributors and retailers) work together to acquire raw materials, convert these into specified final products, and deliver the final products to retailers (Beamon, 1998). Supply chains are characterized by a forward flow of products and a backward flow of information (Simchi-Levi *et al*., 1999; Min and Zhou, 2002). Supply chain theory offers tools to map, manage and analyze processes in the production of spatial data, including initial data collection, pre-processing, value addition, inventory control, transportation and distribution of geographic products. In this study, we consider the SDI to be a supply chain. We apply general supply chain theory and show where we deviate (or specialize) due to the nature of spatial data and services as a product.

2. Background

2.1 Spatial Data Infrastructures

An SDI may be defined as a framework to facilitate the management of information assets with a focus on better communication channels for the community for sharing and using data sets (Rajabifard *et al*., 2002). *"SDI is about facilitation and coordination of the exchange, sharing, accessibility, and use of spatial data within the spatial data community with standardization and reutilization as important functional properties"* (Crompvoets *et al*., 2010). Steudler *et al.*, (2008) defines an SDI as follows: *"The SDI is fundamentally a concept about facilitating and coordinating the exchange and sharing of spatial data between stakeholders from different jurisdictional levels in the spatial data community.*

A spatial data infrastructure delivers data and services that can be used by various agencies, including the public, for a variety of applications. SDIs focus on fulfilling the goals of different users, yet these goals often conflict when user requirements are considered individually. This is the SDI complexity issue, which propagates due to the fact that they are dynamic, multi-disciplinary and comprise many components. The definition of SDI can change depending on the implementation objectives of the users (Hendriks, 2012).

2.2 The Supply Chain and the Value Chain

A supply chain is an integrated system which synchronizes a series of inter-related business processes in order to: (1) acquire raw materials and parts; (2) transform these raw materials and parts into finished products; (3) add value to these products; (4) distribute and promote these products to either retailers or customers; (5) facilitate information exchange among various business entities (e.g. suppliers, manufacturers, distributors, third-party logistics providers and retailers) (Min and Zhou, 2002). The supply chain is characterized by a forward flow of goods and backward flow of information (Simchi-Levi *et al*., 1999; Min and Zhou, 2002), as depicted in Figure 1. A supply chain does not necessarily represent a linear chain of one-on-one relationships, but rather a web of multiple networks and relationships (Min and Zhou, 2002). A supply chain map enhances *"the strategic planning process, ease distribution of key information, facilitate supply chain redesign or modification, clarify channel dynamics, provide a common perspective, enhance communications, enable monitoring of supply chain strategy and provide a basis for supply chain analysis*" (Gardner and Cooper, 2003: p39).

Supply chain management is closely linked to logistics. According to Frazelle (2002), logistics "*is the flow of material, information, and money between consumers and suppliers*". Supply chain management builds upon the logistics framework with an aim to achieve linkage and co-ordination between the processes of other entities in the network namely suppliers, customers, and the organization itself (Christopher, 2011). A supply chain in essence involves all entities or actors who are involved directly or indirectly in fulfilling a customer's request making the customer an integral part of the chain. The required data characteristics by the end user determine the costs of production (NAS, 2004).

The terms supply chain and value chain are often used synonymously and interchangeably within logistics and supply-chain management literature, however, a supply chain is much more integrated and broader than a value chain. The value chain focuses on the processes that add value to products as they progress through the chain. On the other hand, the supply chain encompasses the value-chain processes, as well as distribution, transportation and inventory control. It also includes the backward flow of feedback information to improve and adjust the supply chain and its products. The value chain can be extended to analyze factors influencing industry performance, including access to and requirements of terminal markets, the legal, regulatory and policy environment, coordination between firms in the industry and the level and quality of support services (Campbell and Kula, 2006).

2.3 Modeling SDI Supply Chains

In spatial data terms, the supply chain "encompasses all activities associated with the flow and transformation of spatial and attribute data from the raw data stage (capturing), through to the end user, as well as the associated information and money

flows." Data, information and money flow up and down the supply chain (Schmitz, 2008). Refer to Figure 1.

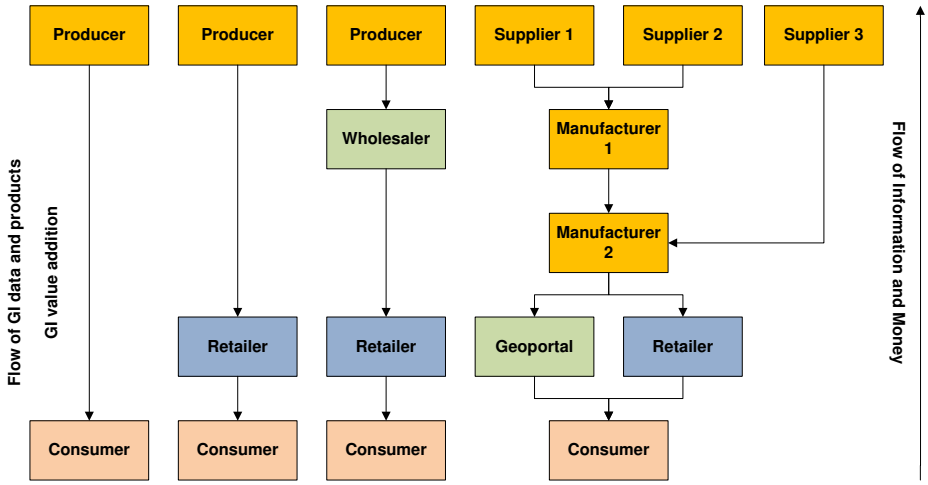

Figure 1. The flow of data, information and money up and down the supply chain, adapted from (Simchi-Levi *et al.*, 1999; Schmitz, 2007)

A geographical information system (GIS) has been defined in supply chain terminology in (Schmitz, 2008) as "a computer-assisted system, combined with appropriate infrastructures, resources and management, that acquires input data from suppliers, performs inventory and warehousing through data storage and retrieval, creates and value adds and delivers geographical and related non-geographical data to customers". The SDI can be viewed as an extended GIS enterprise or rather a countrywide distributed network of geographical information systems comprising both visible and virtual networks of acquiring raw materials (input data), manufacturing (processing, transformation and value addition to geographic information) and delivery of various data and customized information products to a wide range of users.

On the far end of the supply chain, the SDI ensures availability of data through clearinghouses and accessibility through multiple channels, including different vendors, value-added resellers and other delivery mechanisms, such as portals. The SDI can extend across international borders, for example, the Infrastructure for Spatial Information in the European Community (INSPIRE) (Welle Donker, 2010) initiative was adopted by 27 member states of the European Union.

The spatial data supply chain in an SDI is not a simple linear sequential structure but rather a complex network (Crompvoets *et al.*, 2010). Spatial data and services are produced through the interaction of different suppliers' suppliers (intermediaries). Each intermediary contributes to the value the geographic information. Economists refer this phenomenon as the 'value chain'. The value chain with respect to spatial data is described in (Krek and Frank, 2000; Genovese *et al.*, 2009). The paradox for

spatial data is that the fixed cost of collecting the data and keeping it current is very high at the beginning of the value chain when the value of the individual products is still low. Put in other words: geographic information is expensive to collect while the dissemination is generally inexpensive (Longhorn and Blakemore, 2008). This impacts on pricing models.

There are a number of similarities between the SDI and the value chain described by Campbell and Kula (2006). These include legal and regulatory issues [licensing, policies and copyright], coordination [data collection, processing and delivery as well as joint efforts between different stakeholders] and support services [information and communication technologies that ensure connectivity and communication between entities or organizations and availability, accessibility and delivery of data]. In this study, we map the land administration supply chain of the SDI in South Africa.

2.4 Related Work

This section reviews prior work in describing and managing SDIs as infrastructures. The push factors are to address the complex nature of SDIs and the manner in which the creation of spatial data can be viewed as a supply chain. The production of spatial data is a value addition process (AZOIC, 2010) in which there is a forward flow of materials and a backward flow of information and money as described in the supply chain definition by Simchi-Levi *et al*., (1999) and Min and Zhou (2002). van Leonen and van Rij (2008) view the creation of framework data sets which are core to the SDI as a networked effort requiring organizations to coordinate data collection. This is similar to the concept of SDIs being supply chains that is presented in this study. The production of spatial data requires collaboration from different stakeholders who participate in the value addition processes of geographic information. This value addition process is part of the supply chain as described in section 2.2.

SDIs are complex systems whose performance needs to be monitored based on theories that can accommodate the complexities. For example, an SDI can be viewed as a complex adaptive system (Grus *et al*., 2008; Grus *et al*., 2010). Grus *et al*., (2010) view SDIs as both complex and dynamic in nature. The dynamic nature of the SDI implies that the behavior of the system is unpredictable which requires assessors that detect and assess both the predictable and unpredictable changes. Supply chains and their management have been evolving as organizations are increasingly connected by information technology. The complexity of the supply chain changes also due to the expansion of businesses across wider geographical networks and the dynamic customer needs. In this regard, supply chains are not static by dynamic in the same context of SDIs as described by Grus *et al*., (2010). Several issues are presented in Kurwakumire (2013) on managing the performance of geographic information infrastructures (GII). One of the major difficulties in the monitoring and management is on defining the boundaries of the infrastructure. On the other hand, supply chain management can assists in making visible the whole production process, even of spatial data, up to delivery to the end user. This can be used as the basis for demarcating the SDI boundaries for the purposes of monitoring performance.

Krek and Frank (2000) described the value chain for spatial data in general, whereas we model the supply chain for land administration data. Furthermore, in our research we model not only the forward moving value chain but also the supply chain of land administration data, which includes feedback information flowing backwards in the chain. Such information can be used to track errors and improve the quality of products.

ANZLIC (2010) views the production and access of spatial data as a value addition process referred to as a value chain. The value chain is defined as *"a modification process of raw spatial data into the final products and services that fit end user requirements"*. The user requirements in this case are the drivers in product development process. The ANZLIC study acknowledges that several actors participate in transforming spatial data into the final product demanded by the customer. Making spatial data products and services available and accessible is one of the objects of the SDI. This view of the production of spatial data is important in this study as it describes the manufacturing process of products, which is necessary when mapping out the SDI supply chain. The SDI is more than a delivery mechanism but also incorporates manufacturing processes, order processing, inventory management and warehousing as partly depicted in Figure 2.

Schmitz (2007, 2008) employs supply chain management in determining whether efficiency and effectiveness of Geographic Information Systems (GIS) units in product delivery can be enhanced. This was achieved through using the Supply Chain Operations Reference model (SCOR) which is an industry-based model endorsed by the Supply Chain Council. The SCOR model uses 5 management principles namely (1) plan, (2) source, (3), make, (4) deliver and (5) return, in managing supply chains. Schmitz analyzes the relevance of supply chains in improving the production and delivery of spatial data by a GIS unit. This is accomplished through studying supply chains from the manufacturing industry in order to determine their applicability as models for managing geographic information systems. The focus in (Schmitz, 2007) was on a GIS unit housed by Eskom, which is an electricity distributor in South Africa. Schmitz (2007, 2008) focused on modeling the supply chain on a corporate level. This study focuses on modeling the supply chain at an inter-organizational level as the SDI extends across different organizations in a wider geographical space and more complex network.

To map the SDI supply chain there is a need to identify the role players (suppliers, customers and service providers) as they collectively form the supply chain through their interactions. Hjelmager *et al*., (2008) identified and described six stakeholders of an SDI and recognized that an individual stakeholder can execute different roles. For example, an organization can act as a policy maker, who sets out rules and policies for an SDI, and at the same time, be a producer of data and services required in an SDI. Sinvula *et al*., (2013) applied the stakeholder model to the Namibian SDI in order to improve the understanding of the SDI. However, they did not model interactions between the stakeholders, which are modeled in this chapter.

The literature reviewed in this section focuses on (1) the networked effort in the creation of spatial data in a value-chain process, (2) SDI performance monitoring, (3) the complex and dynamic nature of SDIs and (4) application of supply chains in assessing GIS units. This review is relevant to this study for the purposes of defining what constitutes the SDI. The necessity of assessing the efficiency and effectiveness of SDIs has also been highlighted. This study utilizes supply-chain management principles in order to make the SDI supply chain visible. This research builds up on the work presented in (Schmitz, 2007) by extending the application of supply chains to mapping the SDI. The SDI, unlike a GIS unit, extends across a wider geographical area and developing a mechanism for making its supply chain visible is necessary to improve the efficiency in providing spatial data. Supply chains are used to map complex business processes which cuts across different companies and industries. SDIs present a similar complex network that resembles an extended enterprise. Supply-chain mapping improves distribution of information and communication within the chain (Gardner and Cooper, 2003).

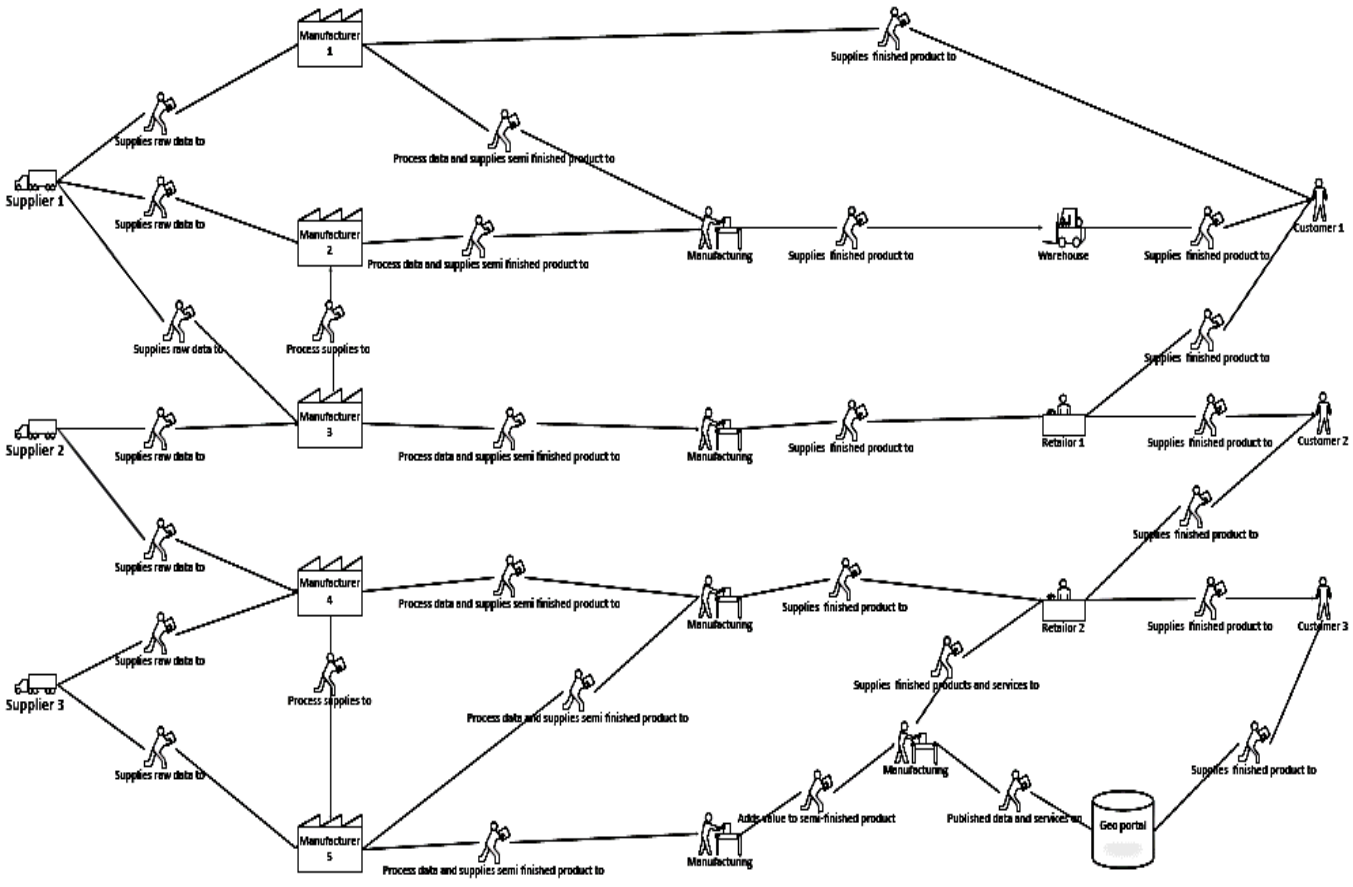

Figure 2. Theoretical SDI supply chain

3. The SDI Supply Chain for Land Administration Data in South Africa

The focal point of this study is the Chief Surveyor General's Office (CSG) since the CSG has been identified as one of the potential custodians of cadastral data in the South African SDI (CSI, 2012). The CSG is the organization where land administration data activities are centralized. Data comes into the CSG mainly through the cadastral survey process. Several stakeholders, including real estate agencies, municipalities and utility companies, utilize these data for their own applications.

3.1 The SDI Supply Chain

The SDI should ensure accessibility to data and services (i.e. products) to the various types of users. Access to reliable data needs to be timely and efficient, i.e. there is a need to optimize not only the transaction time for a data request, but also the supply chain from raw data collection to final user product. Figure 2 maps the SDI supply chain at the highest level of abstraction. It shows the transformation of raw spatial data into various final products. This is achieved through a value-addition process. The concept presented in Figure 2 is that within an SDI, there are several stakeholders who interact in a network to prepare the final product required by the customer as detailed by ANZLIC (2010), Hjelmager *et al.*, (2008) and Krek and Frank (2000). Different customers desire differing products, which are delivered to them through various means. Figure 2 depicts the processes of sourcing raw materials, transforming them to final products and distribution to the end user. This is a precise reflection of the supply chain definition by Beamon (1998) given in section 1 of this chapter.

There are various costs that are incurred as products progress from one supplier to another to include raw material, transportation, production and inventory costs (Simchi-Levi *et al.*, 1999) and these need to be kept at a minimum as they affect the price of the final product. The SDI supply chain should be cost effective, so that services will be less costly to the customer. Modeling the SDI offers a mechanism for improving communication between the different role players in the supply chain process. In this context, spatial data organizations do not operate in isolation, but in a network, thus communication improves the operational efficiency of the supply chain. Users are key to the SDI supply chain as they determine the type of product and characteristics as well as when it is required.

3.2 The SDI Supply Chain for Land Administration Data in South Africa

In this section, we consider the supply chain for the production of cadastral data set by the CSG. This data set has been identified as one of the core geospatial datasets of the South African SDI (CSI, 2012). We consider a use case for a request for land development done through the City of Johannesburg (CoJ) (Joburg, 2011). We map only the production processes until the data is in the data store at the CSG.

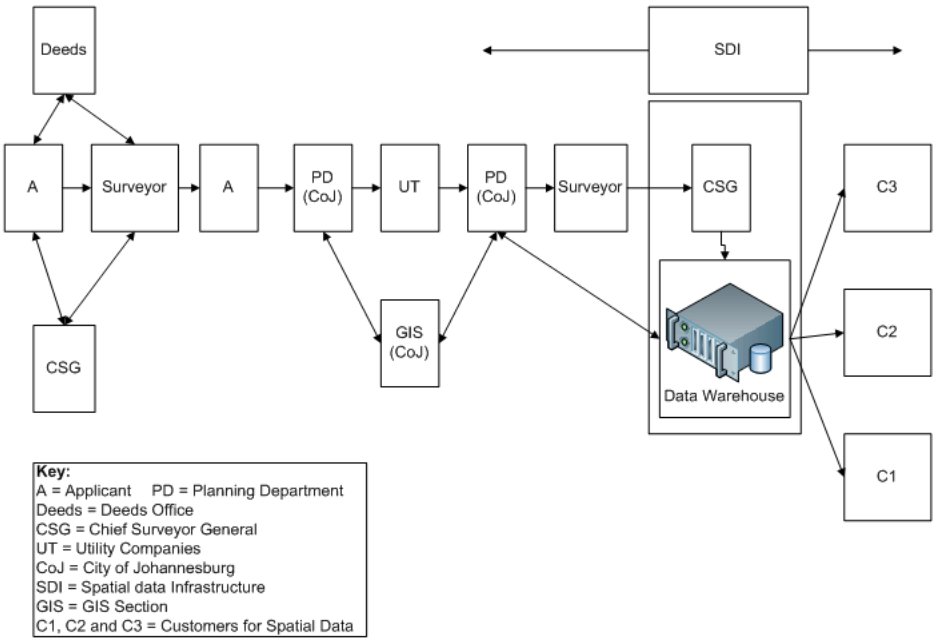

Figure 3. Thread diagram illustrating the land administration supply chain

The process is described in subsequent paragraphs and the corresponding supply chain is illustrated in Figure 3. An applicant who wants to develop all or part of his land into a township lodges his application to the planning department (PD) at CoJ and the application details are captured into the development management database. As part of the application portfolio, the applicant submits proof of ownership or authority to request the development. This is in the form of a title deed collected from the Deeds Office and a survey diagram from the CSG.

The PD evaluates the application through the Township Application System (TAS), assigns a reference number and feedback is to be sent to the applicant within a given timeframe. The PD requests a township layout plan from the applicant and in order to obtain this, the applicant has to appoint a professional land surveyor to prepare the plan. The surveyor may need to request additional data from the CSG to aid in developing the layout plan such as diagrams of adjacent properties, coordinate lists or topographical maps. The proposed layout plan is submitted to the PD at the CoJ by the applicant and captured as a proposed layout by the GIS department at CoJ, a copy is stored with the application documents at the PD, which is part of development planning.

The PD and the GIS department then propose a name for the new township. The PD sends the proposed township layout to utility companies for them to check if their bylaws have been abided to. The PD also checks adherence to zoning restrictions, servitudes and other public rights. The utility companies then provide feedback on

their evaluation of the proposed layout back to the PD. If all conditions are successfully met, the applicant gets positive feedback and the layout is sent to the CSG for approval. If there is negative feedback, the applicant has to ensure adjustments are done to the proposed layout according to the feedback.

The CSG evaluates the layout plan and other submitted documents for quality (accuracy, consistency, adherence to land survey records and completeness) and if all conditions are met to the minimum quality requirements, the layout is approved by the CSG. The approved documents are captured into the databank at the CSG. If conditions are not met, the submitted survey records are returned to the surveyor for correction until the required standards are met. The PD accesses the approved layout from the CSG through the website which is captured by the GIS department as part of the cadastral information. The application and approved layout is then published in the government gazette.

4. Discussion

Supply chains allow us to visualize the different players that form the SDI network and the different ways in which they interact. The land administration supply chain in Figure 3 identifies the different players: the applicant, the Deeds Office, the CSG, the professional land surveyor, the PD, and utility companies.

The chain provides a mechanism for tracking errors and improving the quality of spatial data. For example, the planning department can identify zoning violations in a submitted layout plan or inconsistencies with existing infrastructure. Quality is improved when the land surveyor corrects the layout plan and resubmits. Managing the supply chain aids in achieving customer value through producing and delivering spatial data products demanded by customers, rather than products planned for production by the manufacturers.

The chain makes it possible to analyze the different sharing or exchange mechanisms between different entities in the chain and the different licensing methods, partnerships and agreements that exist. For example, commercial companies wishing to add value and resell spatial data buy licenses while there is free access to the utility companies during the evaluation of a township application. Consultants who perform contract work for CoJ sign data declarations and have unpaid access to the data requirements for a project.

Figure 3 is the use case for the SDI supply chain, but focusing on land administration data. It conceptualizes the supply chain map presented in Figure 2 but utilizes a real world example. The SDI is a complex network, but to demonstrate the concept of supply chain mapping, the thread diagram (refer to Figure 3), is necessary. The simple illustration shows the interaction of seven actors within the SDI network before the final products can be availed to the customers. The diagram only demonstrates the

flow of information, products and money, yet there are other costs as shown in Figure 4.

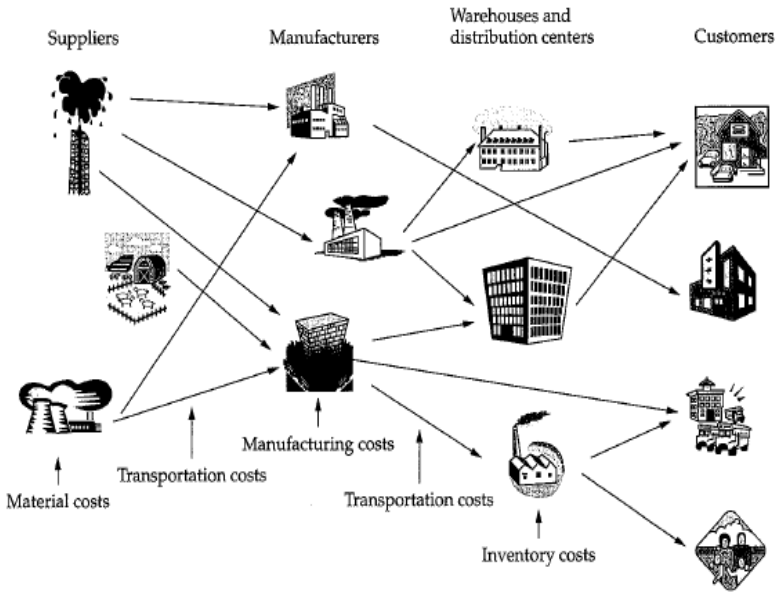

Figure 4. The logistics network, Source: (Simchi-Levi *et al*., 1999)

From each customer, it is possible to start a new supply chain, which is an extension of this use case. For example, consider customer 1 to be a real estate agency. Customer 1 would be interested in cadastral maps, street data and property information, which is integrated (*further value addition*) and deployed on a website that property seekers use. In this regard, customer 1 is now a supplier of spatial data to an even possible, wider audience.

Analyzing the SDI supply chain as a whole can assist in identifying suppliers that seemed invisible and new markets to distribute products, thus improving the business intelligence for both suppliers and customers.

Managing the SDI as a supply chain network can enable geographic products to be developed and delivered efficiently to customers while achieving some resource optimization. Efficiency in this context refers to delivery of the demanded product, in the right quantity, quality and time period that meets the satisfaction of the customer or user.

The use case presented is not comprehensive enough to demonstrate the complexity of the SDI. It is possible to map better this supply chain using the SCOR model utilized in (Schmitz, 2007). It is then possible to visualize the value addition from one entity to another, the warehousing and distribution costs and mechanisms, the products that are being demanded and delivery times. There are quality control procedures available

at each step in the chain and important service providers such as telecommunications and courier companies, which make the SDI network, work. The SCOR model makes a wide range of stakeholders and processes in the supply chain visible. This brings an opportunity for better managing the SDI.

5. Conclusion

One of the objectives of the SDI is to make data available and accessible to all levels of society since everyone now virtually uses spatial information. To improve data accessibility, data must be affordable to industry, government and the public. This requires the SDI to be as cost effective as possible. Cost effectiveness of the SDI can be analyzed and improved through studying the SDI supply chain. The objective of this study was to define them in supply-chain terminology. We demonstrated the chain of events in the land development process that leads to the production of cadastral data in the CSG's data store. The progression of actual events does not follow such a simple a linear sequence but is actually a more complex one. We plan to expand that flow to model it with a reputable supply chain model.

In the past, supply chains were studied and analyzed for a particular single manufacturing process and for a particular product or organization. This has changed, as there is increased attention in studying the supply chain as a whole (Beamon 1998, Min and Zhou, 2002). In other words, rather than studying a particular geographic information system in an organization or a particular process within the GIS network, it is also necessary to look at the integrated system as a whole. This can be done using the Supply Chain Operations Reference model (SCOR) illustrated in (Gardner and Cooper, 2003; Schmitz, 2008), since it makes it possible to map and analyze the supply chain as a whole. In future work, we plan to apply the SCOR model to analyze processes in the SDI supply chain network as a whole rather than only analyzing single discrete processes.

References

Akinyemi F. O., (2011), Evaluating Access to Spatial Data Information in Rwanda. *URISA Journal,* 23(2): 30-47.

ANZLIC (2010), *Economic Assessment of Spatial Data Pricing and Access – Stage 1 Report: Principles, Issues and Alternative Models*, The Spatial Information Council.

Beamon, B.M., (1998), Supply Chain Design and Analysis: Models and Methods. *International Journal of Production Economics,* 55: 281-294.

Campbell, R. and Kula, O., (2006), A Value Chain Approach to Economic Development, ACDI VOCA: Expanding Opportunities Worldwide.

Campagna, M., (2006), GIS for Sustainable Development. CRC Press, Boca Raton, FL, USA. ISBN 978-0-8493-3051-3.

Christopher, M., (2011), Logistics and Supply Chain Management – 4th Edition, England: Prentice Hall.

Crompvoets, J., de Man, E. and Macharis, C., (2010), Value of Spatial Data: Networked Performance beyond Economic Rhetoric. *International Journal of Spatial Data Infrastructures Research*, 5: 96-119.

CSI (2012), Stakeholder Survey on Defining the Criteria and Identifying Core Geospatial Datasets and Data Custodians in South Africa, Committee for Spatial Information: South Africa.

Frazelle, E.H., (2002), Supply Chain Strategy – The Logistics of Supply Chain Management, New York: McGraw-Hill Trade.

Gardner, J.T. and Cooper, M.C., (2003), Strategic Supply Chain Mapping Approaches. *Journal of Business Logistics*, 24(2): 37-64.

Genovese, E., Cotteret, G., Roche, S., Caron C. and Fleick, R., (2008), Evaluating the Socio-Economic Impact of Geographic Information: A Classification of the Literature. *Article under Review for the International Journal of Spatial Data Infrastructures Research*.

Genovese, E., Roche, S. and Caron, C., (2009), The Value Chain Approach to Evaluate the Economic Impact of Geographic Information: Towards a New Visual Tool.

Grus, L., Crompvoets, J., and Bregt, A.K., (2008), Theoretical Introduction to the Multi-view Framework to Assess SDIs, in Crompvoets J, Rajabifard, A., van Loenen, B. and Fernández, T.D. (Eds.), A Multi-View Framework to Assess SDIs: Wageningen University.

Grus, L., Crompvoets, J. and Bregt, A.K., (2010), Spatial Data Infrastructures as Complex Adaptive Systems. *International Journal of Geographical Information Science,* 24(3): 439-463.

Hendriks, P.H.J., Dessers, E. and van Hootegem, G., (2012), Reconsidering the Definition of a Spatial Data Infrastructure. *International Journal of Geographical Information Science*, 26(8): 1479-1494.

Hjelmager, J., Moellering, H., Delgado, T., Cooper, A.K., Rajabifard, A., Rapant, P., Danko, D., Huet, M., Laurent, D., Aalders, H.J.G.L., Iwaniak, A., Abad, P.,

Düren, U. and Martynenko, A., (2008), An initial Formal Model for Spatial Data Infrastructures. *International Journal of Geographical Information Science*, 22(11): 1295-1309.

Joburg (2011), Corporate Geo-Informatics – Progress Report, CoJ Printing Department.

Krek, A. and Frank, A. U., (2000), The Production of Geographic Information – The Value Tree. *Journal for Spatial Information and Decision Making*, 13(3): 10-12.

Kurwakumire, E., (2013), Evaluating GIS – A Case of Uganda Public Sector, In Otukei R, Gidudu A and Musinguzi M., (Eds.), Proceedings of the Advances in Geomatics Research Conference, Makerere University, Kampala, Uganda, ISBN:978-9970-475-00-1.

Longhorn, R. and Blakemore, (2008), Geographic Information – Value, Pricing, Production and Consumption, Boca Raton: CRC Press.

Makanga, P. and Smit, J., (2010), A Review of the Status of Spatial Data Infrastructure Implementation in Africa, *Research Article – SACJ*, 45: 357-365.

Min, H. and Zhou, G., (2002), Supply Chain Modeling: Past, Present and Future. *Computers & Industrial Engineering*, 43(1): 231-249.

NAS (2009), Ensuring Integrity, Accessibility and Stewardship of Research Data in the Digital Age. Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, National Academy of Sciences.

NAS (2004), The Geographic Data Market – Licensing Geographic Data and Services, National Academy of Sciences.

Rajabifard, A., Feeney, M.E.F., and Williamson, I.P., (2002), Directions for the Future of SDI Development. *International Journal of Applied Earth Observation and Geoinformation*, 4: 11-22.

Schmitz, P., (2007), The use of Supply Chains and Supply Chain Management to Improve the Efficiency and Effectiveness of GIS units, Geography Department: University of Johannesburg.

Schmitz, P., (2008), Using Supply Chain Management to Enable GIS Units to Improve Their Response to Their Customer's Needs, Proceedings of the Academic Track of the 2008 Free and Open Source Software for Geospatial Conference, Capetown, South Africa.

Simchi-Levi, D., Simchi-Levi, E. and Kaminsky, P., (1999), Designing and Managing the Supply Chain – Concepts, Strategies, and Cases, United States: McGraw-Hill.

Sinvula, K.M., Coetzee, S., Cooper, A.K., Nangolo, E., Owusu-Banahene, W., Rautenbach, V. and Hipondoka, M., (2013), A Contextual ICA Stakeholder Model Approach for the Namibian Spatial Data Infrastructure (NamSDI). 26th International Cartographic Conference (ICC), August 25-30, Dresden, Germany.

Steudler, D., Rajabifard, A., and Williamson, I.P., (2008), Evaluation and Performance Indicators to Assess Spatial Data Infrastructure Initiatives, in Crompvoets J., Rajabifard, A., van Loenen, B. and Fernández, T.D., A Multi-View Framework to Assess SDIs: Wageningen University.

Van Loenen, B. and van Rij, E., (2008), Assessment of Spatial Data Infrastructures From an Organisational Perspective, in Crompvoets J., Rajabifard, A., van Loenen, B., and Fernández, T.D., A Multi-View Framework to Assess SDIs: Wageningen University.

Welle Donker, F., (2010), Public Sector Information Access Policies in Europe. In Fitzgerald B (Ed.), Access to Public Sector Information: Law, Technology & Policy. Volume 1 (pp. 253-279). Sydney: Sydney University Press.

Welle Donker, F., (2009), Public Sector Geo Web Services: Which Business Model Will Pay for a Free Lunch?, in SDI Convergence, GSDI, http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/143.pdf.

Welle Donker, F., van Loenen, B., and Zevenbergen, J., (2010), Geo Shared licenses: A base for better access to public sector geoinformation for value added resellers in Europe. *Environment and Planning B: Planning and Design*, 37: 326-343.

CHAPTER 3

Design and Development of a Spatial Metadata Automation Framework Applied in Australia

Hamed Olfat[1&2], Abbas Rajabifard[1], Mohsen Kalantari[1] and Chris Pettit[1]

[1] Centre for SDIs and Land Administration, Department of Infrastructure Engineering,
The University of Melbourne
olfath@unimelb.edu.au; abbas.r@unimelb.edu.au; saeidks@unimelb.edu.au;
cpettit@unimelb.edu.au

[2] Land Victoria, Department of Transport, Planning and Local Infrastructure
hamed.olfat@dtpli.vic.gov.au

Abstract

Current approaches struggle to effectively and efficiently manage metadata creation, updating and improvement for the ever-growing amount of spatial data created and exchanged between people or organizations within the Spatial Data Infrastructures (SDIs) and data sharing platforms. In order to overcome the main challenges regarding spatial metadata management, this chapter presents the outcomes of a research project undertaken by the authors at the University of Melbourne. The chapter first explores the results of a case study investigation in the context of Australia to identify the spatial metadata management and automation requirements. Then, it reviews the results of assessing a number of metadata management tools, which are commonly used within the geospatial community, against a set of criteria developed for this research. The chapter then investigates the design and development of a framework and associated approaches and tools to facilitate and automate spatial metadata creation, updating (in real time with dataset modification), and enrichment (through the end users' interactions). This framework took advantage of GML and Web 2.0 technologies. Finally, a metadata system designed and implemented for the Australian Urban Research Infrastructure Network (AURIN) portal based on the outcomes of the research project is reviewed in the chapter.

KEYWORDS: spatial, metadata, automation, GML, Web 2.0, AURIN

1. Introduction

Spatial information is necessary to make sound decisions at the local, regional and global levels (Nebert 2004). Therefore, the amount of spatial datasets being created and exchanged between organizations or people is increasing considerably. According to a released study by Daratech (2011) for the period of 2004–2010, the overall growth of geospatial industry has increased by 11% in the areas of data, software and services. The report highlighted that the spatial data is the fastest growing segment of the geospatial industry and is definitely becoming a major contributor to the overall growth of the industry.

As more spatial data is produced, it becomes more important to manage and locate such resources (Göbel and Lutze, 1998). The role spatial metadata plays in the management and location of these resources has been widely acknowledged (Tsou 2002, Limbach *et al.,* 2004). Batcheller (2008) also agrees with this view and states that metadata is often employed by institutions to organize, maintain and document their spatial resources internally, and may also provide a vehicle for exposing marketable data assets externally when contributed to online geospatial exchange initiatives.

However, the current approaches struggle to effectively manage metadata creation, updates, and improvement for an ever-growing amount of data created and shared in the Spatial Data Infrastructures (SDIs) and data-sharing platforms. Among the available approaches, the manual entry, enrich and update approach has been considered monotonous, time-consuming, and a labor-intensive task (West and Hess 2002, Guptill 1999). Also, existing semi-automatic metadata approaches mainly concentrate on specific dataset formats to extract a limited number of metadata values (e.g. bounding box). Moreover, metadata is commonly collected and created in a separate process from the spatial data lifecycle, which requires the metadata author or responsible party to put extra effort into gathering necessary data for metadata creation and updating. In addition, dataset creation and editing are detached from metadata creation and editing procedures, necessitating diligent updating practices involving at a minimum, two separate applications (Rajabifard *et al.,* 2009). Metadata and related spatial data are often stored and maintained separately using a detached data model that fails to provide automatic and simultaneous metadata updating when a dataset is modified. In addition to these challenges, Cooper *et al.,* (2011) discuss that users are not involved in the development of standards, such as assessing quality or documenting metadata. Kalantari *et al.,* (2010) also argue that the users are disconnected from the spatial metadata creation and improvement process.

In order to address these challenges, a research project entitled 'spatial metadata automation' was undertaken by the authors in four phases. The first (Conceptual) phase investigated the requirements of spatial metadata automation. In this phase, to establish the theoretical background of the research, along with an extensive literature review, a case study was undertaken in Australia in order to identify the current status of spatial metadata management and the requirements for the spatial metadata automation. Also, a number of spatial metadata management tools were selected and

assessed against a set of criteria developed for this research. Finally, the results achieved from the first phase were integrated and the main challenges regarding the spatial metadata management and automation were determined. In the second (Design) phase, a spatial metadata automation framework was designed and developed to overcome the identified main challenges. In the third (Implementation) phase, two prototype systems were implemented to prove the conceptual design of the framework. In the final (Evaluation) phase, a set of criteria was developed for the assessment of the prototype systems. In this phase, two questionnaires were designed and distributed among the organizations that participated in the Australian case study as well as other interested parties. The results of this survey are being analyzed to identify the areas which need improvement. The outcomes of this research project were then applied to design and develop a spatial metadata tool for the Australian Urban Research Infrastructure Network (AURIN) platform. In this chapter, we will provide a summary of each of the phases of the research and discuss the development of the spatial metadata tool in the context of the AURIN platform application.

2. Australian Case Study

The case study strategy was selected for investigating the current status of spatial metadata management and automation requirements within the geospatial community. Australia was selected as the case study area for undertaking this investigation as the research team had ready access to a number of partner organizations and there has been a significant effort in spatial metadata standards and systems implementation in Australia.

Within the context of the Australian case study area, twelve organizations were identified to participate in the survey based on a snowball sampling method. These organizations were the Office of Spatial Data Policy, Victorian Departments of Primary Industries and Sustainability and Environment, ACT Planning and Land Authority, PSMA Australia Limited, VicRoads, Bureau of Meteorology, Tasmanian Department of Primary Industries, Parks, Water and Environment, Australian Hydrographic Service, South Australian Department for Transport, Energy and Infrastructure, Western Australian Land Information System (WALIS), and Sinclair Knight Merz Pty Ltd.

In order to undertake the survey, a Web-based questionnaire was designed and distributed among the participating organizations. The survey identified the main issues and challenges associated with managing spatial metadata and collected the priority requirements for spatial metadata automation for the participating organizations outlined above. The structure of the questionnaire and the results of the survey are discussed in detail by Olfat *et al.,* (2010a). To complement the requirements analysis, a number of commonly used spatial metadata management tools were also examined to understand current functionalities and automation capabilities and shortcoming of these tools. The following section reviews the results of that examination.

3. Spatial Metadata Tools Assessment

Following the main challenges and requirements of spatial metadata management identified during the case study; in order to assess the current status of metadata tools, a set of criteria was suggested. These criteria are outlined in Table 1.

Criteria	Description
Support for the integration of metadata creation with spatial data lifecycle	Whether the tool provides the organizations with the facilities to generate metadata values parallel to their spatial data lifecycle. This integration could result in reducing the burden of metadata creation for metadata authors through associating the generation of each metadata element to its relevant responsible party. Also, the integration could have the potential to overcome the problem of missing or incomplete metadata through recognizing the stage to generate and update metadata within the data lifecycle.
Support for the integrated data model	Whether the tool provides an integrated data model for storing spatial datasets and metadata. This data model should allow datasets and metadata to be coupled in a common environment, which could result in managing and maintaining them together.
Support for automatic metadata creation	Whether the tool supports the automatic generation of spatial metadata values using different sources such as dataset file and pre-defined metadata. This could result in saving required time and resources for creating metadata and also increasing the quality of metadata by reducing the human error.
Support for automatic metadata updating when dataset changes	Whether the tool automatically synchronizes the metadata with any changes to the dataset in real time. This could result in the spatial metadata always being up-to-date.
Support for interaction with end users to improve the content of metadata	Whether the tool engages end users to improve the content of metadata during the data discovery process. For instance, the interaction might include the functionalities for tagging datasets with new search words or commenting on the datasets. Interaction with end users could result in connecting the end users to metadata creation and maintenance process, facilitating the discovery process by involving end users' knowledge of datasets, and making the metadata management tool more user-friendly. The end user here means people seeking spatial datasets and not the spatial data cataloguers.

Table 1. Set of criteria proposed for examining the selected spatial metadata management tools

An investigation was then undertaken to select the broadly used metadata management tools in the geospatial community that could be adapted to suit the spatial metadata needs of the case study area – Australia. The results from the case study investigations were also used for the selection purpose.

According to Rajabifard *et al*., (2007), CatMDEdit, and GeoNetwork have been recognized as the international metadata management tools that could be adapted to

suit the needs of ANZLIC, as the peak intergovernmental organization for the collection, management and use of spatial information in Australia and New Zealand. Also, according to findings of case study investigations, ANZMet Lite, GeoNetwork, BlueNetMEST, and ESRI ArcCatalog were frequently used by the participating organizations in Australia (Olfat *et al*., 2010a).

Thus, all these tools were assessed against the criteria proposed in Table 1. In addition to these tools, GeoNode and European Open Source Metadata Editor (EUOSME) were also included in the assessment process. Moreover, another tool, which has been recently developed in Australia namely 'xMet Client', was included in the evaluation.

As a result of this investigation, the spatial metadata management tools selected to be reviewed and examined against the criteria were categorized in two main groups including international and Australian. Table 2 illustrates the summary of results for tools assessment.

Spatial metadata management tool		**Criteria**				
		Support for integration with spatial data lifecycle	**Support for integrated data model**	**Support for automatic metadata creation**	**Support for automatic metadata updating**	**Support for interaction with end users**
International	GeoNetwork opensource	NO	NO	YES – Only when metadata are harvested.	YES – Only when metadata are harvested; but not in real time.	NO
	GeoNode	NO	NO – Only metadata and spatial data are associated to each other.	YES – Only for a limited number of elements (e.g. date/time, Bounding Box, distribution URL, and contact information).	NO	NO
	ESRI ArcCatalog	NO	NO – Only metadata and spatial data are stored in the same place.	YES – Depending on the input dataset formats for the elements outlined by ESRI (2002).	YES – Depending on the input dataset formats for the elements outlined by ESRI (2002); but not in real time.	N/A

	CatMDEdit	NO	NO	YES – Depending on the input dataset formats for the elements outlined by CatMDEdit (2011).	NO	N/A
	EUOSME	NO	NO	NO	NO	N/A
Australian	BlueNetMEST	NO	NO	YES – Only when metadata are harvested.	YES – Only when metadata are harvested; but not in real time.	NO
	ANZMet Lite	NO	NO – Only 'linked' metadata is stored in the same directory which the dataset is stored.	YES – Depending on the input dataset formats for the elements outlined by OSDM (2009).	NO	N/A
	xMet Client	NO	NO	NO	NO	N/A

Table 2. Summary of results of examining selected metadata management tools against the criteria

The integration of results achieved during the case study and tools assessment resulted in identifying the main challenges which needed to be addressed in the research. These main challenges are described in next section.

4. Summary of Main Challenges

The identified challenges can be classified in five categories.

4.1 Relationship between Metadata Management and the Spatial Data Lifecycle

Metadata describes different aspects of the dataset such as identification, quality, citation, extent, constraints, etc (ISO 19115: 2003). Therefore, ideally metadata should be part of a spatial dataset and its values should be generated and updated with any change to the dataset from the very first stages of the data lifecycle (Olfat *et al.,* 2012b). Producing metadata afterwards is difficult and may be a laborious task (Taussi 2007). However, the results of the case study and tools assessment reveal that metadata generation is commonly undertaken after the dataset is fully created or is ready to be published over the Web at one point of time, which is not an incessant

practice parallel to the data lifecycle. Collecting metadata later requires considerable effort and not all the information might be available (Timpf *et al.,* 1996) and the metadata gathered in this way is often missing or incomplete (Rajabifard *et al.,* 2009).

4.2 Use of an Integrated Metadata Data Model

The conceptual phase showed that the current metadata generation/updating approach is rooted in a detached data model. In a 'detached data model' the spatial data and its associated metadata are stored separately in different files or databases which make them either without a relationship with one another or to have only a common identifier. In contrast, in an 'integrated data model' spatial data and metadata can be mapped to and stored in a middleware, so that with any change in the data the metadata can be updated at the same time (Kalantari *et al.,* 2009). Moreover, it was deduced that the current approach is entirely dependent on the knowledge of the metadata author or responsible party about the dataset (Olfat *et al.,* 2012a).

As a result, designing and implementing an integrated metadata data model that can be used for automating the metadata updating process at the same time as the dataset's modification will benefit organizations.

4.3 Real-time Spatial Data and Metadata Updating

The results of the survey ran in the context of Australia and showed that a large portion of spatial data custodians and creators still update metadata through a separate activity from the dataset modification (Olfat *et al*., 2010a). It was also found that separate teams have the responsibility to update dataset and metadata for the same organization. This results in a delay between dataset and metadata updating time and therefore prevents the metadata from being always up-to-date, reliable, and precise. Following this current approach, the organizations need extra resources in terms of budget and time to undertake further effort to update metadata after any change to the dataset.

In addition, the results of spatial metadata tools assessment showed that these tools lack the support for real-time spatial data and metadata updating. Within the assessed tools, the ESRI ArcCatalog, which has also been commonly used by different researchers (Batcheller 2008, Batcheller *et al*., 2007, Westbrooks 2004), synchronizes metadata with the latest status of spatial data. However, the synchronization process requires human intervention to run (ESRI 2010). Therefore, metadata and spatial data updating would not be real-time in nature.

4.4 Dependency of Metadata Automation Methods on Dataset Format

As a result of systematically reviewing the metadata automation research and development activities and selected spatial metadata tools, it can be surmised that the

existing automation tools are highly restricted to dataset formats to extract metadata values. For instance, CatMDEdit automates metadata generation only for Shapefile, DGN, ECW, FICC, GeoTIFF, GIF/GFW, JPG/JGW, and PNG/PGW formats (CatMDEdit 2011), or GeoNode automatically generates a few metadata values for the Shapefiles. Therefore, an agnostic dataset format automatic approach to create and update metadata will play an important role to address this issue.

4.5 Interaction with End Users for Metadata Creation and Improvement

The results of assessing the selected spatial metadata tools, along with the case study investigations in the context of Australia, indicated that the current tools are not sufficiently user-friendly. The end users are also disconnected from the spatial metadata creation and improvement process. These tools need more interaction with the users to improve the content of metadata; especially 'keyword' metadata element (Kalantari *et al*., 2010), which is the main gateway for discovering and finding datasets over the Web.

In order to address the above challenges, a spatial metadata automation framework and associated tools were designed and developed which are presented in the next section.

5. Spatial Metadata Automation Framework

As illustrated in Figure 1, the core of framework contains the 'spatial data lifecycle'. The framework through the 'Lifecycle-centric Spatial Metadata Creation' approach aims to integrate the metadata creation with the steps involved in the spatial data lifecycle. Also, the framework aims to address the challenge of lack of a real-time dataset and metadata updating through the 'Automatic Spatial Metadata Updating' approach, which would be dataset format agnostic and integrated with the spatial data lifecycle. Also, in order to address the current challenge of using a detached data model for spatial data and metadata storage the framework develops an integrated data model in which the spatial dataset and its related metadata can be managed and updated together. Finally, the developed framework focuses on engaging the end users to improve the content of metadata to address the challenge of end users' disconnection from the metadata creation and improvement process. This would be undertaken through the 'Automatic Spatial Metadata Enrichment' approach, which is also integrated with the spatial data lifecycle.

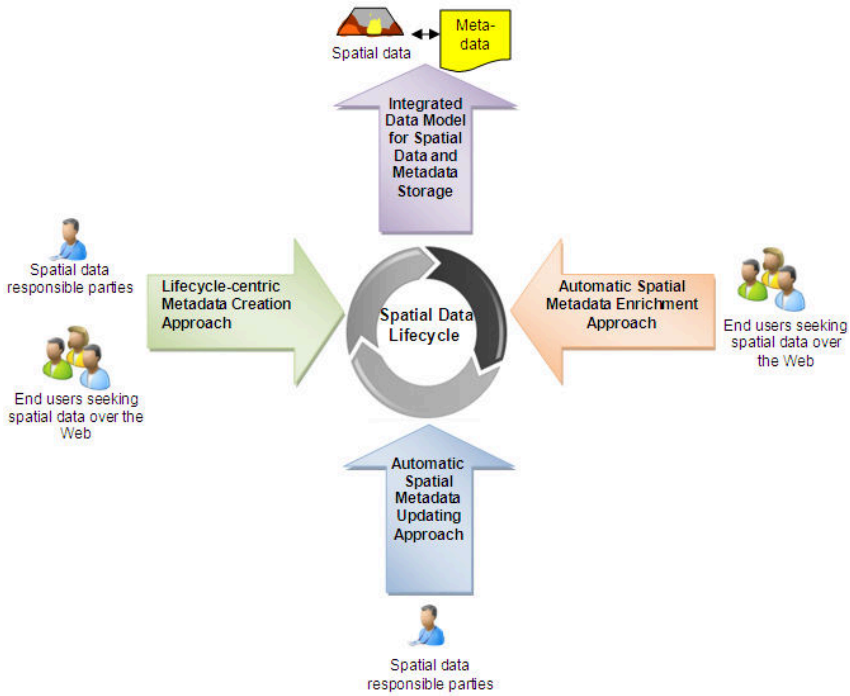

Figure 1. A framework to address the identified challenges regarding spatial metadata management and automation

The components of the framework are described in more details in following sections.

5.1 Lifecycle-centric Spatial Metadata Creation Approach

The 'Lifecycle-centric Spatial Metadata Creation' approach, as illustrated in Figure 2, aims to integrate the metadata creation with the steps involved in the spatial data lifecycle. In this regard, a generic spatial data lifecycle was developed based on the Australian Government Information Interoperability Framework (AGIMO 2006) in the 'Planning and Policy Making', 'Data Collection', 'Spatial Dataset Creation', 'Storage', 'Publication', 'Discovery and Access', 'Utilization', and 'Maintenance' steps.

The elements recommended by the ISO 19115: 2003 were then reviewed systematically and mapped against the steps of the generic spatial data lifecycle. According to the results of this investigation discussed by Olfat *et al*., (2012b), the highest number of metadata elements should be created within the spatial dataset creation step. Planning and policy making, dataset maintenance, publication, data collection, dataset storage, utilization, and discovery and access are respectively the next steps with the highest number of elements.

As illustrated in Figure 2, the responsible party for each step of the spatial data lifecycle would be in charge of creating (and maintaining) related metadata values within those steps. The case study investigations showed that the responsible parties for creating and maintaining spatial data can be categorized into two main groups: an internal team within the organization and an external team. Also, based on the definition of 'discovery and access' and 'utilization' steps of the lifecycle discussed by Olfat *et al*., (2012b) the end users are mainly engaged in creating (and improving) the content of metadata within these two steps.

As a result of the lifecycle-centric approach, the metadata will be completed over time in conjunction with the spatial data lifecycle and therefore, it is more likely to be accurate and up-to-date. The proposed approach brings forth some advantages for the organizations. This approach will support the generation and updating of a wide range of the ISO metadata elements. It also has the potential to overcome the problem of missing or incomplete metadata through recognizing the specific step to generate and update metadata within the data lifecycle. Moreover, it most likely reduces the burden of metadata creation for metadata authors by involving the spatial data responsible parties and interacting with the end users in creating and updating metadata values. However, in order for the lifecycle-centric approach to work properly a metadata entry/edit tool needs to be designed and developed to provide a combination of available manual/semi-automatic/automatic metadata creation methods and an appropriate level of access for the responsible parties involved in the spatial data lifecycle.

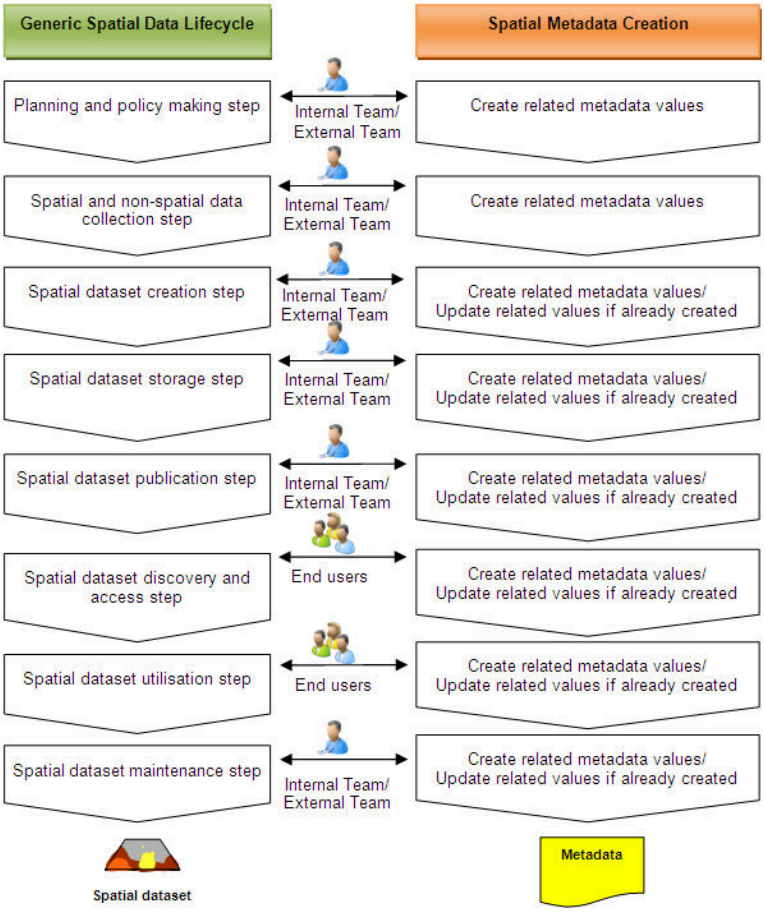

Figure 2. Life-cycle centric spatial metadata creation approach

According to Figure 1, an automatic approach for updating metadata in conjunction with the dataset modification and an integrated data model for storing metadata and dataset are other components of the framework developed for addressing the spatial metadata challenges. The next section explores these two components.

5.2 Automatic Spatial Metadata Updating (Synchronization) Approach

Real-time spatial metadata and dataset updating plays a significant role in accessing the most up-to-date and precise metadata in any sharing platform. The metadata synchronization approach, as an automatic process by which properties of a spatial dataset are read from both the back end (where dataset is stored) and the front end (where the modification environment is up and running) and written into its spatial metadata at the same time as any modification of the dataset (Olfat *et al.,* 2010b; Kalantari *et al.,* 2010), is designed to address this need.

According to the aim of metadata synchronization approach, the prerequisite for taking this approach into account would be designing and building an integrated data model for storing metadata and the dataset related to each other. Accordingly, a Geography Markup Language (GML) application schema was developed to support the integrated data model (Olfat *et al.,* 2013). Through this data model, each dataset would be related to its metadata record. Having the relationship between these two sources and accommodating dataset geometries, attributes and metadata values into a middleware would result in a comprehensive dataset, which can also be exchanged over the Web between different spatial systems as well as end users. Comprehensive datasets are datasets that are associated by three fundamental components: geometries (and topologies), attributes and metadata.

By transferring this comprehensive dataset to a user interface through the middleware the users (spatial data responsible parties) would be able to represent and edit dataset and metadata concurrently. While modifying a dataset, the responsible parties should be able to see the modification reflection on corresponding metadata values simultaneously and automatically. Some of the metadata values affected by the dataset modification should be updated at the front end (e.g. date of revision and lineage) and the others should be updated at the back end (e.g. bounding box) via synchronization scripts. Those elements that are updated at the front end would be transferred to the back end (metadata table) through the middleware and those updated at the back end would be directly replaced on the metadata table in the database.

By having the synchronization approach in place, after any dataset modification the new values for geometry, attributes and metadata would be transferred to and stored at the back end and are shown at the same time on the user interface. The conceptual design for metadata synchronization approach is illustrated in Figure 3 (modified after Olfat *et al.,* 2013).

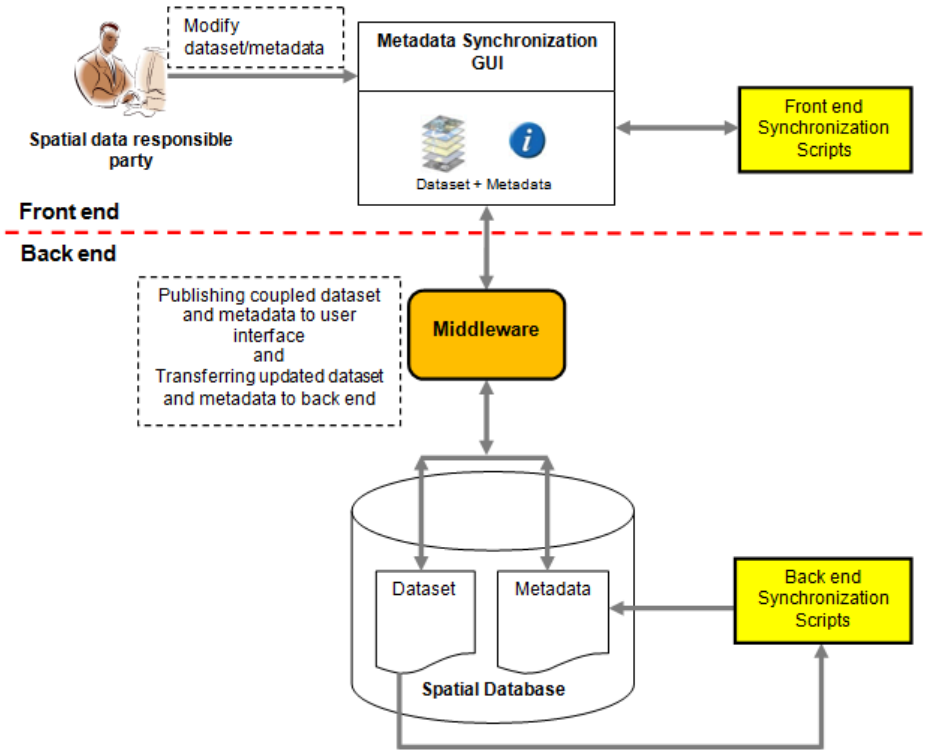

Figure 3. Conceptual design for metadata synchronization approach (modified after Olfat *et al.,* 2013), proposed to be adopted in 'Dataset Modification' step within the spatial data lifecycle

In order to prove the above conceptual design, a prototype system was implemented within the GeoNetwork open-source environment using the three-layer architecture of storage, service, and application layers, as illustrated in Figure 4. In the storage layer two databases were built; one for storing the spatial dataset together with metadata and another one for storing the corresponding metadata in the GeoNetwork. In the service layer, the Web servers for supporting the required services (including WFS-Transactional and CSW) were employed. Also, a new Web service namely 'SYNC' was developed in the service layer to synchronize the metadata catalogs stored in both databases in the storage layer. In the application layer, an interface was developed and integrated with the GeoNetwork interface to display the spatial data and metadata coming in GML format (as the output of integrated data model developed for bundling dataset and metadata). This interface also provided the end users with the facility to modify the vector dataset and see the reflection on a subset of ISO 19115: 2003 metadata elements (e.g. date of revision, lineage statement, and bounding box) automatically and simultaneously. Through the 'SYNC' Web service the corresponding metadata catalog stored in the GeoNetwork database was also updated with any change to the dataset.

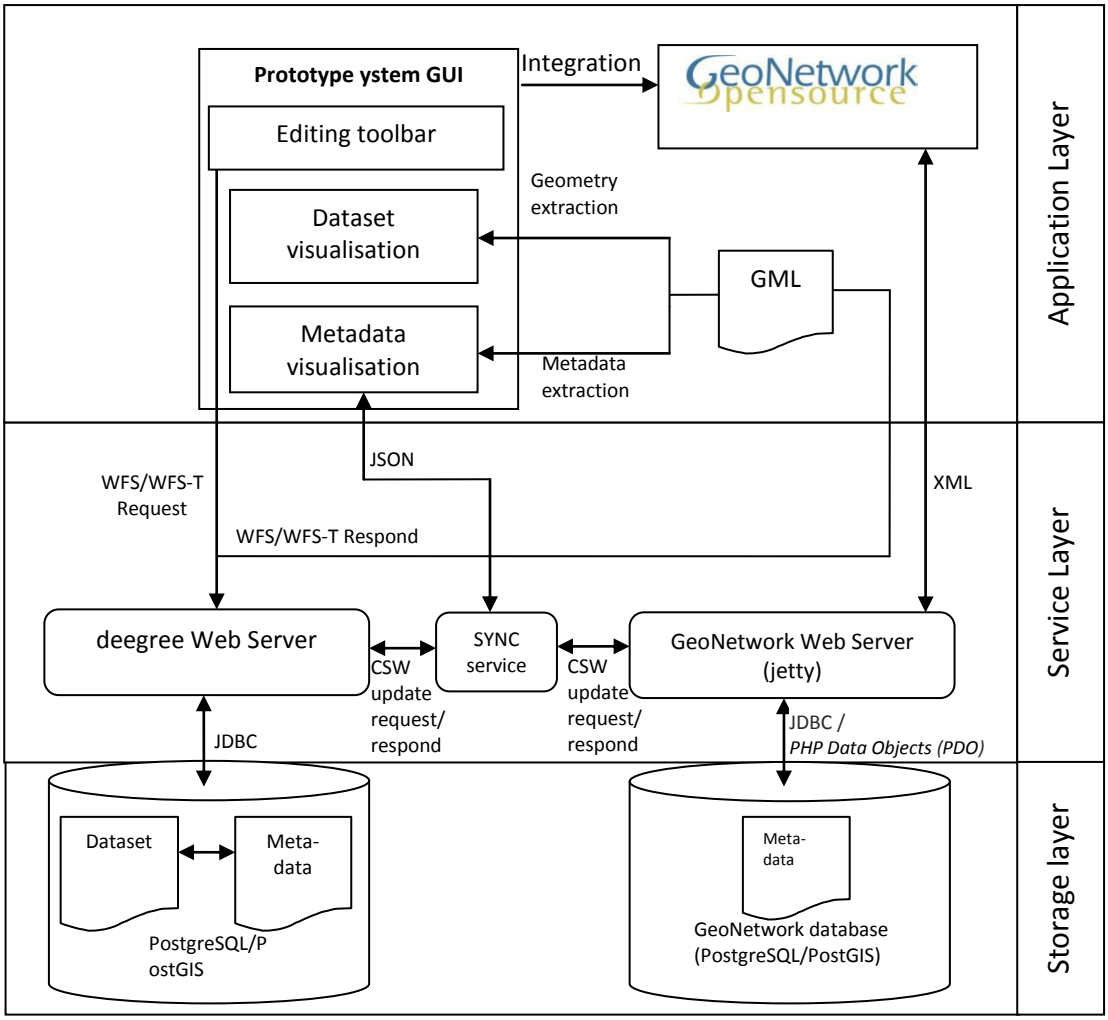

Figure 4. Architecture for automatic spatial metadata updating prototype system

In order to implement the system the open source technologies (e.g. GeoNetwork opensource catalog, deegree WFS-T and CSW server, PostgreSQL, PostGIS, OpenLayers and GeoExt) were used. Figure 5 illustrates the overview of the prototype system Graphical User Interface (GUI).

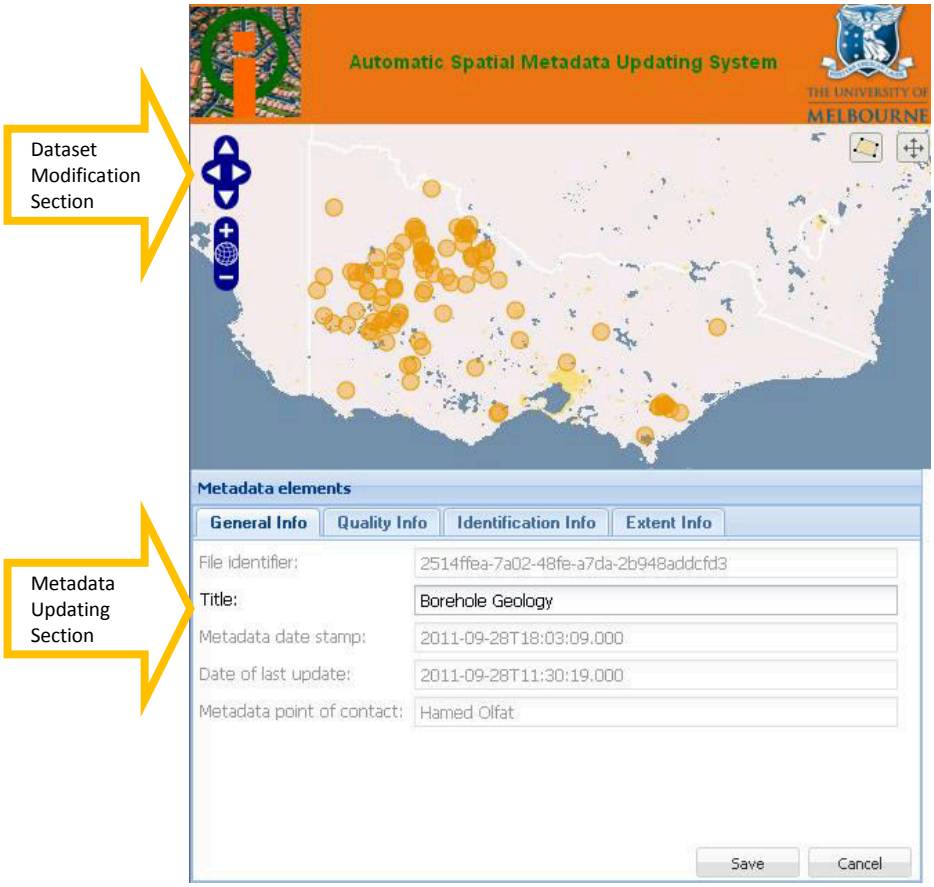

Figure 5. Automatic spatial metadata updating (synchronization) prototype system GUI

The functionalities implemented for the automatic spatial metadata updating approach would most likely provide the geospatial community with a variety of benefits including a reduced burden of manual metadata updating after dataset modification, facilitation of the interoperability by publishing the datasets in GML and regardless of any specific dataset format, and enablement of data responsible parties to publish and share datasets along with attributes and metadata in a single document. This approach also helps to avoid missing, incomplete, out-of-date and unreliable metadata. Moreover, having the metadata synchronization approach in place could give a peace of mind to data responsible parties, due to the metadata always being current with dataset changes. The synchronization could also provide a better discovery service to users seeking spatial datasets over the Web by providing them with the most recent version of metadata.

The next section presents the last component of the spatial metadata automation framework, namely 'Automatic Spatial Metadata Enrichment' and its associated prototype system.

5.3 Automatic Spatial Metadata Enrichment Approach

In order to design the automatic spatial metadata enrichment approach, there was a need to define the metadata element(s) of which their value(s) could be improved by the end users seeking spatial data over the Web. According to the results of mapping the ISO 19115: 2003 metadata elements against the generic spatial data lifecycle steps, the 'descriptive keyword' was the only metadata element in which its value could be improved by the end users during the 'dataset discovery and access' step of the spatial data lifecycle. The descriptive keyword element is one of the mandatory elements recommended by the ISO 19115: 2003 standard that should be embedded into each spatial metadata record and is defined as 'commonly used word(s) or formalized word(s) or phrase(s) used to describe the subject' (ISO 19115: 2003).

The spatial data discovery systems usually support making a variety of queries via basic and advanced search modes on the spatial metadata records to retrieve the characteristics of the most appropriate datasets for the end users. Accordingly, identifying the appropriate keywords to describe the spatial datasets is fundamental within any sharing platform, but has become increasingly problematic (Chi 2009). The appropriate keyword for any spatial dataset means the keyword that is consistent with the content of the dataset and can reveal its essence and applications. In addition, an appropriate keyword should address the probable queries made by users from diverse categories. Moreover, a keyword should have a popular meaning that most of the users and data responsible party agree on (Kalantari *et al.,* 2010).

Currently, the keyword metadata element is created by the metadata responsible parties in two ways: 1) using a library of search words during metadata creation (e.g. the ANZLIC search words library within the ANZMET Lite and xMet Client tools), and 2) using responsible parties' own opinions about the datasets. Although, using the first method will standardize the process of keyword allocation among the metadata responsible parties, it can restrict the end users to select a keyword from a defined range in which they may or may not be familiar with. Also, the second approach to keyword creation requires the end users to have the same knowledge and insight as the responsible parties have about the whole library keywords and the areas they cover.

To improve the current process of keyword creation, Kalantari *et al.,* (2010) designed two complementary models namely 'indirect' and 'direct' rooted in the tagging and folksonomy features of Web 2.0. The indirect model is designed to recognize and tag the most popular search words for describing datasets through monitoring the end users' behavior during the data discovery process without their understanding. This model contains three stages as illustrated in Figure 6. However, the direct model enables the end users to interact with the metadata records and rate (to agree or disagree with) the tagged search words and also to directly add value to the keyword metadata element. Through the direct model, the end users are also able to comment on the datasets and create a new additional metadata.

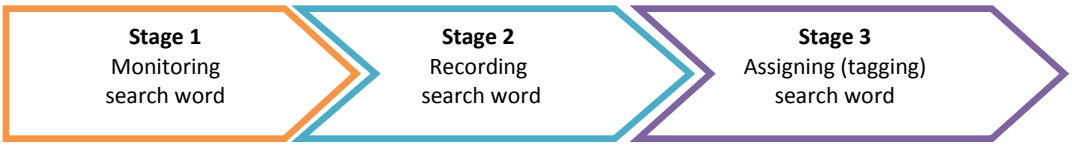

Figure 6. Stages of the indirect model for spatial metadata enrichment

The tagged search words resulted from both direct and indirect models will be also visualized in a tag cloud. Using the tag cloud which includes the most popular keywords describing the datasets within a spatial data discovery system, will enhance and facilitate the spatial dataset discovery and retrieval process. In the tag cloud, the end users will be able to retrieve the metadata associated to any tag through clicking on the tag that acts similarly to a hyperlink. In addition, the end users can have the possibility to visualize the tags which are assigned to the same dataset by hovering the mouse over a search word in the cloud. Figure 7 illustrates this concept.

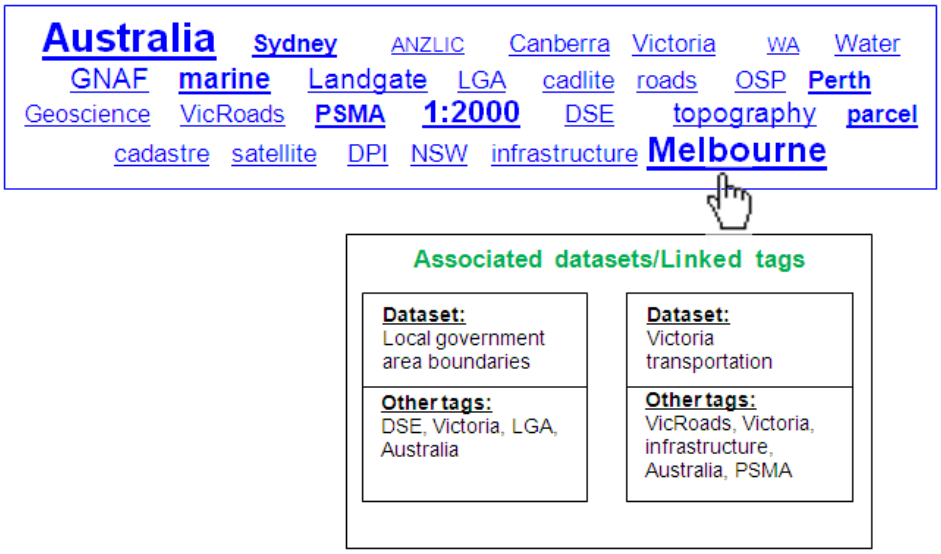

Figure 7. Visualizing a search word's associated datasets and linked tags

One of the critical considerations for developing the automatic spatial metadata enrichment approach would be the involvement of the spatial data responsible parties to assess the recorded search words and remove the noise and spam before they are assigned to the metadata records as new 'metadata keyword' values.

In order to prove the concept of the automatic spatial metadata enrichment approach, two add-ons were implemented within GeoNetwork in a way that could be simply installed by other users on a working GeoNetwork platform. These add-ons are here described.

5.3.1 Agree/Disagree/Tag Cloud Add-on

This add-on automatically observes the end users' interaction with the spatial data discovery process within the GeoNetwork and collects the end users' feedback on metadata records. It monitors every interaction of end users including exploration of metadata details, spending time on reading metadata, and downloading the actual spatial data. It also asks the end users to submit new search words, and agree or disagree with the existing tagged search words. Therefore, this add-on provides the end users with a way to enrich metadata records with tagged search words (implicitly or explicitly), so that subsequent users benefit from this enrichment in their own searches. The tagged search words are represented in a cloud, which is a graphically weighted list of search words (tags). Once the relevant search words are recorded, they should be assigned to their related metadata records as a new 'keyword' metadata value. The assignment of recorded search words depends on the specific threshold considered by the data catalog administrator or metadata/data responsible party.

5.3.2 Suggestion List Add-on

This add-on was designed and implemented within the GeoNetwork to provide the end user with a suggestion list based on previously searched terms. This happens while typing a search word in the data catalog search box. Using this add-on, all subsequent searches benefit from previous searches. This facility most likely provides a user-generated context for metadata improvement and automation.

Through the add-ons developed within the GeoNetwork the end users are able to interact with the data discovery system for creating and improving the content of 'keyword' metadata element. They can also share their knowledge about datasets by agreeing or disagreeing with the relevance of the existing tagged search words or adding new search words to the datasets. Also, the tag cloud potentially provides the end users with the capabilities to visit the most popular search words, gain ideas regarding the available data, and access the datasets more quickly and simply.

In the last phase of the research project, the prototype systems were evaluated in terms of usability, effectiveness and efficiency. The results of the evaluation survey are being analyzed and will be published in future.

Some of the outcomes of the spatial metadata automation research project were also used by the authors in designing and developing a metadata tool for the Australian Urban Research Infrastructure Network (AURIN) platform. The next section briefly reviews this tool.

6. Spatial Metadata Tool for Australian Urban Research Infrastructure Network (AURIN)

The Australian Urban Research Infrastructure Network (AURIN) is a $20 million project funded by the Australian Government's Super Science scheme. AURIN aims to provide built environments and urban researchers, designers and planners with infrastructure to facilitate access to a distributed network of aggregated and disaggregated datasets and information services (Pettit *et al*., 2013).

In the earliest stages of the AURIN project, design and implementation of a spatial metadata tool was defined as an integral technical requirement to enable the AURIN portal data administrators and custodians to create and edit associated metadata and therefore register datasets into the AURIN federated data architecture (Sinnott *et al*., 2013). Once the business requirements of the project were identified, a Metadata Profile was prepared for collecting the required metadata records for the AURIN platform based on the instructions recommended by the AS/NZS ISO 19115: 2005 (Figure 8).

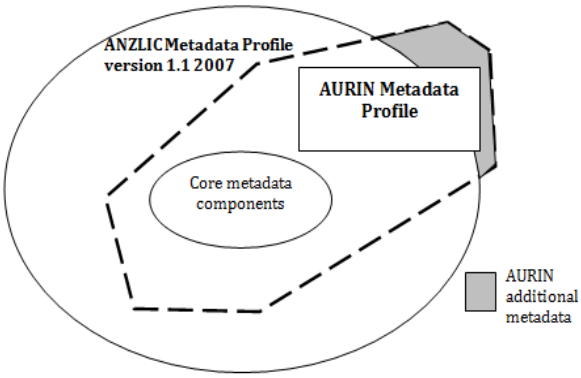

Figure 8. AURIN metadata profile structure, adopted from (AS/NZS ISO 19115: 2005)

Based on the functional and non-functional business requirements, a metadata tool was developed in Java (back end) and JavaScript (front end) using Jetty server and PostgreSQL database. Figure 9 illustrates an overview of the metadata tool GUI.

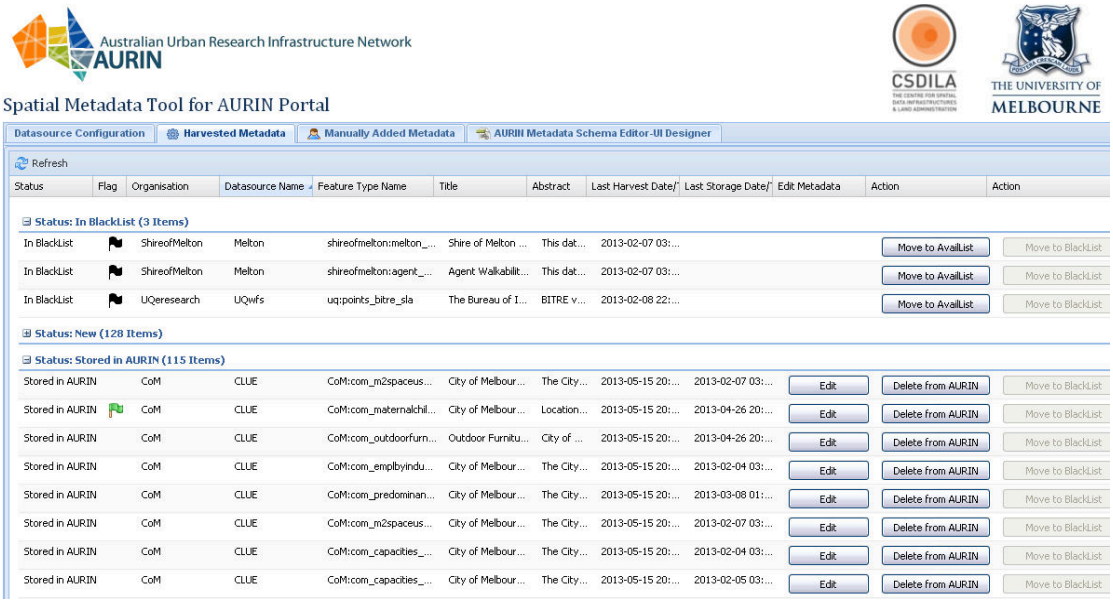

Figure 9. GUI of metadata tool for AURIN platform

This tool provides the AURIN metadata administrators and data custodians with different functionalities. Some of these functionalities are summarized as following:

- Providing metadata entry/edit user interface designer – this tool enables the AURIN administrators to dynamically design the metadata entry/edit user interface. Creating different tabs and groups and inserting metadata elements (from AURIN Profile) in each group are supported by this functionality.
- Data sources configuration – this tool enables the users to configure WFS and RDBMS data sources for harvesting and creating metadata.
- Automatic harvesting metadata – this tool uses the GeoNetwork opensource engine to harvest the metadata records from WFS data sources. The harvested metadata elements are then mapped to AURIN Metadata Profile using a stylesheet (xslt file) implemented for this purpose.
- Management of harvested metadata records – this tool provides the users with different options to manage the harvested metadata. These options are 'edit metadata content', 'store metadata in AURIN repository', 'delete metadata from AURIN repository', 'move metadata to black list (not being harvested in the future)', and 'move metadata record from black list to available list'.
- Metadata creation – this tool enables the users to manually create metadata records for RDBMS data sources. Once the RDBMS data source is configured, the users are able to create metadata for the datasets included in this data source.

- Harvesting dataset attributes – this tool provides services to automatically harvest dataset attributes for both WFS and RDBMS data sources. These attributes are then shown in the metadata entry/edit user interface along with the other metadata values.
- Re-harvesting metadata records – this tool automatically harvests metadata from WFS data sources in user-defined periods.
- Automatic detection of metadata changes – this tool is able to detect the changes in metadata content after any re-harvesting and notify the AURIN users about these changes by flagging the updated metadata values. The tool also provides the users with functions to affect or ignore the changes for each updated metadata value.
- Editing AURIN metadata schema – following the dynamic nature of AURIN Metadata Profile, this tool provides the AURIN administrators with a user-friendly and easy-to-use interface to edit the metadata schema.

The AURIN metadata tool is currently used by different data custodians involved in the AURIN project. For example the Population Health and Information Development Unit (PHIDU), at the University of Adelaide have over 150 datasets available for end user to search, discover and download from the AURIN portal. The data custodians as PHIDU have used the AURIN metadata tool to register each of the datasets into the federated data architecture.

7. Conclusion

In order to address the main challenges and requirements regarding the spatial metadata management and automation identified by the 'spatial metadata automation' research project, a new framework was presented in this chapter. The framework was capable of integrating metadata creation with the spatial data lifecycle via a 'lifecycle-centric spatial metadata creation' approach, updating metadata in real time with dataset modification through an 'automatic spatial metadata updating (synchronization)' approach, and involving the end users in creating and improving the content of keyword metadata using an 'automatic spatial metadata enrichment' approach.

The framework took advantages of a new GML-based integrated data model for storing and bundling spatial data and metadata. This data model replaced the detached data model for storing and delivering spatial data and metadata separately and had the potential to address the challenges regarding spatial data interoperability, dataset format dependency, and real-time dataset and metadata updating. However, the research showed that for implementing a robust GML-based integrated data model for spatial data and metadata storage there is a need for a standard data model to store the (XML-based) ISO 19115: 2003 compliant metadata records in a relational database. The geospatial community currently lacks such a data model. Also, it was realized that there is a need to maintain metadata values at the feature-level, in addition to the dataset-level, at the same time with any change to the dataset.

The new framework was facilitated through Web 2.0 features (folksonomy and tagging) to interact with end users for creating and improving the content of keyword metadata element. Based on this, the end users could share their knowledge about datasets, visit the most popular search words, gather ideas regarding the available data, and access the datasets more quickly and simply. However, results from the evaluation phase of the research indicate that the automatic enrichment approach proposed in this research could be improved by integration with an ontology-based data discovery system. It is suggested that the integration will contribute to address the issues of ambiguity and heterogeneity of user-generated search words and facilitate the spatial data discovery through recording, assigning and visualizing the most relevant search words for describing datasets.

The chapter also reports on a real world application of the metadata tool recently used in Australia which was implemented to address the AURIN portal administrators' and data custodians' requirements regarding the metadata management. This tool provides the AURIN administrators and data custodian with a user-friendly and easy-to-follow environment to dynamically design a user interface for metadata entry and edit. Having used the metadata automation methods (e.g. harvesting metadata and attributes for WFS data sources), the AURIN metadata tool minimizes the amount of metadata values which need to be created and updated manually by end users.

Acknowledgements

The authors wish to acknowledge the support of the Australian Research Council Linkage Project (LP0883933) on Automating and Integrating Spatial Data and Metadata Processes, its industry partners and the members of the Centre for Spatial Data Infrastructures and Land Administration in the Department of Infrastructure Engineering, The University of Melbourne, in the preparation of this chapter and the associated research. However, the views expressed in the chapter are those of the authors and do not necessarily reflect those of these groups. The authors also wish to acknowledge the support from the AURIN project in funding the development of the AURIN metadata tool reported in this chapter.

References

AGIMO (2006), *Australian Government Information Interoperability Framework.* Australian Government Information Management Office (AGIMO), Department of Finance and Administration, Australian Government.

AS/NZS ISO 19115: 2005, *Geographic information – Metadata.* Standards Australia.

Batcheller, J.K., (2008), Automating Geospatial Metadata Generation – An Integrated Data Management and Documentation Approach. *Computers and Geosciences,* 34: 387-398.

Batcheller, J.K., Dunfey, R.I., Reitsma, F. and Gittings, B.M., (2007), Automating Geospatial Metadata Using ESRI's ArcGIS and Microsoft's .NET. *10th AGILE International Conference on Geographic Information Science.* Aalborg University, Denmark.

Brooke, J., (1996), SUS: A "Quick and Dirty'" Usability Scale. In: Jordan, P.W., *et al.* eds. Usability Evaluation in Industry. London: Taylor & Francis, 189-194.

CatMDEdit (2011), *User's Manual- Management of Repositories* [online]. Available from: http://iaaa.cps.unizar.es/software/index.php/Management_of_repositories#Extracting_automatic_metadata_from_data_files [Accessed 27 August 2011].

Chi, Y. L., (2009), A Consumer-centric Design Approach to Develop Comprehensive Knowledge-based Systems for Keyword Discovery. *Expert Systems with Applications,* 36(2): 2481-2493.

Cooper, A. K., Coetzee, S., Kaczmarek, I., Kourie, D., Iwaniak, A. and Kubik, T., (2011), Challenges for Quality in Volunteered Geographic Information. *AfricaGEO 2011.* Cape Town, South Africa.

Daratech (2011), *GIS/Geospatial Sales up 10.3% to US$4.4 billion Growth Forecast to Top 8.3% in 2011.* Cambridge, MA, USA: Daratech, Inc.

ESRI (2002), Creating a Custom Metadata Synchronizer. *An ESRI ® White Paper* [online]. Available from: http://www.esri.com/library/whitepapers/pdfs/metadata.pdf [Accessed 10 September 2009].

ESRI (2010), *Metadata Synchronization* [online]. Available from: http://webhelp.esri.com/arcgiSDEsktop/9.3/index.cfm?TopicName=Metadata_synchronization [Accessed 27 October 2011].

Göbel, S. and Lutze, K., (1998), Development of Meta Databases for Geospatial Data In The WWW. *6th ACM International Symposium on Advances in Geographic Information Systems 1998.* Washington DC, USA, 94-99.

Guptill, S.C., (1999), Metadata and Data Catalogues In Geographical Information Systems. *In:* Longley, P. A.*, et al.* eds. *Geographical Information Systems: Principles, Techniques, Management and Applications.* second ed.: John Wiley & Sons, Inc, 677-692.

ISO19115: 2003, *Geographic Information – Metadata.* Geneva, Switzerland: International Standardization Organization (ISO).

Kalantari, M., Olfat, H. and Rajabifard, A., (2010), Automatic Metadata Enrichment: Reducing Spatial Metadata Creation Burden through Spatial Folksonomies. *In:* Rajabifard, A.*, et al.* eds. *Spatially Enabling Society.* Luven University Press, 119-129.

Kalantari, M., Rajabifard, A. and Olfat, H., (2009), Spatial Metadata Automation: A New Approach. *In:* Ostendorf, B.*, et al.* eds. *Spatial Science Conference 2009 (SSC2009).* Adelaide, Australia, 629-635.

Limbach, T., Krawczyk, A. and Surowiec, G., (2004), Metadata lifecycle management with GIS context. *10th EC GI & GIS Workshop.* ESDI State of the Art, Warsaw, Poland.

Nebert, D., (2004), *Developing Spatial Data Infrastructures: The SDI Cookbook Ver 2.0.* Technical Working Group Chair, GSDI.

Olfat, H., Kalantari, M., Rajabifard, A., Senot, H. and Williamson, I.P., (2013), A GML-based Approach to Automate Spatial Metadata Updating. *International Journal of Geographical Information Science,* 27(2): 231-250.

Olfat, H., Kalantari, M., Rajabifard, A., Senot, H. and Williamson, I.P., (2012a), Spatial Metadata Automation: A Key to Spatially Enabling Platform. *International Journal of Spatial Data Infrastructures Research,* 7: 173-195.

Olfat, H., Kalantari, M., Rajabifard, A. and Williamson, I.P., (2012b), Towards a Foundation for Spatial Metadata Automation. *Journal of Spatial Science,* 57(1): 65-81.

Olfat, H., Kalantari, M., Rajabifard, A., Williamson, I.P., Pettit, C. and Williams, S., (2010a), Exploring the Key Areas of Spatial Metadata Automation Research in Australia. *GSDI-12 World Conference.* Garden City, Singapore.

Olfat, H., Rajabifard, A. and Kalantari, M., (2010b), Automatic Spatial Metadata Update: A New Approach. *XXIV FIG International Congress 2010.* Sydney, Australia.

OSDM (2009), *ANZLIC Metadata Profile Short User Guide for the ANZMet Lite* [online]. Available from: www.osdm.gov.au/Metadata/08_ANZLIC_MetaProfile_Short_%20ANZMetLite.pdf/?id=1002 [Accessed 24 August 2011].

Pettit, C.J., Tomko, M., Sinnott, R. and Stimson, R., (2013), Building an e-Infrastructure to Support Urban Research in Australia: a Lens Centered View, *Surveying & Spatial Sciences Conference 2013*, 15th -19th April, 2013, Canberra.

Rajabifard, A., Kalantari, M. and Binns, A., (2009), SDI and Metadata Entry and Updating Tools. *In:* Loenen, B.V., Besemer, J.W.J. and Zevenbergen, J.A. eds. *SDI Convergence. Research, Emerging Trends, and Critical Assessment.* Delft, 121-136.

Rajabifard, A., Kalantari, M., Binns, A. and Bennett, R., (2007), *Final Report on BlueNet-Anzlic ver 1.0.* Melbourne: The University of Melbourne.

Sinnott, R.O. Bayliss, C., Bromage, A., Galang, G., Grazioli, G., Greenwood, P., Macauley, A., Mannix, D., Morandini, L., Nino-Ruiz, M., Pettit, C., Tomko, M. Sarwar, M., Stimson, R., Voorsluys, W. and Widjaj, I., (2013), The Urban Research Gateway for Australia: Development of a Federated, Multi-disciplinary Research e-Infrastructure, International Workshop on Science Gateways, *Proceeding on the 5th International Workshop on Science Gateways*, Zurich, Switzerland, 3-5th June, 2013.

Taussi, M., (2007), *Automatic Production of Metadata Out of Geographic Datasets.* (MSc). Helsinki University of Technology.

Timpf, S., Raubal, M. and Kuhn, W., (1996), Experiences with Metadata. *7th International Symposium on Spatial Data Handling, SDH'96.* Delft, The Netherlands, 12B.31-12B.43.

Tsou, M.H., (2002), An Operational Metadata Framework for Searching, Indexing, and Retrieving Distributed Geographic Information Services on The Internet. *Lecture Notes in Computer Science.* Springer, Berlin, 313-332.

West, L.A. and Hess, T.J., (2002), Metadata as a Knowledge Management Tool: Supporting Intelligent Agent and End User Access to Spatial Data. *Decision Support Systems,* 32: 247-264.

Westbrooks, E.L., (2004), Distributing and Synchronizing Heterogeneous Metadata in Geospatial Information Repositories for Access. *In:* Hillman, D. I. and Westbrooks, E. L. eds. *Metadata in Practice.* Chicago.

CHAPTER 4

A Description of Spatial Data Infrastructure Stakeholders in Ghana Using the ICA Model

Wiafe Owusu-Banahene[1&2], Foster Mensah[3], Serena Coetzee[1], Antony K Cooper[1&4], Victoria Rautenbach[1], Kisco M Sinvula[1], Emma Nangolo[5] and Martin Hippondoka[6]

[1] Centre for Geoinformation Science, University of Pretoria, Pretoria, 0002, South Africa

[2] Department of Computer Engineering, University of Ghana, Legon, PMB, Legon Accra, Ghana

[3] Centre for Remote Sensing and Geographic Information Services, University of Ghana, Legon, PMB 17, Legon, Accra, Ghana

[4] Built Environment, CSIR, PO Box 395, Pretoria 0001, South Africa

[5] Independent Researcher, Windhoek, Namibia

[6] Department of Geography, History and Environmental Studies, University of Namibia, Namibia.

osobana@gmail.com; fmensah@ug.edu.gh; serena.coetzee@up.ac.za; acooper@csir.co.za; victoria.rautenbach@gmail.com; kisco_77@msn.com; ronnameke@hotmail.com; mhipondoka@unam.na

Abstract

The National Framework for Geospatial Information Management (NAFGIM) was a spatial data infrastructure (SDI) initiative in Ghana which started around the year 2000. NAFGIM was developed as an integral part of a national effort to manage spatial data pertaining to the environment and natural resources. It sought to bring together technology, policies, institutional resources and standards to enhance the production, storage, access and utilization of geographic data and information. NAFGIM is no longer functional but Ghana is again embarking on another SDI initiative. This precipitated key diagnostic questions, such as: what led to NAFGIM's decline and how can lessons learnt from NAFGIM inform current SDI developments in Ghana? The International Cartographic Association (ICA) has developed formal models of an SDI, including identifying six types of SDI stakeholders and their specializations. The ICA model has been applied to describe the Namibian SDI (NamSDI). In this chapter, we follow this work and use the ICA model to describe the types of stakeholders in NAFGIM, their contributions, roles and impact. Current SDI developments in Ghana

can benefit from this stakeholder analysis, because most of the NAFGIM stakeholders are still relevant in current SDI developments as part of the ongoing Land Administration Project. Research results confirm the value of modeling stakeholders of an SDI: Stakeholders are identified and their roles assigned, potential conflicts are identified and can be pro-actively mitigated, facilitating harmonized stakeholder participation. The results also contribute to understanding commonalities between stakeholders in different SDIs generally. This is important because SDIs provide access to the geographic information that is essential for sustainable development and for advancing science.

KEYWORDS: spatial data infrastructure (SDI), Ghana, stakeholder, NAFGIM, ICA model

1. Introduction

A spatial data infrastructure (SDI) refers to the infrastructure, i.e. the basic physical and organizational structures required to facilitate and coordinate the efficient and effective discovery and use of spatial data (Rajabifard *et al*., 2006; Jackson and Gardner, 2011). The concept of an SDI has been around for two decades and the definitions are still evolving. Georgiadou *et al.,* (2005) defined an SDI as a combination of technology, systems, standards, networks, people, policies, organizational aspects, geo-referenced data, and delivery mechanisms to end users.

Ghana is a sub-saharan African country (shown in Figure 1) with a land surface area of 239,460km^2, a population of 24.66 million (Ghana Statistical Service, 2010) and a Gross Domestic Product (GDP) of GHS 73,101.9 million[1] (Ghana Statistical Service (2013).

Ghana came close to the establishment of a legally mandated SDI through efforts by the government, the World Bank and other donors to respond to the challenges of striking a balance between economic development and sustainable management of renewable resources. Through these efforts the National Framework for Geospatial Information Management (NAFGIM) was started in 2000 and later became the *de facto* SDI. Indeed, Masser (2005) highlights NAFGIM as one of the early SDIs in Africa. Currently, NAFGIM is no longer functional but Ghana is again embarking on an SDI initiative. What led to the failure of NAFGIM? How can lessons learnt from NAFGIM inform current SDI developments in Ghana?

The Commission on Geoinformation Infrastructures and Standards of the International Cartographic Association (ICA) has been using the Reference Model for Open Distributed Processing (RM-ODP) (ISO/IEC 10746-1:1998) to develop formal models of an SDI from the enterprise and information viewpoints of RM ODP (Hjelmager *et al*., 2008), and from the computational viewpoint (Cooper *et al*., 2013). These viewpoints

[1] Equivalent to USD 36.5 Billion at the time of writing this chapter.

contribute toward a more holistic interpretation of an SDI, independent of specific SDI legislation, technology and implementations (Cooper *et al*., 2013).

Specifically, the enterprise viewpoint describes the purpose, scope and policies for an SDI, and the relationship of an SDI to its environment, its role and the associated policies (Hjelmager *et al*., 2008). A key part of the enterprise viewpoint analysis was to identify the general roles of stakeholders within and around an SDI: Policy Maker, Producer, Provider, Broker, Value-added Reseller (VAR) and End User (Hjelmager *et al*., 2008). The Commission also identified 37 special cases of these general roles (Cooper *et al*., 2011), which were used as the template to clarify the different stakeholders for this investigation of NAFGIM and are included in Tables 1 and 2. While Cooper *et al*., (2011) is readily available online, we have included the descriptions of all 37 special cases in an annex. Further, it was beyond the scope of our research described here to interrogate or test the model in detail. Nevertheless, while we did find the model to be robust and useful, we have identified some issues with the model, as outlined below in Section 4 that we will take forward in further research.

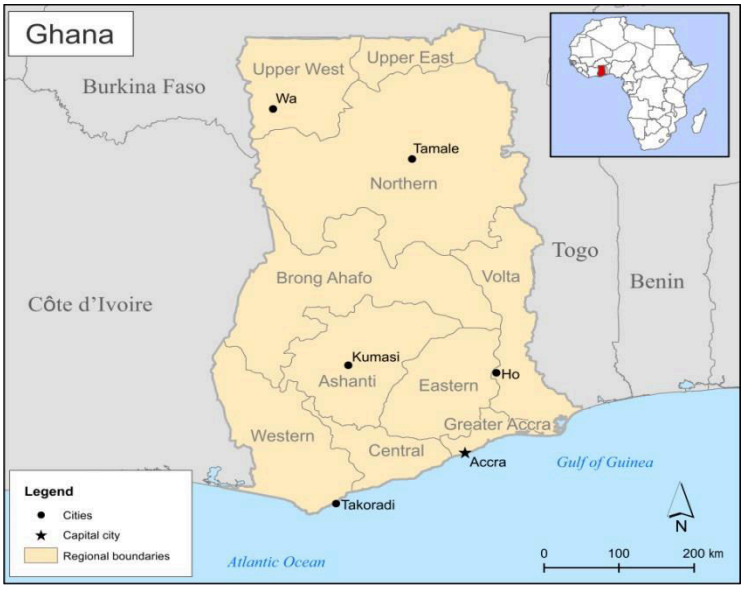

Figure 1. Location of Ghana in Africa

The ICA model has been applied to the Namibian Spatial Data Infrastructure (NamSDI) (Sinvula *et al*., 2013). The motivation for conducting the NamSDI stakeholder's analysis was to contribute towards the successful implementation of SDI in Namibia from a scientific perspective. The ICA model used in contextualizing a policy and legislative dependent NamSDI was robust, in which the roles, interests and motivation of stakeholders involved in NamSDI were identified. This contributed significantly to the holistic interpretation of NamSDI based on specific SDI legislation, technology and strategic implementations. For example, the Government of the Republic of Namibia

(GRN) was the notable policy maker and producer of fundamental (base/reference) spatial datasets, through various line ministries and state owned agencies (Sinvula *et al*., 2013).

In this chapter, the ICA model is used to describe and analyses SDI stakeholders in Ghana, building upon previous research on the Namibian SDI (Sinvula *et al*., 2012, 2013). To date, such an analysis has been published only for Namibian SDI stakeholders and we followed the structure and methodology applied in (Sinvula *et al.,* 2013). The analysis of SDI stakeholders in Ghana contributes to a better understanding of which stakeholders were involved in NAFGIM and what their contributions and roles were. Lessons for the current initiative and the future can be learnt from the types of stakeholders involved and their impact on NAFGIM. The analysis enables us to comment on the stakeholder types in the ICA model, e.g. are the same types of stakeholders involved in a developing country (e.g. Ghana and Namibia) as in developed countries (which influenced the development of the ICA model)?

The remainder of the chapter is structured as follows: section 2 presents SDI developments in Ghana (summarized in Figure 2); section 3 describes the methodology applied; section 4 describes the stakeholders; section 5 discusses the results and section 6 provides the conclusion.

2. SDI Developments in Ghana

We have compiled Figure 2 to show a timeline of the SDI developments in Ghana from 1988 through to today, which are discussed in detail here. As part of plans by the World Bank and other international donors to promote the development of Environmental Information Systems in Sub-Saharan Africa (EIS-SSA), a continent-wide program to support a series of National Environmental Action Plans (NEAPs) in Africa was established (Ezigbalke, 2004). It started in the late 1980s to early 1990s in response to the challenges of striking a balance between economic development and sustainable management of renewable resources. In March 1988, the Government of Ghana initiated preparation of a NEAP which was adopted in 1991.

The NEAP preparation identified land information availability as a priority and provided an opportunity for a more coherent framework on environmental and resource management information. In 1991 when the NEAP was finalized for Ghana, a National Environmental Information System (NEIS) was proposed to rectify the deficiencies on the state of environmental information. This led to the design of the environmental information system (EIS) development, a sub-component of the Environmental Resource Management System (ERMS) of the Ghana Environmental Resource Management Project (GERMP), a five-year project to implement the NEAP which started in 1993. The EIS was aimed at strengthening institutions involved in the collection, processing and analysis of environmental information and at the creation of core datasets for environmental planning in Ghana.

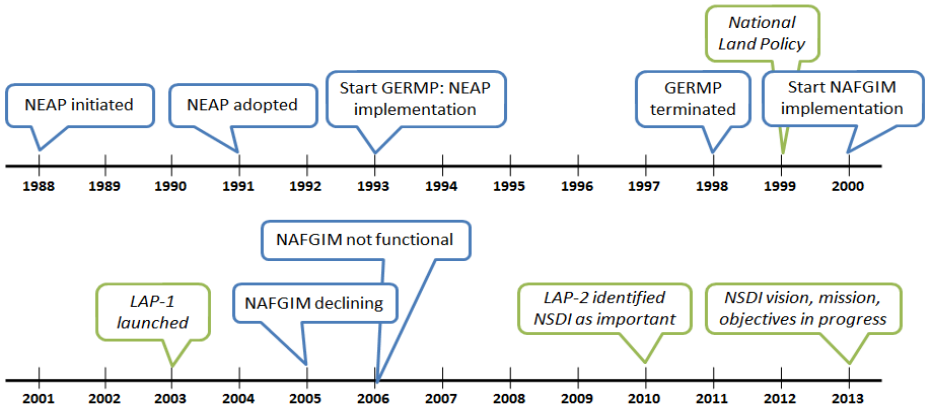

Figure 2. SDI developments in Ghana from 1988 until today

The Survey Department, the Lands Commission, the Soil Research Institute, the Meteorological Services Department, and the Centre for Remote Sensing and Geographic Information Services (CERSGIS) were identified to produce and collate the relevant land-related datasets for the project, under the sponsorship of the Government of Ghana, the World Bank and the Danish International Development Agency (DANIDA).

As the project proceeded, the participating institutions felt the need for, and initiated the creation of, a framework for sharing data and for coordinating the production and harmonization of their spatial data products. This initiative led to the birth of NAFGIM. NAFGIM's implementation started in April 2000 with a secretariat, a steering committee and an inter-agency forum. The secretariat, comprising a secretary, technical staff and a coordinator, was located at the Environmental Protection Agency (EPA). The steering committee constituted the policy-making body, while the inter-agency forum constituted a loose configuration of geospatial data producers and users that established mechanisms for the harmonized exchange of inter-sectoral information.

NAFGIM worked through technical workgroups that dealt with broad thematic areas. NAFGIM evolved to become the *de facto* SDI of Ghana (Ezigbalike, 2004; EPA, 2005; Cooper and Gavin, 2005; UNEP, 2010) and presented an opportunity for the establishment of a legally mandated SDI in Ghana. Between 2000 and 2005, Crompvoets and Bregt (2007) periodically conducted surveys, taking inventory of national clearinghouses on the Web by measuring eleven characteristics of a clearinghouse, such as the number of data suppliers, the number of datasets available and the number of monthly visitors. From these characteristics a clearinghouse suitability index was calculated from 2002 to 2005. The index showed that NAFGIM was declining: in 2005 it scored 21, 14 points lower than in 2002. Currently, NAFGIM is no longer functional (Karikari, 2006; Yawson *et al*., 2010).

In 2003, the Land Administration Project (LAP-1) was launched as a long term (15-25 years) Land Administration Program to implement policy actions recommended in the National Land Policy document of June 1999 (Ministry of Lands and Forestry, 1999). The programme appraisal recognized the need for up to date maps to support critical on-going land administration operations in support of agriculture, forestry, environmental management, urban and regional planning, mining, municipal services, storm water and sewerage, property tax, building permits, valuation systems, titling and deeds registration, infrastructure systems such as electricity, telecommunications, water, gas and real property maps, all potentially supporting land markets and national development.

In 2010, the programme therefore identified as part of LAP-2, the development of a surveying and mapping policy, a geodetic reference network, continuously operating reference stations, a national spatial data infrastructure, production of digitized base maps and the establishment of a street addressing system as important activities for consolidating, regulating and strengthening land administration and management systems in Ghana (Ministry of Lands and Natural Resources, 2011). This presented another opportunity for the establishment of a legally mandated SDI in Ghana, after the demise of NAFGIM.

Today, a consultant has been engaged and consultations are currently ongoing to develop the National SDI (NSDI) vision, mission statements and objectives for Ghana. These NSDI guiding principles to build NSDI partnerships, to ensure the creation of adequate capacity to empower the NSDI, to raise awareness and to communicate in the most effective ways to ensure that NSDI objectives are met, and to develop a technological framework to enhance access to spatial data, its use and sharing. In this chapter we use GERMP and NAFGIM as examples to describe SDI stakeholders in Ghana.

3. Methodology

Hjelmager *et al*., (2008) identified and described six stakeholders in the enterprise viewpoint of an SDI using Unified Modeling Language (UML) case diagrams and recognized that an individual stakeholder can execute different roles. For example, an organization can act as a policy maker, who sets out rules and policies for an SDI, and at the same time, be a producer of data and services required in an SDI. Cooper *et al*., (2011) took this further by identifying various special cases of these general roles which they termed ‘sub-types’ and ‘sub-sub types’.

The ICA’s model describes the characteristics of an SDI at a high level of abstraction. The objective here is to model SDI stakeholders in Ghana as sub-types of the ICA’s model stakeholders. Such a modeling exercise improves the understanding of stakeholders. Completion of the exercise also allows one to comment on the behavior and applicability of the abstract ICA model to a specific SDI instance, even though this exercise was not meant to be an examination of the model.

The stakeholder types and sub-types in the SDI in Ghana were described by associating ICA stakeholder types with NAFGIM stakeholders in Ghana and the roles they played in NAFGIM when it was functional. The identification of stakeholders was done based on direct observations in the form of personal involvement and experiences, impressions and literature (refer to Table 2). In the discussion, the current developments and stakeholders of LAP-2 are discussed and compared to NAFGIM's stakeholders. The end result is an identification and discussion of past and potential future stakeholders and their roles in an SDI in Ghana.

4. Stakeholders in SDI in Ghana

In this section, we describe according to the ICA model, the stakeholders that participated in NAFGIM. Table 1 shows the description of the six stakeholders in the ICA model. Table 2 shows ICA stakeholder types, stakeholder sub-types and stakeholder sub-sub-types with examples identified from NAFGIM. We describe the individual stakeholder types in NAFGIM in sub-section 4.1 through to sub-section 4.6. The identification of stakeholders was done based on direct observations in the form of personal involvement and experiences, impressions and literature such as Environmental Protection Agency (EPA) (2005); Ezigbalike (2004); Lance and Bassolé (2006); Ministry of Lands and Forestry, (1999); Cooper and Gavin (2005); United Nations Environment Programme (UNEP) Report (2010) and Yawson *et al*., (2010).

Stakeholder	Description
Policy maker	*A stakeholder who sets the policy pursued by an SDI and all its stakeholders*
Producer	*A stakeholder who produces SDI data or services*
Provider	*A stakeholder who provides data or services to users through an SDI*
Broker	*A stakeholder who brings users and providers together and assists in the negotiation of contracts between them*
Value-added reseller (VAR)	*A stakeholder who adds some new feature to an existing product or group of products, and then makes it available as a new product*
End user	*A stakeholder who uses the SDI for its intended purpose*

Table 1. Types of stakeholders in the ICA model (Hjelmager *et al*., 2008)

4.1 Policy Maker

NAFGIM was established through the implementation of NEAPs and EIS in Ghana. In March 1988 the Government of Ghana initiated the NEAP which was adopted in 1991. The Government of Ghana is therefore the ***policy maker***. Ghana currently has no SDI Act but Parliament is the *legislator* that is expected to pass bills into acts. The ***decision***

maker in NAFGIM was a steering committee which acted as the policy-making body. NAFGIM had a ***secretariat*** comprising a secretary, technical staff and a coordinator who was located at the EPA. The Government of Ghana, the World Bank and DANIDA were the ***champions*** of NAFGIM, as can be seen from Table 2. They were motivated by the necessity to promote sustainable development in the country.

4.2 Producer

Under the GERMP, which was part of the EIS-SSA, the major official data producers were the Survey Department, the Lands Commission, the Soil Research Institute, the Meteorological Services Department, and CERSGIS, who were brought together to produce and collate the relevant land-related datasets for the project. The Survey Department is responsible for producing the geodetic framework, aerial photographs and digital elevation model (with the Soil Research Institute). It also produces international, national, regional, district, metropolitan, municipal and town boundaries. Additional public data-producing institutions participated in NAFGIM, including the EPA, were the Department of Feeder Roads, the Water Research Institute, the Forestry Commission, the Soil Research Institute, the Ghana Statistical Service, the Electoral Commission and the Ghana Meteorological Services Department.

Some private companies, such as Rudan Engineering and GeoTech, were involved in NAFGIM as contractors or agents who worked for the Survey Department. As shown in Table 2, CTK Network Aviation Ltd was a **commercial mapping agency** that invested in the production of data for NAFGIM, hoping to get a return on investment from the government in future. No ***Community Interest*** and ***Crowd Sourced*** producers were identified for NAFGIM.

We identified CERSGIS as a stakeholder motivated by ***special interest*****,** as it produced maps of the social infrastructure for local communities. The production of flood hazard maps for the Western region of the country made CERSGIS perform the role of a stakeholder motivated by ***process***. No ***passive producer*** was identified. The NAFGIM Secretariat received revision notices and also acted as ***database administrator***. We identified ***interested amateur***, ***expert amateur*** and ***expert professional*** producers for NAFGIM, as shown in Table 2, but not any ***neophyte*** and ***interested amateur*** producers.

4.3 Provider

All the official data producers were identified as stakeholders of NAFGIM as they ***provided data*** and ***services*** for their own use and for the use by others**.** EPA was a ***distributor of data*** packaged by CERSGIS, as shown in Table 2. For example, CERSGIS packaged the datasets developed under GERMP according to districts and regions through District and Regional Information Systems, referred to as 'Regional and District Packaging'. These datasets included topographical data, land cover/land use data, soil

and land suitability data, land ownership data and meteorological data and were packaged on CD-ROMs for distribution by the EPA. 'Ghana country at a glance' was distributed freely on CD-ROM and via a website (which is not functional any more). The NAFGIM Secretariat acted as a ***data arbiter***. We did not identify a ***service distributor*** or ***service arbiter*** in NAFGIM.

4.4 Broker

No ***crowdsourcing facilitator*** was identified. Private companies, such as Rudan Engineering, CTK Network Aviation Ltd and GeoTech acted as ***clients/users finders, providers finders*** and ***négociants***, as presented in Table 2. The NAFGIM Secretariat played the role of both ***cataloguer*** and ***harvester.***

4.5 Value-Added Reseller

CERSGIS was a ***publisher*** of satellite imagery. Satellite imageries were processed by CERSGIS into products such as satellite images of Ghana from 1990 to 2000 and AVHRR data were re-sampled and geo-rectified. We did not identify any ***service integrator. Data and metadata aggregator/integrator*** *value-added resellers* included EPA, CERSGIS, the Soil Research Institute, the Meteorological Services Department, the Survey Department (now the Survey and Mapping Division of the Lands Commission) and the Lands Commission, as indicated in Table 2. Examples of data that were aggregated and/or integrated were a land cover atlas for Ghana – 1998, a land cover/land use data – 2003, and a land suitability atlas and bulletins.

4.6 End User

Citizens, visitors, government employees, consultants and private companies were identified as ***naïve consumers*** (when using whatever is available with limited ability to determine the quality of the data or services (Cooper *et al.*, 2011)) or ***advanced users***. These are shown in Table 2.

Stakeholder type	Stakeholder sub-type	Stakeholder sub- type	Examples
Policy Maker	Legislator		Parliament of Ghana
	Decision maker		NAFGIM steering committee
	Secretariat		NAFGIM Secretariat within EPA
	Champion		Government of Ghana, DANIDA, World Bank
Producer	Status	Official mapping agency	• Survey Department • Lands Commission • Soil Research Institute • Meteorological Services Department • CERSGIS

			• EPA • Department of Feeder Roads • Additional public institutions that participated in NAFGIM
		Commercial mapping agency	CTK Network Aviation Ltd
		Community interest	None
		Crowd source	None
	Motivation		
		Special interest	CERSGIS (e.g. community-based social infrastructure mapping)
		Economic	All of the data producers
		Process	CERSGIS (e.g. flood hazard and health risk maps for selected districts of Ghana)
	Role		
		Captor of raw data	All of the data producers
		Submitter of revision notice	Submitted to the NAFGIM Secretariat
		Passive producer	None
		Database administrator	NAFGIM Secretariat
	Skill	Neophyte	Unlikely
		Interested amateur	Unlikely
		Expert amateur	Special interest data and rainfall data
		Expert professional	Many examples
		Expert authority	Many examples
Provider			
	Data provider		
		A producer that is its own data provider	• Survey Department • Lands Commission • Soil Research Institute • Meteorological Services Department • CERSGIS • EPA • Department of Feeder Roads Additional public institutions that participated in NAFGIM
		Data distributor	EPA: CERSGIS packaged datasets for EPA to distribute e.g. • 'Regional and District Packaging' of GERMP data on CD-ROM, • 'Ghana country at a glance' free data distribution
		Data arbiter	NAFGIM Secretariat

	Service provider	A producer that is its own service provider	All of the official data producers of NAFGIM
		Service distributor	None yet
		Service arbiter	None yet
Broker			
	Crowd-sourcing facilitator		None yet
	Finder	Clients/users finder	Private companies
		Providers finder	Private companies
	Harvester		NAFGIM Secretariat
	Cataloger		NAFGIM Secretariat
	Négociant		• Private companies e.g. Rudan Engineering, CTK Network Aviation Ltd and GeoTech • Technical advisors in ministries • Public-private partnerships (PPP)
Value-added reseller (VAR)			
	Publisher		CERSGIS- satellite imageries (image processing) e.g. satellite images of Ghana 1990 and 2000, Advanced Very High Resolution Radiometer (AVHRR) data re-sampled and geo-rectified
	Aggregator/ integrator	Service integrator	None
		Data and metadata aggregator/ integrator	• EPA • CERSGIS • Soil Research Institute • Meteorological Services Department • Survey Department (now Survey and Mapping Division of Lands Commission) • Lands Commission Examples of data aggregated/integrated: - Land cover atlas for Ghana –1998 - Land cover/land use data –2003 - Land suitability atlas and bulletins
End user			
	Naive consumer		Citizens and visitors, government employees, consultants and private companies
	Advanced user		Citizens and visitors, government employees, consultants and private companies

Table 2. Stakeholder types and sub-types in the SDI in Ghana

5. Discussion

The modeling of SDI stakeholders in GERMP and NAFGIM remains relevant today. The current SDI developments in Ghana under LAP-2 are expected to introduce some new SDI stakeholders, though many of the stakeholders identified under NAFGIM will maintain their roles. For example, regarding *policy makers*, the *legislator* is still the Parliament of Ghana, but the LAP Secretariat is now the stakeholder in the role of *decision maker* and *secretariat* under LAP-2. The *champions* are the Government of Ghana and the World Bank. DANIDA is no longer playing the role of a champion as it did under NAFGIM. On the *producer* side, all the *official mapping agencies* identified under NAFGIM will maintain their roles. However, the Survey Department under NAFGIM is the Survey and Mapping Division of the Lands Commission. In NAFGIM, CTK was a *commercial* mapping agency, but there is no such agency under LAP-2.

Furthermore, potential *passive producers*, *neophyte* and *interested amateur* stakeholders are expected to participate in the current SDI developments in Ghana. An example is the Google platform introduced in Ghana recently. Users of GPS and mobile devices are also expected to contribute data to the SDI in Ghana.

The collective knowledge of the participants in the modelling exercise contributes to the completeness of the model and provides a snapshot of their collective knowledge about SDI stakeholders. Nevertheless, one has to assume that unless there is an official SDI with clear delineation of what is 'within' and what is 'outside' the SDI, there could be additional stakeholders and SDI-related activities that are not yet represented in the model. This confirms that SDI-related activities exist independently of an official SDI but that there is a need for a coordinating role, for example, to provide a central point of access to metadata about available datasets. In the case of Ghana, the LAP-2 work aims to provide such a technological framework which will enhance access to spatial data, its use and sharing.

The ICA model failed to take into account the level of geographical information systems, historical initiatives of national SDIs and developmental contexts of countries. It is more at an abstract level and more applicable to the developed and industrialized nations, for example, as many end users in developing countries are naïve in terms of spatial data and ICT. The stakeholder roles and interests were more subjective and to some extent not factual.

Furthermore, some stakeholders and sub-types are not included in the ICA model. For example, the stakeholder *producer* does not include sub-type *services*, even though the definition included the production of services. Thus the *producer of services* is not included in the ICA model. Suppliers of hardware and software were also excluded from the model. Moreover, some of the definitions such as *community interest* and *crowd sourcing* are so close to each other and should therefore be re-examined. For instance, should community interest and crowd source be combined as NGO/not for profit? A source of ambiguity we encountered was the fact that stakeholder sub-types are not mutually exclusive, e.g. status, motivation and role describe different aspects

of a stakeholder, but do not represent a sub-classification (tree) under the producer stakeholder type (as we had initially understood).

The analysis of the SDI stakeholders in Ghana revealed that data sharing, collection and distribution activities can be coordinated without a legal mandate, as long as projects continue and funds are available, because NAFGIM did not have a legal mandate, but functioned effectively for some years. Coordinated data production and sharing took place in Ghana in the 1990s under GERMP and spawned the development of the NAFGIM framework focusing on environmental and sustainable development information. When GERMP and related projects ended, funds, motivation and the imperative to sustain NAFGIM faded. The question is whether LAP with its planned legal mandate and inclusion of all kinds of spatial data will succeed in the future.

Society requires SDI and data for sustainable development and science cannot progress without SDI and data. In these aspects, this research contributes to understanding SDI stakeholders and their commonalities. It helps identify stakeholder participation upfront. The modeling exercise can be used to avoid repeating past mistakes (e.g. when drafting policies) and to minimize stakeholder conflict.

6. Conclusion

Our experiences show that there is value in modeling the stakeholders in an SDI: It clarifies who the stakeholders are and what their roles and contributions are or could be in the SDI of the country. In the case of Ghana, current SDI developments can benefit from the stakeholder analysis of NAFGIM presented in this chapter. The project determination that led to the deterioration of NAFGIM has already informed policies and strategies of current SDI developments. The current SDI developments in Ghana can also benefit from a comparison of SDI stakeholders in different countries which we seek to do in future studies.

The modeling exercise described in this chapter not only improves the understanding of stakeholders in Ghana, it also serves to test the behavior and applicability of the abstract ICA model to a specific SDI instance. In future, we aim to compare the results of applying the ICA model to SDIs in Namibia, Ghana and other countries to further identify key aspects of SDIs, to improve the understanding of SDI stakeholders and to make recommendations for the improvement of the ICA model, such as to deal with the issues we identified here.

Acknowledgements

We acknowledge with gratitude funding received through the Joint Research Grant under the SA-Namibia Research Partnership Programme towards this study. The map in Figure 1 was produced by Yvette Bevis of the Centre for Geoinformation Science,

University of Pretoria, South Africa. We would also like to thank the anonymous referees for their useful comments.

References

Allotey, J.A., (1999), *Developing A National Framework for Geo-Spatial Information in Ghana* Report on First Meeting of the Committee on Development Information (CODI) 28 June-2 July 1999, Addis Ababa, Ethiopia.

Coleman, D.J., Georgiadou, Y. and Labonte, J., (2009), Volunteered Geographic Information: The Nature and Motivation of Producers. *International Journal of Spatial Data Infrastructures Research*, Special Issue on GSDI-11, 4.

Cooper, A.K. and Gavin, E.J.O., (2005), Spatial Metadata in Africa and the Middle East, in Moellering, H., Aalders, H.J.G. and Crane, A. (Eds) *World Spatial Metadata Standards*. Amsterdam: Elsevier.

Cooper, A.K., Moellering, H., Hjelmager, J., Rapant, P., Delgado, T., Laurent, D., Coetzee, S., Danko, D. M., Düren, U., Iwaniak, A., Brodeur, J., Abad, P., Huet, M. and Rajabifard, A., (2013), A Spatial Data Infrastructure Model from the Computational Viewpoint, *International Journal of Geographical Information Science*, in Press.

Cooper, A.K., Rapant, P., Hjelmager, J., Laurent, D., Iwaniak, A., Coetzee, S., Moellering, H. and Düren, U., (2011), Extending the Formal Model of a Spatial Data Infrastructure to Include Volunteered Geographical Information, *25th International Cartographic Conference (ICC)*, Paris, France, 4-8 July 2011.

Crompvoets, J. and Bregt, A., (2007), National Spatial Data Clearinghouses, 2000-2005',in Onsrud, H. (Ed.) *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ESRI Press, pp. 133-146.

Engelhardt, F., (2012), *Creating an Environmental Geographic Information System for the City of Kumasi, Ghana*. Master Thesis in Sustainable Development at Uppsala University.

Environmental Protection Agency (EPA), (2005), Country Paper – GHANA, Environmental Statistics in Perspective. *Proceedings of Workshop on Environmental Statistics for the ECOWAS Region, 28 February-4 March, 2005, Dakar – Senegal*, Available from http://unstats.un.org/unsd/environment/ghana.pdf [Accessed 8 April 2013].

Ezigbalike, D., (2004), Advances In SDI Development In Africa, *Proceedings 7th GISDECO, 10-12 May 2004, Johor, Malaysia,* Available from

http://www.gisdevelopment.net/proceedings/gisdeco/2004/keynote/ezipf.htm [Accessed 8 April 2013].

Georgiadou, Y., Puri, S.K. and Sahay, S., (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.

Ghana Statistical Service (2013), *Revised Gross Domestic Product 2012,* Ghana Statistical Service (April 2013), Available from http://www.statsghana.gov.gh/docfiles/GDP/revised_gdp_2012.pdf [Accesed 14 May 2013].

Ghana Statistical Service (2010), Population and Housing Census. Population by Region, District, Age Groups and Sex, 2010*, Ghana Statistical Service*, Available from http://www.statsghana.gov.gh/docfiles/pop_by_region_district_age_groups_and_sex_2010.pdf [Accesed 14 May 2013].

Hjelmager, J., Moellering, H., Cooper, A.K., Delgado, T., Rajabifard, A., Rapant P., Danko, D., Huet, M., Laurent, D., Aalders, H.J.G.L., Iwaniak, A., Abad, P., Düren, U. and Martynenko, A., (2008), An initial formal model for spatial data infrastructures. *International Journal of Geographical Information Science* 22(11&12): 1295-1309.

ISO/IEC 10746-1:1998, *Information Technology – Open Distributed Processing – Reference Model: Overview*, International Organization for Standardization (ISO), Geneva, Switzerland.

Jackson, M.J. and Gardner, Z., (2011), *World Bank SDI Report*, World Bank, Available from http://lgosmgb2.nottingham.ac.uk/elogeowiki/index.php/World_Bank_SDI_Report [Accessed 11 May 2013].

Karikari, I.B., (2006), Ghana's Land Administration Project (LAP) and Land Information Systems (LIS) Implementation: The Issues. Article of the Month*, International Federation of Surveyors*, Available from http://www.fig.net/pub/monthly_articles/february_2006/karikari_february_2006.pdf [Accessed: 9 May 2013].

Karikari, I., Stillwell, J. and Carver , S., (2005), The Application of GIS in the Lands Sector of a Developing Country: Challenges Facing Land Administrators in Ghana. *International Journal of Geographical Information Science,* 19(3): 343-362.

Lance, K. and Bassolé, A., (2006), SDI and National Information and Communication Infrastructure (NICI) Integration in Africa. *Information Technology for Development,* 12(4): 333-338.

Masser, I., (2005), *GIS Worlds: Creating Spatial Data Infrastructures*. ESRI Press, Redlands CA, ISBN 1-58948-11-4.

Ministry of Lands and Forestry (1999), Land Policy Document. Ministry of Lands and Forestry.

Ministry of Lands and Natural Resources (2011), Land Administration Project – Phase Two. Project Implementation Manual.

Nebert, D., (2004), *Developing Spatial data Infrastructures: The SDI Cookbook*. GSDI, Available from http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf [Accessed 10 May 2013].

Rajabifard A., Feeney, M. and Williamson, I.P., (2006), Directions for the Future of SDI development. *International Journal of Applied Earth Observation and Geoinformation*, 4: 11-22.

Sinvula, K.M., Coetzee, S., Cooper, A.K. and Hipondoka, M., (2012), Exploring the Potential Suitability of an SDI Model in Context of the National Spatial Data Infrastructure (NamSDI) of Namibia, *GISSA Ukubuzana 2012 Conference*, Ekurhuleni, South Africa, 2-4 October 2012.

Sinvula, K.M., Coetzee, S., Cooper, A.K., Nangolo, E., Owusu-Banahene, W., Rautenbach, V. and Hipondoka, M., (2013), A Contextual ICA Stakeholder Model Approach for the Namibian Spatial Data Infrastructure (NamSDI). *26th International Cartographic Conference (ICC)*, August 25-30, Dresden, Germany.

United Nations Environment Programme (UNEP) Report (2010), *Best Practices and Lessons Learnt from the Implementation of Africa Environment Information Network*, Available from http://www.unep.org/dewa/africa/docs/en/Best_Practices_and_Lessons_Learnt_from_the_Implementation_of_AEIN.pdf [Accessed 8 April 2013].

Yawson, D.O., Armah F.A. and Dadzie S.K.N., (2010), Ghana's Right To Information Bill: Opportunity For SDI as a Technical Infrastructure. *International Journal of Spatial Data Infrastructures Research*, 5: 326-346.

Annex 1. The Types and Sub-types of Stakeholders in an SDI, Taken verbatim from Cooper *et al.,* (2011).

Stakeholder type	Stakeholder sub-type	Stakeholder sub- sub-type	Description
Policy Maker			*A stakeholder who sets the policy pursued by an SDI and all its stakeholders, such as developing policies for VGI, soliciting for VGI, acceptance criteria, quality assurance (e.g. verification against other, independent VGI), etc.*
	Legislator		An 'external' authority (not obviously perceived as being part of the SDI, but in practice, a key stakeholder) that determines the framework within which the SDI has to exist, but the Legislator does not necessarily understand anything about the SDI. For INSPIRE, this would be the European Parliament.
	Decision maker		A participant in the SDI who makes policies (including initiating the SDI) and who understands geospatial data and the applications, constraints, etc. The Decision Maker is often a committee of representatives of stakeholder communities. For INSPIRE, this would be the INSPIRE Committee (IC).
	Secretariat		The 'glue' of the SDI keeping it all together. The Secretariat is often a department in government with the mandate and budget to support the SDI, and that can contract out services. Especially for an SDI of VGI, the Secretariat can start informally and then crystallize once funding is available to pay for participation (as happened with OpenStreetMap, for example, which only received core funding in its second year of operations [OpenStreetMap 2010]). For INSPIRE at the European level, this would be the Joint Research Centre (JRC), as the overall technical coordinator, and Eurostat, as the overall implementation coordinator. Specific roles of the Secretariat include: • Supporting and monitoring the implementation of policies, etc. • Facilitating communication between stakeholders, particularly to provide feedback

			(e.g. quality or popularity of a data set, viability of a data product specification, responses to draft policies). • Building the actual SDI (generally through contractors). • Ensuring the smooth running of processes. • Classification of stakeholders.
	Champion		Promotes the SDI, such as encouraging citizens to contribute VGI. The Champion does not necessarily have a mandate, but could be motivated by the need to promote social justice, by environmental awareness, or by commercial interest. The Champion could be the initiator of the SDI.
Producer			*A stakeholder who produces SDI data or services, such as a lay person who generates VGI.*
	Status		
		Official mapping agency	An organization with the budget, resources, expertise and mandate to perform mass data production across the whole of the area of interest, normally to a consistent specification across the whole area. These include topographical, cadastral, hydrographic, meteorological, geological, hydrological, social statistical, environmental and other mapping agencies. These are at all levels of government (local, provincial, national, regional and global).
		Commercial mapping agency	A for-profit organization producing data and products for its identified markets.
		Community interest	Produce general base data or specialized data with broad or narrow coverage, especially as VGI. Exhibits the 'long tail', with many contributors of small data sets and few contributors of most of the data. There will be many more End Users than Producers.
		Crowd source	Issue an open call for data to anyone (the crowd), often according to a specification and often with a reward (not necessarily financial). This includes citizen science projects.
	Motivation		
		Special interest	Produce data for their local area and/or for a narrow interest, such as to protect the environment, empower a community (e.g.

			asset-based community development) or counteract bias in official sources of data.
		Economic	Produce data for economic or financial reasons, such as for direct financial reward (e.g. as an employee, on contract or to sell), promoting awareness of a business (locations, products, services, special offers and opening hours), and End Users unwilling to pay for institutional data.
		Process	Produce data because of particular interest in the data capture processes *per se*, such as training for students (as a way to motivate them), or the mapping parties that combine data capture with social events.
	Role		
		Captor of raw data	Produce data such as locations measured by GPS or drawn from background images, categorization and description of features, photos and images.
		Submitter of revision notice	Submit a notice to revise or correct data in an SDI, performed most often by citizens to improve the data of their immediate environment. An example is swisstopo (Guélat 2009). This would comprise many contributors of very small data sets.
		Passive producer	Produce data through their mobile devices being tracked by a service provider, such as cellular telephones or in-car navigation devices, to monitor traffic flows, assess telecommunication network congestion, or for other purposes. Clearly, this raises ethical issues concerning informed consent, uninformed consent, surreptitious tracking and privacy.
		Data base administrator	Ensure that the database specifications are respected (e.g. by providing rules to integrate data in the database and by checking these rules are respected, by ensuring consistency checks, etc.).
	Skill		Coleman *et al* [2009] categorize the skill levels of users that are producers (which they identify with the neologism, producers), as (in their ordering):
		Neophyte	No formal background in a subject, but with the interest, time and willingness to offer opinions or data.
		Interested amateur	'Discovered' an interest in a subject and begun reading background literature, consulting colleagues and experts, experimenting with applications and gaining

			experience in appreciating the subject.
		Expert amateur	May know a great deal about a subject and practice it with passion on occasion, but does not rely on it for a living.
		Expert professional	Studied and practices the subject, relying on that knowledge for a living, and may be sued if their products, opinions and/or recommendations are proven inadequate, incorrect or libelous.
		Expert authority	Widely studied and long practiced a subject and now recognized to possess an established record of providing high-quality products and services and/or well-informed opinions – and stands to lose that reputation and perhaps their livelihood if that credibility is lost, even temporarily.
Provider			A stakeholder who provides data or services, produced by others or itself, to users through an SDI. Examples include an aggregator of VGI, such as Ushahidi, and the provider of the infrastructure for collecting VGI, such as OpenStreetMap.
	Data provider		
		A producer that is its own data provider	This is the classical model used by a national mapping agency.
		Data distributor	Holds the catalogs and data of Producers, to take the administrative burden away from the Producers in dealing with users. The Distributor does not assess the data they are redistributing; they are merely an agent for the Producer. This would include dissemination through a website or on CD-ROM, etc.
		Data arbiter	Selects datasets from Producers according their published criteria (i.e. performing quality assurance and even certification), but does not add value in any other way.
	Service provider		
		A producer that is its own service provider	This is the typical model used by a location-based service (LBS) provider (e.g. find a service or facility available where I am now).
		Service distributor	Makes services available through their website or runs the services internally for clients. The cloud-computing model is typical.
		Service arbiter	Selects services from Producers according their published criteria (i.e. performing

			quality assurance and even certification) and provides them through their website, but does not add value in any other way.
Broker			A stakeholder who brings End Users and Providers together and assists in the negotiation of contracts between them. They are specialized publishers and can maintain metadata records on behalf of an owner of a product. Their functions include harvesting metadata from Producers and Providers, creating catalogs, and providing services based on these catalogs. An example for VGI is a community-based organization that enables the members of its community to provide updates and corrections to the published information of their local authority, such as addresses.
	Crowd-sourcing facilitator		Such as Amazon Mechanical Turk, which allows businesses to access an on-demand, scalable work force by advertising small "human intelligence tasks" to be completed [Amazon 2010].
	Finder		
		Clients/users finder	Promotes and sells a portfolio of data and services from Producers, Providers and VARs, to End Users.
		Providers finder	Sources data or services for an SDI. In South Africa, for example, the State Information Technology Agency (SITA) has a mandate to procure services for government departments, providing tender evaluation and management, etc.
	Harvester		Harvests metadata on data and services and integrates them.
	Cataloguer		Builds and maintains a catalog.
	Négociant		A stakeholder who brings End Users and Providers together and assists in the negotiation of contracts between them. They are specialized publishers and can maintain metadata records on behalf of an owner of a product. Their functions include harvesting metadata from Producers and Providers, creating catalogs and providing services based on these catalogs. A VGI example is a community-based organization that enables the members of its community to provide updates and corrections to the published information of their local authority.
Value-added			A stakeholder who adds some new feature

reseller (VAR)			to an existing product or group of products, and then makes it available as a new product. An example is searching for, evaluating and integrating VGI (possibly also with official information), to create a new data set or product. It is important to realize that a VAR does not necessarily sell its products, but could generate its income from other sources (e.g. support services).
	Publisher		Takes data from various sources, and integrates and edits them to produce a new product, such as an atlas or a location-based service (LBS). A Publisher could add some of its own data.
	Aggregator/ integrator		
		Service integrator	Chains services together. Would often reside in the cloud.
		Data and metadata aggregator/ integrator	Selects, edits, enhances and combines data into a new offering: • Conflation of datasets (selecting the 'best' versions of features and attributes from across several data sets). • Aggregation of metadata (more complex to do for VGI because of the multitude of Producers and the patchwork nature of their contributions). • Integration of different data sets and their metadata.
End user			A stakeholder who uses the SDI for its intended purpose. Many End Users cannot differentiate between VGI and official information, unless they are told explicitly, and hence would use VGI transparently. End Users tend to use VGI for 'quick and dirty' purposes, such as navigation, because there are no issues of copyright or liability.
	Naive consumer		Uses whatever is available with limited ability to determine the quality of the data or services.
	Advanced user		Has expert domain and/or geospatial expertise and hence can make informed decisions about the data and services to use and can provide informed, technical criticism of the data and services. They often use a GIS or other advanced software.

CHAPTER 5

Analyzing Organizational Levers of Spatial Enablement

Ezra Dessers[1], Joep Crompvoets[2] and Geert Van Hootegem[3]

[1] KU Leuven, Centre for Sociological Research

[2] KU Leuven, Public Management Institute

[3] KU Leuven, Centre for Sociological Research

ezra.dessers@soc.kuleuven.be; joep.crompvoets@soc.kuleuven.be;
geert.vanhootegem@soc.kuleuven.be

Abstract

It is widely recognized that the effectiveness of a Spatial Data Infrastructure (SDI) depends on the uptake of spatial data use and sharing by organizations in support of their processes. In this chapter, spatial enablement refers to the extent to which spatial data handling supports the overall objectives of an organizational process. Case study research findings are presented which indicate that the presence of an integrated (as opposed to fragmented) process in which the spatial data related activities are embedded, can be related to a higher level of spatial enablement of the process. The mode of dividing tasks and the way of allocating spatial data related activities within a process can thus be regarded as organizational levers of spatial enablement.

KEYWORDS: spatial enablement, organizational levers, task division

1. Introduction

Spatial Data Infrastructure (SDI) is a complex concept with many facets, but essentially SDIs are about facilitating and coordinating spatial information flows (Crompvoets *et al.*, 2004; Masser, 2005). The effectiveness of an SDI depends on the uptake of spatial

data use and sharing by organizations in support of their activities (Harvey *et al*., 2006). These activities are part of the organization's processes. A process can be defined as the sequence of steps involved in producing products and services (Daft, 2001). Process performance then refers to these products and services in connection to what is expected from them by their users and society at large. Processes are defined as spatially enabled, when a high performing integration of spatial data is included in these processes (Dessers *et al*., 2012). This chapter analyzes the spatial enablement of processes within organizations. The main objective is to identify potential organizational levers of spatial enablement. The chapter is divided into six parts. While this introduction briefly described the background of the chapter, the next part introduces the research questions. In the third part, a research design is proposed that should allow answering these research questions. The fourth part describes the operationalization of the main concepts and variables, while the fifth part presents the results of an extensive case study that was performed in the region of Flanders (Belgium), between 2009 and 2011 (Dessers *et al*., 2012). The chapter ends in the sixth part with some concluding remarks.

2. Research Questions

An organization can be seen as a dynamic network of interdependent elements between which interactions occur. The complexity of the network is the result of splitting up processes into tasks that are carried out by the different elements of the network. Van Amelsvoort (2000) explains that, as the complexity of the network increases, also the coordination of the network becomes more complex and difficult to realize. If the number of separate elements involved in a specific process chain increases, the coordination of the process becomes more difficult to handle, as the risk for interferences rises. Due to the dependencies in the network, interferences risk to be passed on to the next element in the chain, resulting in escalation effects. Van Amelsvoort's approach explains why the overall coordination of a specific process might be hampered by process fragmentation.

The challenge of coordinating and implementing spatial data related activities in such a process might be even bigger, since the spatial enablement of a specific process could be regarded as an aspect-related objective. In systems theory (Luhmann, 1984), sub-systems refer to a subset of elements of a larger system (such as an department in an organization), while aspect systems refer to so-called relations between system elements (such as the economic aspect, the political aspect or the technical aspect) (In 't Veld, 1994). Figure 1 illustrates how an organizational sub-system may have several aspect systems, and conversely, how a certain aspect system may touch upon various sub-systems. The present chapter focuses on the spatial data related aspect system, in the sense that, in the process chain, spatial enablement refers to the spatial data related aspect of the various process activities. Other aspects may include funding, personnel management and quality assurance. The aspect-related objectives are in general not the primary objectives of the organizational sub-systems concerned. An aspect-related objective, like the spatial enablement of a specific process, would have

to be aligned with the core objectives of the process (which are directly related to the delivery of a certain product or service to the client or end user). For instance, the first objective of a spatial planner would be the design of land use plans, and not so much whether and how the resulting spatial datasets will become accessible according to the principles of the governing SDI. It should be noted that the concept of spatial enablement in fact also refers to the adoption of SDI objectives in the context of a specific process (with regard to data sharing, or the application of certain technical standards and procedures). De Vries (2009) for instance, reported that the European SDI 'INSPIRE' is adopted with varying success in processes within and between organizations.

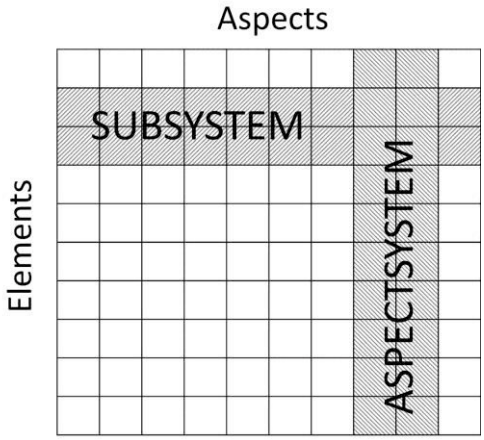

Figure 1. Sub-systems and aspect systems (adapted from: In 't Veld, 1994)

Based on a system-theoretical approach, it seems preferable to embed the aspect-related activities in the process activities. With regard to spatial enablement, this means that each organizational unit should be able to perform the needed spatial data related activities (possibly supported by a specialized spatial data unit at the organizational level).

This chapter investigates the proposition that a move towards a less fragmented task division, combined with an embedding of the spatial data related activities (further referred to as 'spatial data function') in the various process steps, contributes to the development of spatially enabled processes. The research questions are: (RQ1) What is the relation between the task division and the level of spatial enablement of a process? (RQ2) What is the relation between the allocation of the spatial data function and the level of spatial enablement of a process? The two research questions will be further developed and explained in Section 4.

3. Research Design

We chose to pursue a case study research design in order to answer the research questions. Case-based research is a widely used method for studying complex contemporary phenomena in their actual context (Yin, 2003). Since this research aims to assess how and why differences in operational process characteristics impact on their level of spatial enablement, the case study seems to be an appropriate research method (Dessers *et al*., 2012). The case study is focused on the public sector in the region of Flanders (Belgium). The case is defined as a process in which spatial data are accessed, used and shared. Three cases were selected: the development of zoning plans; the management of traffic accident registrations; and the mapping of flood areas. Within each case, a further selection was made of six to eight organizations.

The development of zoning plans is the first case in this research. A zoning plan is generally aimed at the development of a specific area ranging from a single parcel to an entire city district (RWO, 2008). Three governmental levels are authorized to develop zoning plans: the regional, the provincial and the municipal level. Each level has delineated powers with regard to spatial planning. Spatial data are used during plan preparation, in the plan design phase, and for exchanging draft plans with various stakeholders. A digital version of the final zoning plan can be used in other processes, like building permit delivery.

In the domain of road safety, spatial data are increasingly used for the registration of traffic accidents in order to monitor the progress in road safety and to evaluate the road safety policy (Van Malderen *et al*., 2009). This second case focuses on the acquisition, use and sharing of traffic accident data. The accident registration process deals with different stakeholders and data flows. The local and federal police are key players in this process as they compile the road accident forms which serve as the basis for the accident registration. All stakeholders involved at later stages in the process make use of these registrations.

The third case refers to the process of compiling, updating, using and distributing flood maps. Rather than just a single flood map, multiple maps exist, each of which play their own role in the policy on flooding. This case examines various organizations that contribute to the identification of recently flooded grounds, create modeled flooding areas, and contribute to the delineation of reservoir areas (or buffer zones) to hold water in the event of flooding.

The selections of these three inter-organizational processes, and the further selection of organizations within each process, were based on the expected variety in inter- and intra-organizational process characteristics. The selection was based on information from exploratory interviews and discussions with key stakeholders, consultation of various documents (such as brochures and annual reports) and survey results (Crompvoets *et al*., 2009; Callens, 2008). During the case study, information was gathered by way of multiple in-depth interviews in each organization, with process owners, GIS operators, team leaders, organizational experts, managers, GIS

coordinators, legal experts, ICT managers and database experts. All interviews were recorded, and transcripts of the interviews were produced. Besides these transcripts, various documents were collected. These documents and the interview transcripts form the raw material for further analysis. Once this raw material was collected, a first descriptive compilation resulted in a report. This report was sent to all the interviewees of the organization for validation purposes. The interviewees were requested to send back their comments, additions and corrections. Once all interviews of a case were finalized, a case workshop was organized for which all the interviewees were invited. At each of the three workshops, all organizations were represented by at least one of the interviewees. At the workshop, those present had the opportunity to discuss the draft version of the report with the researcher. The updated reports were the basis for the actual research analyses.

4. Operationalization

A central concept in both research questions is spatial enablement, which refers to the extent to which spatial data handling supports the overall process objectives (referred to as 'contribution to process performance' by Vandenbroucke *et al*., (2012)). Based on performance management literature (Bolwijn *et al*., 1990; Bekkers, 1998; Toonen, 2003) three variables were chosen for the operationalization of spatial enablement. The values of these variables are based on assessments by process owners and participants. The variable 'efficiency and quality' refers to the decrease of the input of people and means in the process, to the reduction of the lead time, to costs cuttings, to the avoidance of errors and confusion, and to the improvement of the output. The variable 'flexibility and innovation' encompasses dealing swiftly with differing requirements and fields of application, quickly adjusting the process when new demands are formulated during the course of the process, or developing multiple alternatives side by side. It also implies changing and ameliorating the proceeding of the process itself, or integrating new technological tools or organizational methods in the process. The variable 'transparency and reliability' is about customer-orientedness, offering the exact information a client is searching for, offering the citizen more insight into the proceeding of the process, improving legal security, clarifying the citizen's rights and obligations, offering him ways to control the process and to easily consult the related data and documents. The resulting values of these three variables were aggregated into a total value for each organization. A five-point rating scale was applied to rank the different organizations.

RQ1 is about the relation between task division and spatial enablement. Task division refers to the allocation of the various process steps within the organization, including production, preparation and support activities. In other words, it is the extent to which the different steps in the execution of the process are fragmented across various organizational units. It is expected that, as the level of fragmentation becomes larger, the challenge of coordinating and implementing spatial data related activities becomes greater, and spatial enablement becomes more difficult to achieve (see Section 2).

RQ2 deals with the relation between the allocation of the spatial data function and the level of spatial enablement. The extent to which the activities of collecting, using and distributing spatial data are integrated in the process is assessed. It should be clear that the variable spatial data function is different from spatial enablement. While spatial data function refers to the position of (possible) spatial data related activities with regard to the other process activities, spatial enablement is about contribution of the spatial data handling to process performance. Whether concentrating spatial data related process tasks in a specialized GIS unit offers the best chances for a high level of spatial enablement, or conversely, de-concentrating them to the teams responsible for the process tasks, is a question that many organization struggle with (Reeve *et al*., 1999; Crosswell, 2009). Therefore, the relation between both variables is made part of the second research question. An assessment is made of the degree to which the spatial data related activities are separated from the organizational units that perform the other process activities.

The research design is qualitative in nature. The set of variables are conceived as tools for guiding the data collection, analytically categorizing the data, identifying regularities and ensuring comparability between the selected organizations (Miles *et al*., 1994). The following method was used to assess the values for the variables division of labor and spatial data function. First, a description was made of the status of the organizations for each variable, based on the interview transcripts and the collected documents (such as organization charts). Second, the organizations were compared in order to assess their relative position on a five-point scale for each variable (low, medium/low, medium, medium/high, high). This assessment was done by the same researcher for all organizations of the three cases, in close consultation with the research team. It should be noted that the qualitative scale was applied as a tool to structure the data, in order to facilitate the comparative analysis. The technique of pattern-matching (Yin, 2003) was then used to compare the empirical patterns with those predicted by the proposition (as formulated at the end of Section 2).

5. Research Findings

Within each of the three cases, selected organizations were compared to study the research questions. First, the findings on the possible relation between task division and spatial enablement are presented (RQ1), and then the results for allocation of the spatial data function (RQ2) are provided. Radar charts are used to display the extent of (dis)similarity between the pattern of the various variables. Each radar chart consists of a sequence of equi-angular spokes, with each spoke representing one of the organizations. The data values on a spoke are proportional to the magnitude of two selected variables for the organization involved. Since the proposition of this research suggests a 'negative' relation between the task division and spatial data function variables and the level of spatial enablement (e.g. a high level of task division is expected to relate to a low level of spatial enablement), following conversion was done in order to reach comparability. The values of the task division and spatial data

function variables are converted as follows: high=5; medium/high=4; medium=3; medium/low=2; low=1. The values of the spatial enablement variable are converted as follows: high=1; medium/high=2; medium=3; medium/low=4; low=5. Due to this conversion, overlapping or parallel patterns indicate a strong relation between the two selected variables, while divergent patterns indicate a weak relation. It should be noted that in case of overlapping lines, the line that represents the spatial enablement variable is at the top, and thus covers the underlying line.

5.1 Task Division and Spatial Enablement

This section presents the findings on the possible relation between task division and spatial enablement (RQ1). In the Zoning Plans case, six organizations were selected: the cities of Genk, Kortrijk and Leuven, the provincial administrations of Limburg and West-Vlaanderen, and the Department of Spatial Planning, Housing and Immovable Heritage of the regional government (in short: RWO). It is apparent from the radar chart in Figure 2 that task division relates to spatial enablement. For three organizations (RWO, Limburg and Leuven) the match is nearly perfect, while the three other organizations show only a small deviation.

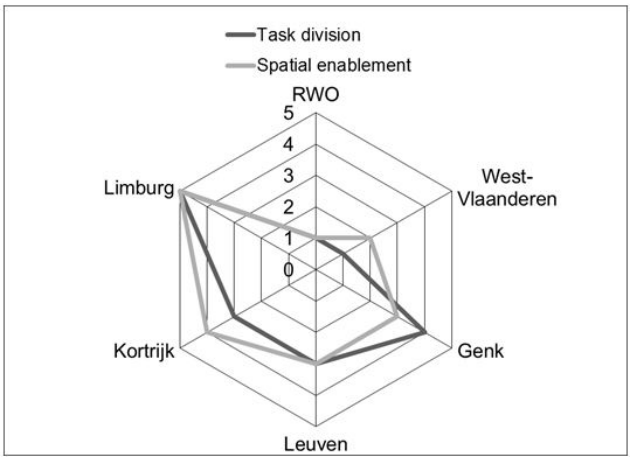

Figure 2. Radar chart comparison of task division and spatial enablement for the Zoning Plans case

Within the Traffic Accidents case, a selection of eight organizations was made: three local police zones (PZ Het Houtsche, PZ VLAS and PZ Leuven), the Federal Police, the national statistics agency (ADSEI), the regional Mobility and Public Works Department (MOW) and two provincial administrations (Vlaams-Brabant and West-Vlaanderen). As can be seen from the radar chart in Figure 3, task division relates to spatial enablement. West-Vlaanderen and PZ Het Houtsche combine the highest level of spatial enablement with a low level of task division. Vlaams-Brabant, PZ Leuven and PZ VLAS have a medium level of spatial enablement, and a (medium/) low level of task

division. And finally, Federal Police, ADSEI and MOW have a (medium/)low level of spatial enablement, and the highest (i.e. medium) level of task division.

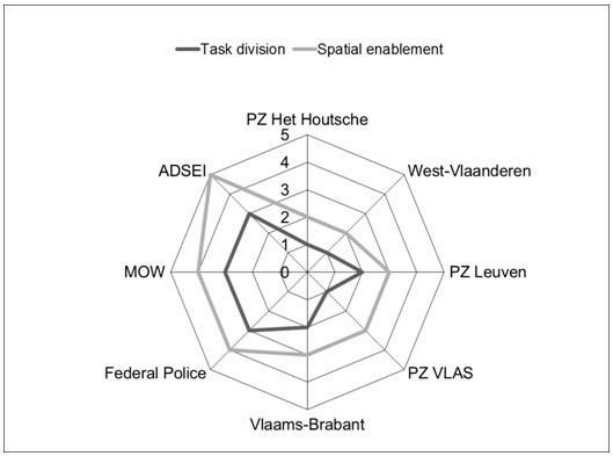

Figure 3. Radar chart comparison of task division and spatial enablement for the Traffic Accidents case

Six organizations were selected in the Flood Maps case: the city of Leuven, the Sint-Truiden Water Board, two provincial administrations (Limburg and Vlaams-Brabant), and two regional agencies (the Flemish Environment Agency and Flanders Hydraulics Research). Again, task division seems to relate to spatial enablement, as shown in Figure 4. The medium/low levels of task division of the Flemish Environment Agency, Limburg and the Sint-Truiden Water Board relate to a high level of spatial enablement, and the medium level of task division of Flanders Hydraulics Research and Vlaams-Brabant relates to a medium(/high) level of spatial enablement. Leuven combines a medium/high value for task division with a low value for spatial enablement.

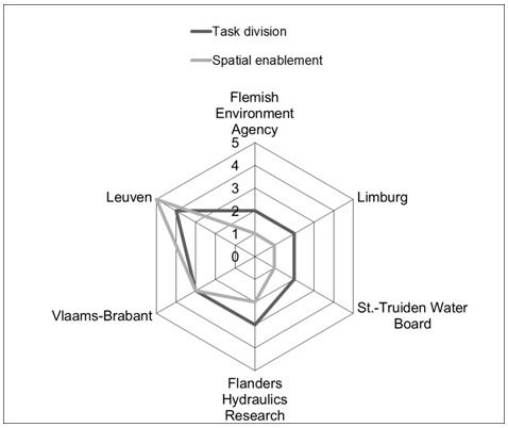

Figure 4. Radar chart comparison of task division and spatial enablement for the Flood Maps case

As for RQ1, the radar charts in Figures 2, 3 and 4 show that the patterns of the two variables are largely parallel, indicating a relation between both. As indicated, some organizations show a small deviation, but no outliers could be identified. The collected data suggest that the presence of an integrated (as opposed to fragmented) process could be related to a higher level of spatial enablement. This relation was found in the three cases.

5.2 Spatial Data Function and Spatial Enablement

This second section provides the results for allocation of the spatial data function (RQ2). It is apparent from the radar chart in Figure 5 that, in the Zoning Plans case, the spatial data function relates to spatial enablement. Again, the match is nearly perfect for RWO, Limburg and Leuven, while the three other organizations only show a small deviation.

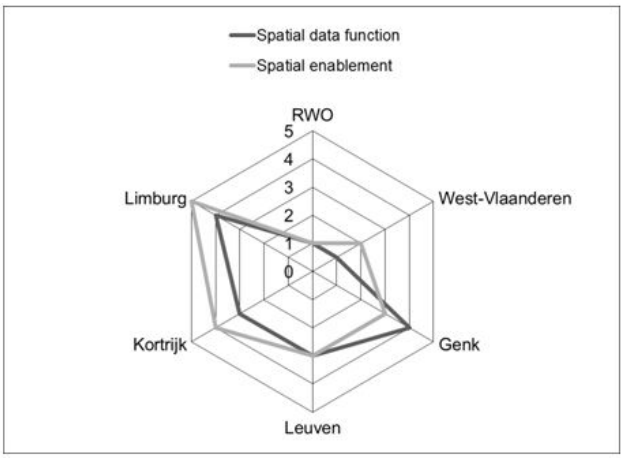

Figure 5. Radar chart comparison of spatial data function and spatial enablement for the Zoning Plans case

In the Traffic Accidents case, the general tendencies seem to prevail, as can be seen from Figure 6. In accordance with the proposition, West-Vlaanderen combines a medium/low spatial data function value with a medium/high level of spatial enablement, while conversely, Federal Police shows a medium/high spatial data function value and a medium/low level of spatial enablement. PZ Leuven and PZ VLAS combine medium values on both variables. However, the values of PZ Het Houtsche, Vlaams-Brabant and MOW slightly deviate from the expected pattern. The intersecting lines of the variables indicate that the variables show equivalent values, or values that only differ one unit (e.g. medium and medium/low). It should be noted that the line of the spatial data function variable in the radar chart shows a gap. ADSEI did not receive a value, since no spatial data were used in support of the traffic accidents process. Because no spatial data related activities could be identified, the variable allocation of the spatial data function was considered to be not applicable to ADSEI.

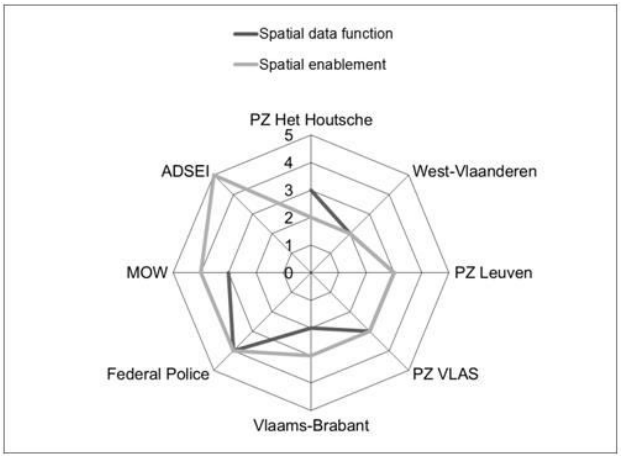

Figure 6. Radar chart comparison of spatial data function and spatial enablement for the Traffic Accidents case

A similar relation was found between spatial data function and spatial enablement in the Flood Maps case, in which a (medium/)low level of separation of the spatial data function could be linked to a (medium/)high level of spatial enablement. A medium level of separation of the spatial data function seems to lead to a medium or low level of spatial enablement. The radar chart in Figure 7 also shows largely coinciding patterns, although the strikingly low level of spatial enablement of Leuven may be somewhat surprising. This result might be due to the limited importance of the Flood Maps process for Leuven (where flooding risks are minimal as a result of its geographical location and of prior interventions).

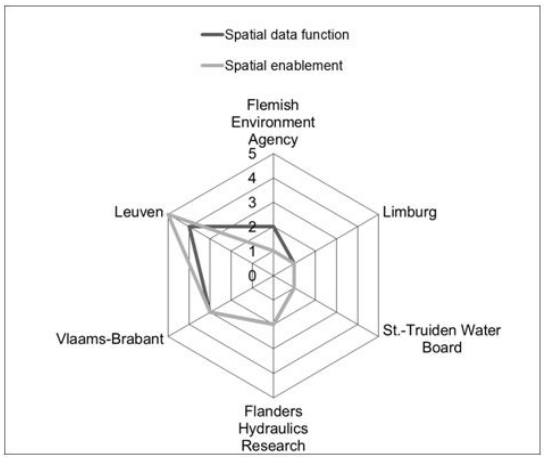

Figure 7. Radar chart comparison of spatial data function and spatial enablement for the Flood Maps case

The radar charts in Figures 5, 6 and 7 show that the pattern of spatial data allocation and the pattern of spatial enablement are largely similar. The collected data suggest a relation between spatial data function and spatial enablement.

6. Conclusion

The presence of an integrated (as opposed to fragmented) process with an embedded spatial data function seems to be related to a higher level of spatial enablement of the process. The similarity between task division and spatial data function may be explained in part by a certain overlap between both variables (since the allocation of the spatial data function is evidently also part of the overall task division). It may indicate that the allocation of the spatial data function in the process depends to a considerable extent on the way the various other functions are allocated. In the Traffic Accidents case this similarity seems to be somewhat less pronounced.

The chapter showed that process characteristics within organizations may impact on the process's spatial enablement. The mode of dividing tasks and the way of allocating spatial data related activities within a process can be regarded as organizational levers of spatial enablement. These findings are of relevance for SDI research and practice, since the uptake of spatial data use and sharing by organizations in support of their activities can be considered to further affect (impede or facilitate) spatial data sharing between organizations (Dessers, 2013).

Acknowledgements

The authors acknowledge the Agency for Innovation by Science and Technology in Flanders (IWT) for sponsoring the SPATIALIST research project of which this chapter is a result. Special thanks are owned to each of the interviewees, for the time they took to share their experiences.

References

Bekkers, V., (1998), *Grenzeloze Overheid. Over Informatisering en Grensveranderingen in het Openbaar Bestuur*, Alphen aan de Rijn: Samson [in Dutch].

Bolwijn, P. and Kumpe, T., (1990), Manufacturing in the 1990s – Productivity, flexibility and Innovation. *Long Range Planning*, 23(4): 44-57.

Callens, H., (2008), GIS in het Lokale Bestuur – Enkele Cijfers, *Presentatie op AGIV Trefdag 2008* [in Dutch].

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

Crompvoets, J., Dessers, E., Geudens, T., Janssen, K., Vancauwenberghe, G., Vandenbroucke, D. and Van hoogenbemt, M., (2009), *Het GDI-netwerk in Vlaanderen. Een Kwantitatieve Verkenning van het Gebruik en de Uitwisseling van Geodata in Vlaanderen*, Leuven: Spatialist, KU Leuven.

Crosswell, P.L., (2009), *The GIS Management Book*, Frankfort, Kentucky: Kessey Dewitt Publications in association with URISA.

Daft, R.L., (2001), *Organization Theory and Design*, Saint-Paul: South-Western College Pub.

de Vries, W., (2009), INSPIRE Obscured by Clouds of Inter-organizational Cooperation with geoICT? *International Journal of Spatial Data Infrastructures Research*, 4: 117-133.

Dessers, E., (2013), *Spatial Data Infrastructures at Work. Analyzing the Spatial Enablement of Public Sector Processes*, Leuven: Leuven University Press.

Dessers, E., Crompvoets, J., Janssen, K., Vancauwenberghe, G., Vandenbroucke, D., Vanhaverbeke, L. and Van Hootegem, G., (2012), A multi-disciplinary research framework for analysing SDI in the context of business processes. *International Journal of Spatial Data Infrastructures Research*, (7): 125-150.

Harvey, F. and Tulloch, D., (2006), Local-government data sharing: Evaluating the foundations of spatial data infrastructures. *International Journal of Geographic Information Science*, 20(7): 743-768.

In 't Veld, J., (1994), *Analyse van Organisatieproblemen. Een Toepassing van Denken in Systemen en Processen*, Houten: Stenfert Kroese [in Dutch].

Luhmann, N., (1984), *Soziale Systeme. Grundriss Einer Algemeinen Theorie*, Frankfurt: Suhrkampf [in German].

Masser, I., (2005), *GIS Worlds. Creating Spatial Data Infrastructures*, Redlands: ESRI Press.

Miles, M.B. and Huberman, M.A., (1994), *Qualitative Data Analysis. An expanded sourcebook*, Thousand Oaks / London / New Delhi: Sage Publications.

Reeve, D. and Petch, J., (1999), GIS, Organizations and People: A Socio-technical Approach. London: Taylor & Francis.

RWO (2008), *Richtlijn Voor de Digitale Uitwisseling van Ruimtelijke Uitvoeringsplannen*, Brussels: RWO [in Dutch].

Toonen, T.A.J., (2003), Administrative Reform: Analytics, in: Peters, G., and Pierre, J. (Eds). *Handbook of Public Administration*. London: Sage Publications. pp. 467-477.

van Amelsvoort, P., (2000), *The Design of Work and Organisation. The Modern Sociotechnical Systems Approach*, Vlijmen: ST-Groep.

Van Malderen, F., Macharis, C., Novikova, E. and Geudens, T., (2009), Accident Registration in Flanders: Methods, Problems and Weaknesses, in Macharis C. and Turcksin (Eds). *Proceedings of the BIVEC-GIBET Transport Research Day Part II*. Brussels: VUB Press, pp. 893-907.

Vandenbroucke, D., Dessers, E., Crompvoets, J., Bregt, A.B., and Van Orshoven J., (2013), Performance Measurement and Spatial Data Infrastructures: Application and Applicability Within the Context of Pork processes. *Computers, Environment and Urban Systems*. 38: 58-66.

Yin, R. K., (2003), *Case Study Research. Design and Methods*, London: Sage Publications.

CHAPTER 6

Towards An Online Self-Assessment Methodology for SDIs

Garfield Giff [1] and John Jackson [2]

[1] GeoInformation Management Consortium

[2] Advanced Mapping Systems, LLC

Garfield.giff@unb.ca; Johnj@advancedmapping.com

Abstract

The chapter investigates and analyzes current SDI assessment activities that focus on stakeholders' assessment and proposes an efficient cost-effective methodology for assessing SDIs from the stakeholders' perspective. Currently SDI organizations assess stakeholder performance based on a readiness or a generalized performance model from the SDI perspective. However, the performance of an SDI organization depends on the performance of stakeholder organizations which are motivated by business fundamentals to pursue enterprise GIS. The authors introduce the concept of a comprehensive integrated enterprise GIS/SDI assessment model from the stakeholders' perspective, and suggest an online approach as a cost-effective method. The chapter then moves on to describe the online assessment model and illustrates its suitability for both stakeholders and SDI assessments. A summary of the benefits of this methodology and areas requiring further development concludes the chapter.

KEYWORDS: SDI Assessment, SDI Stakeholders Assessment, Online Assessment

1. Introduction

In today's society where the emphasis is on accountability and performance, managers of public sector programs are required to demonstrate the performance of these programs, as well as, the impact they are having within their sectors. Similarly, in the *Spatial Data Infrastructure* (SDI) community the demand for reporting on the performance and benefits of SDIs is becoming a growing challenge for SDI managers, coordinators, practitioners and stakeholders. The measuring and monitoring of performance is vital to the successful management of an organization in that operational data analyzed over a time frame provides key insight into whether or not the organization is efficiently and effectively achieving its objectives of producing the desired services and products in a manner that has positive impacts on its customers, stakeholders, constituents, employees, and management. Measuring and monitoring progress is a proactive, best practices management technique used to understand business environment and operation trends that can assist managers in identifying issues before they become a problem. For example, a drop in training attendance may indicate that different training is needed or training is not meeting expectations.

In the case of the SDI, as programs mature and gain more political and funding support, so does the pressure to show progress/success to sponsors. While performance measures provide the facts for SDI sponsors to make wise investments and management decisions accordingly, measures also provide SDI coordinators with insight into the additional support needs of both the SDI (a collective of components) and its stakeholders.

Therefore, the challenge going forward for SDI stakeholders, coordinators, and practitioners is to develop functional (particularly cost-effective) frameworks to measure and monitor the performance of this complex – multiple component, multiple stakeholder and multi-dimensional – infrastructure known as the SDI. To date, the SDI community has taken the initiative in terms of developing and implementing a number of measurement frameworks to monitor the performances of SDIs across the globe. Examples of these can be seen in the works of Crompvoets *et al*., (2004); Delgado *et al*., (2005); Crompvoets *et al*., (2008); Delgado *et al*., (2008); Giff (2008), Giff and Crompvoets (2008); Vandenbroucke *et al*., (2008); Johnson and Kline (2009); KLD Consulting (2009); ADSIC (2010); Castelein and Manso Callejo (2010); Geodata (2010); Morera Amaya (2011); GeoConnections (2011); KU Leuven (2011); Toomanian *et al*., (2011) and the INSPIRE State of Play Reports. It should be noted that this is not a comprehensive list of measuring and monitoring activities throughout the global SDI community since a number of these activities have not been made public.

While the efforts by the SDI community must be commended, those same efforts ignore measuring and monitoring from the stakeholder's perspective. That is, the majority of current SDI performance frameworks do not consider the operational performance of each stakeholder. This is an important aspect of SDI performance since a chain is as strong as its weakest link. Therefore, it follows that an SDI is as strong as its weakest contributing stakeholder. Therefore, it is important to determine the

capabilities of the stakeholders to provide accurate and reliable spatial information and services, as well as, their capabilities to utilize the services offered by the SDI.

This chapter will investigate and analyze current SDI assessment activities that focus on stakeholders' assessment, as well as, propose an efficient cost-effective methodology for assessing SDIs from the stakeholders' perspective. This will be achieved through an introduction to SDI assessment, followed by a review of current SDI assessment activities. An introduction to the concept of SDI assessment from the stakeholders' perspective is then presented along with examples of the application of this method. The chapter then introduces to the readers a methodology for effectively assessing SDIs from the stakeholders' perspective. The chapter closes with a summary of the benefits of this methodology and areas requiring further development.

2. Current SDI Assessment Activities

Over the past ten years, SDI assessment activities have increased significantly, resulting in the development and application of a number of different methodologies for assessing SDIs. The book entitled *A Multi-View Framework to Assess Spatial Data Infrastructures* documents the most common methodologies in use and provides examples of their application. Additionally, a summary of recent SDI assessment activities across the globe can be viewed in (Giff and Crompvoets, 2013).

In general, SDI assessment methodologies may be classified into the two distinctive categories of Readiness and Performance Evaluation (Giff and Crompvoets, 2013). Within these two categories different methodologies are employed to measure and monitor performance. The methodology selected is usually based on the skills of the personnel involved, ease of use, cost, the required results, and the time it takes to perform the evaluation.

2.1 SDI Readiness Assessment Methodologies

A readiness assessment is a fact-gathering exercise carried out to determine the as-is status of a program. It provides insight into whether or not the governance structure, policies, tools and personnel are in place to achieve the stated objectives. That is, the program's readiness to perform the activities necessary to achieve the predetermined set of goals. In the case of an SDI, a readiness assessment provides information on whether or not the key components are in place to achieve the objectives of the SDI, as well as the level of completeness of their implementation (Giff and Crompvoets, 2013). This explains why early SDI evaluations were mainly readiness assessments.

2.2 SDI Performance Assessment Methodologies

An SDI performance assessment goes beyond identifying whether or not key components or desired components have been implemented but seek to determine if these selected components are performing and the level to which they are performing

to meet the objectives of the SDI. That is, an SDI performance assessment is carried out to determine if the SDI is achieving its objectives. This knowledge of whether or not an SDI is achieving the desired outputs, outcomes and impact is usually determining through the usage of performance indicators that are consistently measured and monitored.

Over the past six years the number of reported SDI performance assessments has increased significantly. This is in part due to demand from the SDI funders for more performance oriented information, as well as, an increase in the body of knowledge on measuring and monitoring the performance of SDIs (Giff and Crompvoets, 2013). Commendably, it was noticed that for the majority of the SDIs that had undergone performance assessment, the utilization of a number of the tools of the performance-based management framework was a key to their success (Giff and Crompvoets, 2013).

Summarizing, methodologies within the readiness category are most widely used for measuring and monitoring SDIs. This is mainly because they are inherently simple and cost-effective to implement and administer. The readiness assessment methodology is the most suited to the application of SDIs in their early implementation phase or to determine the capacity of an SDI to achieve predetermined targets. The main weakness of the readiness methodology, however, is that it does not provide sufficient information on: a) the level to which the defined targets are being achieved and b) the actual usage of the SDI or the usage of individual components. This can only be achieved through a performance assessment.

To date, the number of performance assessment methodologies employed in the SDI community is very limited. However, the need for actual SDI performance assessment is growing with the SDI communities in Canada, Abu Dhabi and Sweden leading the way. It is also worth mentioning that in recognition of the need for SDI assessment the INSPIRE group is incrementally transforming the state of play methodology into a performance-based assessment (Vandenbroucke *et al*., 2008) and (Giff and Crompvoets, 2013). Table 1 below provides a listing of the most documented SDI assessment methodologies and their application. This listing of SDI assessment methodologies – widely used by the SDI community – is classified according to the two categories of readiness and performance. It should be noted that the list is not comprehensive but is based on SDI assessment reports that have been made public.

Category	Methodology	Countries Assessed
Readiness	Clearinghouse Readiness	This methodology was applied to 67 countries across the globe. See Crompvoets and Bregt (2007) for details.
	Clearinghouse Suitability Index	Methodology was applied to 83 Countries across the globe. See Crompvoets and Bregt (2008) for details.
	The SDI Readiness Model	Methodology applied to 27 countries worldwide including 17 from the Americas. Most comprehensive application was to the country of Cuba. See Delgado *et al.*, (2008) for details.
	INSPIRE State of Play	Applied to the member countries of the European Union and Croatia, Macedonia, Turkey, Iceland, Norway, Switzerland, and Liechtenstein. Vandenbroucke *et al.*, (2008) provides more details on this methodology.
Performance Assessment	The GeoConnections Framework	Methodology applied to the assessment of the Canadian Geospatial Data Infrastructure. See GeoConnections (2013), and Giff and Crompvoets (2013) for more details.
	The GeoMaturity Model	This multi-level assessment framework was applied the SDI of the Emirate of Abu Dhabi. For more details see ADSIC (2010) and Giff and Crompvoets (2013).
	Balance Scorecard	Applied to the assessment of the Swedish SDI**.** For more information on the application of the BSC to the assessment of the Swedish SDI see Toomanian *et al.*, (2011), Geodata (2010) and Giff and Crompvoets (2013).
	Multi-view Framework	Applied to the assessment of the Netherlands' SDI in 2008. More details of this application can be found in Grus *et al.*, (2010) and Castelein and Manso Callejo (2010). Additional information on the multi-view framework can be found in (Crompovets *et al.*, 2008).

Table 1. A Listing of the most widely used SDI assessment methodologies

3. The Need for Stakeholders' Performance Information in SDI Assessment

SDI assessment requires the input and corporation of the multiple stakeholders involved in the different levels or aspects of the SDI. A performance assessment of this multi-faceted and multi-dimensional infrastructure requires SDI coordinators/assessment team to investigate the performance of the different components of the SDI and aggregate these results to represent performance information at the various levels of the SDI.

Currently, the methodologies used for collecting SDI performance information are usually cumbersome, time consuming and represent a one-off process; with the more common techniques being questionnaires, interviews and literature reviews. Generally, SDI coordinators, managers and stakeholders find these techniques to be tedious and time consuming. In addition, for the SDI coordinators/assessment team, communicating effectively and receiving the necessary cooperation from the different stakeholders can also be challenging. Therefore, if SDI assessments are to evolve into performance-based assessments and provide beneficial results then the performance information collection processes must be simplified and offer greater incentives to both stakeholders and the SDI coordinators/assessment team. This can be achieved through the employment of online self-assessment programs at the different levels of the SDI (Figure 1) with the results of each level capable of being 'rolled up' or summarized to the next higher (parent) level thus contributing to a holistic set of performance information.

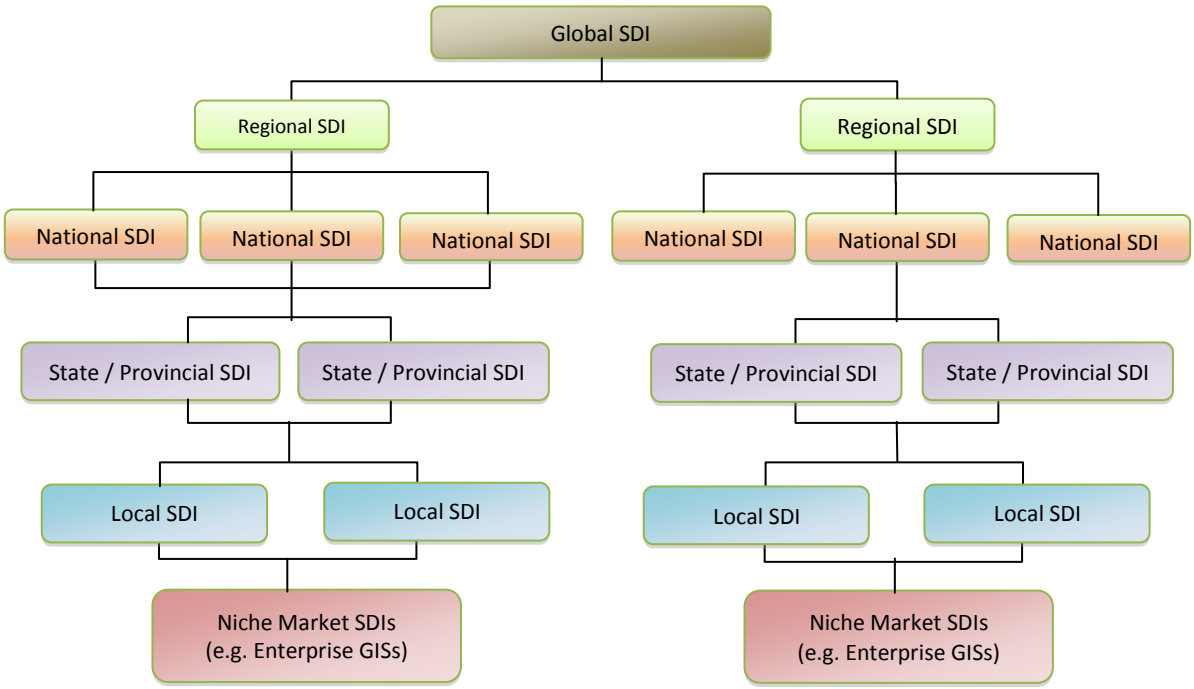

Figure 1. SDI hierarchical structure (from Giff (2005))

From Figure 1 it is clear that an SDI belongs to a hierarchical structure ranging for the Global SDI as the eldest parent to niche market SDIs (e.g. enterprise GISs) as the youngest child (See Rajabifard (2002) for more information on the hierarchical structure of SDIs). Based on the hierarchical concept of an SDI as defined by Rajabifard *et al*., (2000), it can be concluded that the performance of an SDI will be influenced by its children. Therefore, it is important when assessing an SDI to include performance information of its children (i.e. lower-tier stakeholders) in the assessment. However, the collection of this type of information is a big challenge to SDI coordinators. This is mainly due to the fact that stakeholders find that assessing their geospatial activities a costly and time-consuming activity. If the cost of assessing geospatial programs becomes more affordable then it is expected that stakeholders' performance information would become more readily available to SDI coordinators. The authors are of the opinion that an *Online Self-Assessment* tool for stakeholders is one method of significantly reducing the cost, and time when performing a stakeholder assessment.

4. An Online Self-Assessment Tool for SDIs

An *Online SDI Self-Assessment Program* is a set of interactive online e-services designed to enable SDI stakeholders and coordinators to store and maintain various performance information – based on the selection of performance indicators specific to their organization. That is, with the Online SDI Self-Assessment Tool an SDI

manager/coordinator or a stakeholder can create a customized set of measures – from the indicators available – that when captured over time would provide performance information and useful trend analysis. This type of information can help managers identify issues and steps to further review or improve a process and/or service offered by the SDI to the stakeholder or vice-versa.

4.1 The Global Application of the Tool

The Online SDI Self-Assessment Tool facilitates the assessment of an organization's geospatial activities at predetermined time intervals (e.g. weekly, monthly, annually, or bi-annually). Performance information collected is stored locally, but can be shared with parent organizations (Figure 1) to allow for a summary of the assessment results or selected details of the assessment to be analyzed by the parent SDIs, as well as, other stakeholders either connected to the parent SDI or not. This is in support of the roll-up concept where relevant performance information of the child entity is passed-on to be part of the assessment of the parent SDI. In addition, the tool can generate performance reports that classify the geospatial status of a stakeholder entity based on a standardized system which is adopted by the masses. This enables the comparison of stakeholders and SDI thus, providing the SDI community with an instrument for identifying and supporting areas within stakeholders' entities that require additional support, as well as, the identification of best practices within the community.

4.2 The Models of the Online SDI Self-Assessment Tool

The Online SDI Self-Assessment Tool consists of two distinct models. The first model is designed to assess the geospatial performance of the lower-tier stakeholders of the SDI (i.e. GISs forming part of the SDI). The second model supports the assessment of the different levels of SDIs (i.e. state or local SDIs). Two distinctive models are required because of the differences in the organizational structures of the entities participating in the different levels of an SDI and the hierarchical relationship of one SDI to many stakeholders. This is an important aspect of the tool since an assessment framework is most effective when it replicates the natural structure of the organization to be assessed (Giff and Crompvoets, 2013). Therefore, in order for the assessment tool to best replicate the organizational structure of the SDI and its stakeholders the authors decided to use two different models.

4.2.1 Stakeholders' Model

The main purpose of the stakeholders' model is to represent the assessment framework which consists of the indicators, the indicator responses, the ranking of the indicator responses and the maturity levels. The authors based the model on the Capacity Maturity Model (CMM) developed by the Carnegie Mellon University (Watts, 1989; Paulk *et al*., 1994). In addition, works by GPC GIS and Abu Dhabi Systems and Information Centre (ADSIC), Even Keel Strategies, URISA, and Makela that focused on

the application of the CMM to the assessment of GISs were also investigated (ADSIC, 2010; Mangan, 2009; URISA, 2013; Makela, 2012). In addition, to the work done on previous maturity models the authors also drew on practical experiences gain in the development and implementation of the ADSIC maturity model. This model was used to assess more than 30 Abu Dhabi SDI stakeholder entities. In summary, the Online Self-Assessment tool's stakeholder model was built on the solid foundation established by the above organizations and the practical and theoretical knowledge of the authors.

In keeping with the concept that the assessment of the geospatial activities within an entity is most efficient and effective when the assessment framework follows the natural organizational structure of the entity, the stakeholders' model consisted of the key aspects of a geospatial organization that will influence its performance. These key areas are the entity's organizational structure, its information management activities, the IT and GIS technology status, the internal usage of geospatial information to enhance the activities of the organization, and the enhancement of customer service through the application of geospatial information (Figure 2).

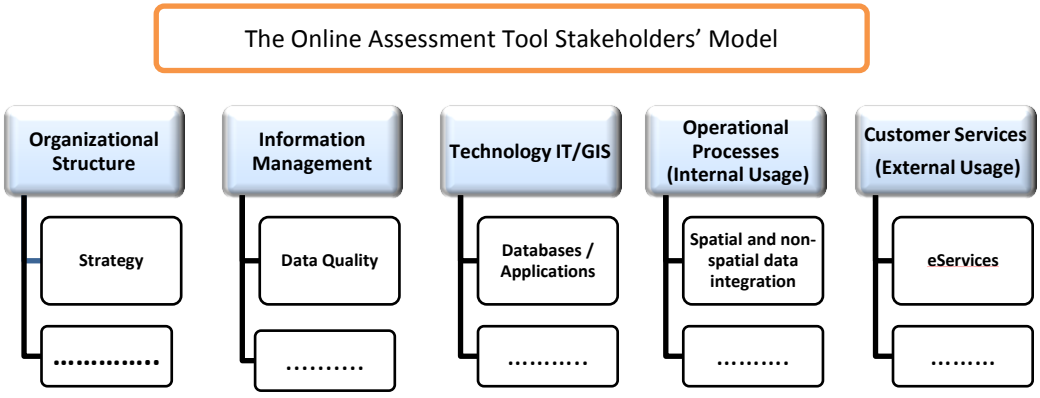

Figure 2. Conceptual level of the stakeholders' model

Figure 2 illustrates the higher levels of the stakeholders' model. However, the complete model is a much more comprehensive architecture consisting of 5-7 indices for each key areas of the organization and with each index being further sub-divided into sub-indices to create a more comprehensive model. Currently, the sub-indices collectively contained over 200+ performance indicators with each indicator having five preset answers (indicator responses). Examples of the indices of the six key areas of a stakeholder entity can be seen in Figure 3.

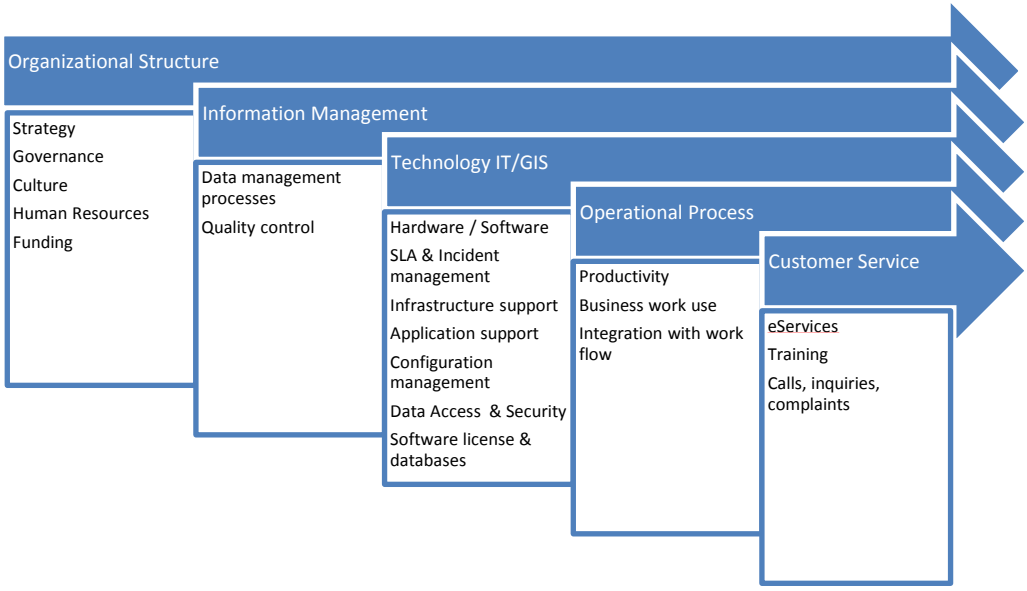

Figure 3. Examples of the indices for key performance areas of a geospatial (stakeholder) organization

Similar to the CMM and the other geospatial maturity models (GPC GIS and Abu Dhabi System Information Center, Even Keel Strategies, URISA, and Makela) investigated by the authors, the online assessment tool consist of a six-level ranking system used to classify the performance status of the organization (Figure 4). The ranking system is based on research and practical application of performance frameworks in the software and GIS sectors. The maturity levels tend to reflect the natural growth path an organization undergoes before evolving into totally utilization of the software necessary to support the optimization of their business processes. These maturity levels, which depict the level at which an organization is developing and utilizing their geospatial capabilities, are as follows:

- *Level 1*: The organization is investigating the benefits of geospatial applications to its activities and the services it offers. That is, managers are aware of the benefits of implementing a GIS in the organization and have discussed with professionals the possibilities of enhancing the business processes with GIS.
- *Level 2*: Geospatial information is applied to solve case-specific problems within the organization. That is, Geospatial information is used in one-off situations and when the specific task is completed the information is not used again.

- *Level 3*: Geospatial information is used by different sections of the organization in performing their daily activities. However, there is no coordination amongst these divisions resulting in duplication.
- *Level 4*: The usage of Geospatial applications throughout the organization is coordinated and managed by a designated unit. However, not all applicable processes are spatially enhanced.
- *Level 5*: All applicable processes are spatially enhanced and GIS is integrated with other systems within the organization.
- *Level 6*: Geospatial *information* usage is optimized and Geospatial information is used by high-level managers in decision-making.

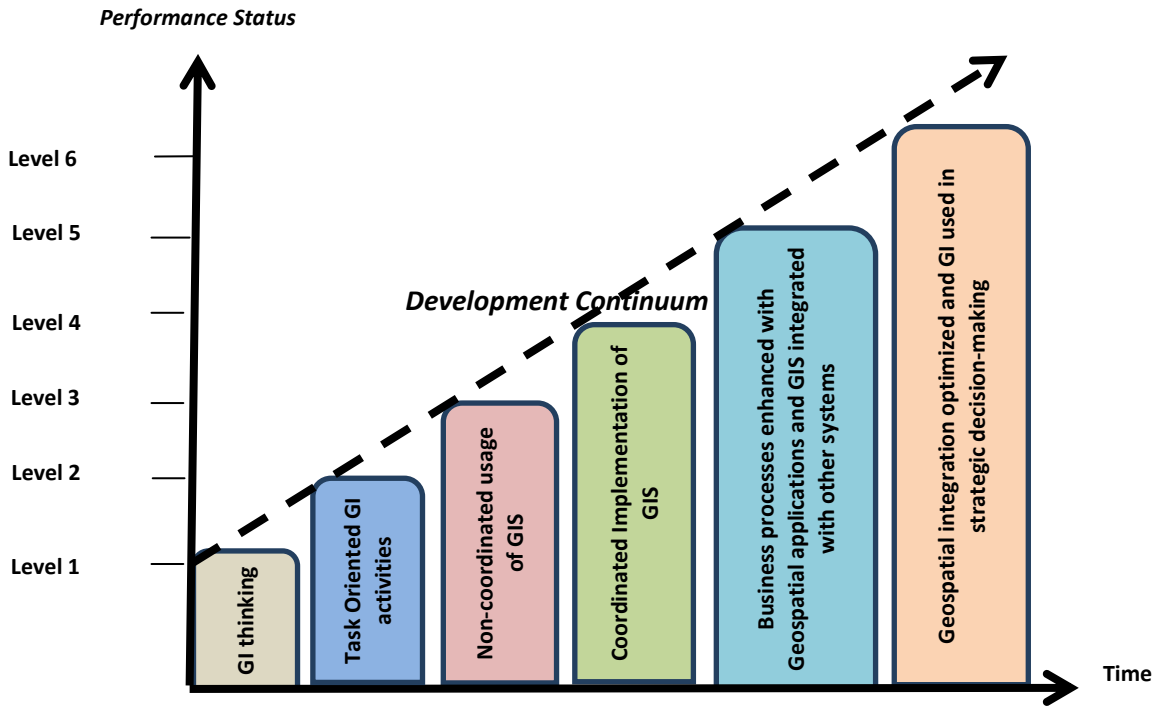

Figure 4. Expected path a stakeholder organization follows through the evolution of GI Usage

Figure 4 defines the evolutionary path an organization usually follows to achieve optimal usage (i.e. Enterprise GIS) of Geospatial Information. The online assessment tool can be used to determine the point in the evolution continuum a stakeholder organization is at – at any given time. This is achieved through the usage of the predetermined, ranked indicator responses. The indictor responses are ranked from 1 to 6 corresponding with the levels along the path to optimization. That is, each response will classify the index or sub-index within one of the levels illustrated in Figure 4.

For example, the indicator "GIS strategic plan in place" will have six predetermined responses which correspond to one of the levels in Figure 4. This will facilitate classifying the organization's GIS strategic planning activities into one of the maturity levels identified in Figure 4. The responses are as follows:

- Organization researching GIS strategic plans: **Level 1**
- Task-orientated GI activities are selected based on the organization's priority: **Level 2**
- Individual sections have performed needs assessment for GI usage: **Level 3**
- The coordinating group have prepared a GIS strategic plan for the organization: **Level 4**
- The drafted GIS strategic plan is circulated across the organization for feedback, additional input, and updating: **Level 5**
- The GIS strategic plan is aligned with and is also a component of the organization's global strategic plan: **Level 6**

Another example of an indicator within the tool and its predetermined responses are as follows:
The indicator "The application of GIS Software licenses through the organization." The responses are as follows:

- Organization reviewed GIS software on the market and have invited vendors to perform demonstration: **Level 1**
- Organization has completed a plan to acquire software for its projected needs but currently using demo software: **Level 2**
- Selected departments have standalone software: **Level 3**
- Detailed plans and guidelines on how software should be used through the organization completed. Pilot project on the way that facilitate access to organization wide software: **Level 4**
- GIS software integrated with key applications across the organization: **Level 5**
- GIS software integrated with all possible applications across the organization: **Level 6**

The ranked indicator responses are also supported by a notes section which allows the users to provide further supportive information which can be analyzed by the performance team. Selected or summarized results from the assessment can then be exported into the SDI model to be used in the assessment of the SDI.

The average value of the results of the indicators of each category (see Figure 3) is used to determine the maturity level the category. For example, the maturity level of the category Information Management will be the average value of the indicators in that category. Similarly, the average value of the five categories identified in Figure 2 and 3 will reflect the maturity level of the organization. That is the Maturity level of an organization is:

The sum of the Maturity levels of the five categories/5

(organizational structure + information management + IT and GIS technology + internal usage of geospatial information + customer service enhancement through the application of geospatial information)/5

4.2.2 The SDI Model

The second component of the online assessment tool – the SDI model – is designed to capture performance information – inclusive of stakeholders' information – with respect to the SDI. The SDI model was built on work done by GPC GIS and ADSIC, GeoConnections, the National States Geographic Information Council, the State of Georgia GIS Committee, INSEAD, MetroGIS, and the Multi-view Frame Work Team (ADSIC, 2010; GeoConnections, 2013; NSGIC, 2011; GISCC, 2008; INSEAD, 2004; KLD Consulting, 2009; and Crompvoets *et al*., 2008). Similar to the stakeholders' model the SDI model is based on the organizational structure of an SDI and encompasses the more advanced performance assessment methodology proposed by Giff and Crompvoets (2013) in their paper Measuring and Monitoring SDIs.

The SDI model was designed to assess the performance of an SDI in two folds. Firstly, the assessment of the readiness of the SDI to achieve its objectives and secondly, the performance of the SDI based on its objectives and actual usage by both stakeholders and non-stakeholders (Figure 5). Figure 5 represents a graphic presentation of the SDI model. From the Figure eight key areas (organizational structure, funding and geo-legal environment, capacity building, infrastructures, data and information environment, standards, objectives and usage) of an SDI can be identified. These key areas are further subdivided into a number of sub-indices and over 100 indicators and indicator responses were designed for these sub-indices. Similar to the stakeholders' model the indicator responses of the SDI model are ranked with respect to the different performance levels of an SDI. That is, each response will classify the performance of a sub-index within the defined performance classification levels of an SDI.

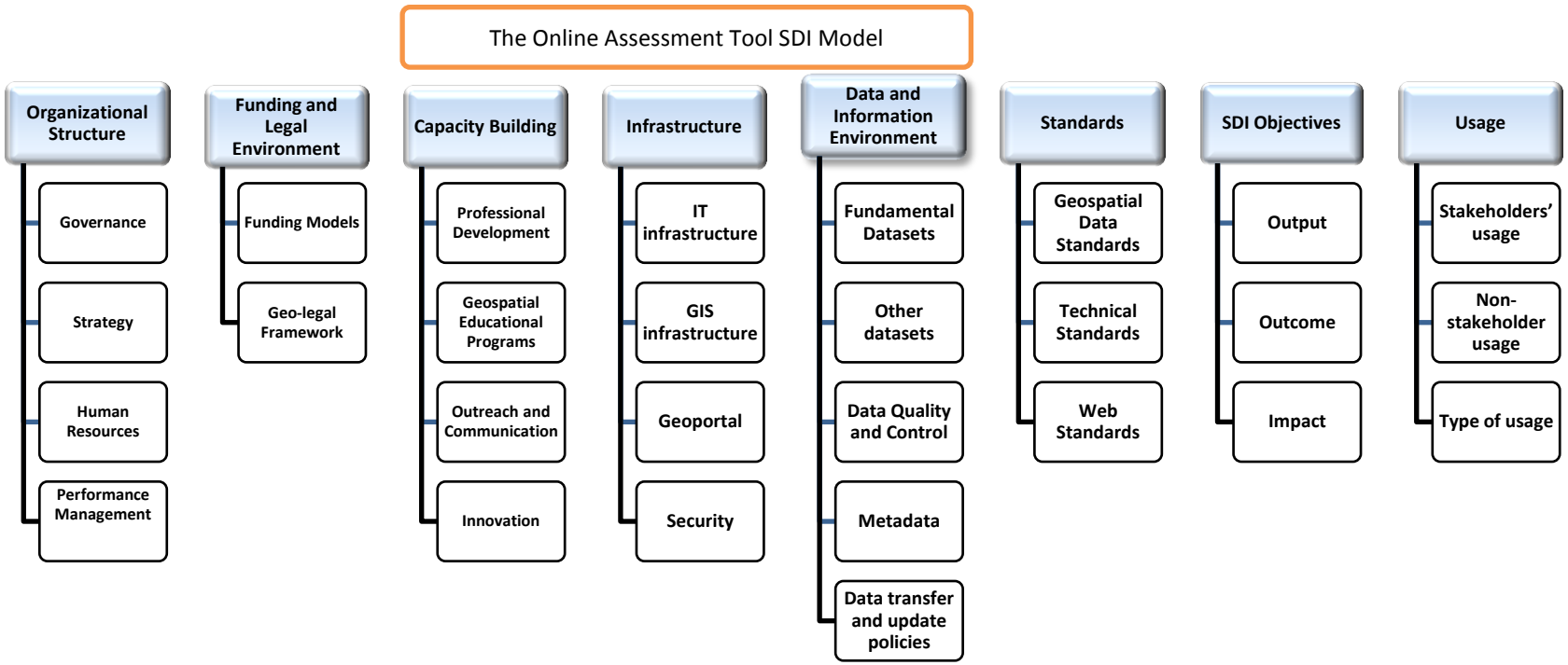

Figure 5. SDI Model for the online assessment tool

The SDI levels used for the online assessment tool are an adaptation of the levels used by GPC GIS and ADSIC in the assessment of the Abu Dhabi Spatial Data Infrastructure. Concise definitions of the adapted levels are as follows:

- *SDI Level 1:* Exploration of the benefits to be gained from formalizing current geospatial ad hoc sharing arrangements carried out by members of the geospatial community. SDI practicality is researched and time is invested in learning more about the organizational and technical issues associated with SDI implementation. SDI champions are born and coordinating body evolves to plot the way forward.
- *SDI Level 2:* Sufficient interest in geospatial sharing has been generated resulting in the start of the innovation-decision process. SDI initiatives are undertaken usually in the formation of working groups to develop policies to facilitate data sharing and the definition of fundamental datasets.
- *SDI Level 3*: Phased implementation of key components (e.g. data, policies and technology) of the SDI. The implementation although driven by the stakeholders is steered by the coordinating body and the outcomes of the working groups. Standards and custodianship of fundamental datasets are clearly defined.
- *SDI Level 4*: SDI accepted by the geospatial community and its significance is growing within the other sectors. An increase number of stakeholders participate in the implementation of the SDI. The SDI is viewed as competent environment for discovering and sharing geospatial information.
- *SDI Level 5*: SDI viewed as a key resource by the stakeholders and the wider geospatial community. Advance applications and technology are implemented to improve viewing, accessing and downloading. The SDI evolves from an information and technology driven facilitator to a service facilitator to meet the needs of the wider society.
- *SDI Level 6*: SDI is matured with seamlessly integration into the activities of government, business and citizens. That is, the SDI facilitates an environment where geospatial information and services are used to by citizens in their everyday activities, by governments in providing services to citizens and by corporate society in their business processes.

As with the stakeholders' model the predefined responses to the indicators are ranked and will place an index or sub-index within one of the above levels. The average scores of the different components are used to determine the overall maturity status of the SDI. The results of each component can also be weighted to emphasize the importance of each component in the final performance status of the SDI.

5. Benefits of Online Self-Assessments Tool

SDI coordinators have a responsibility to encourage and incentivize stakeholders in the development of their geospatial activities. This is because the more efficient and effective the stakeholders are, the more efficient and effective the SDI will be in

delivering geospatial information and services to the community. The Online Self-Assessment program is a smart tool available to both SDI coordinators and geospatial managers (stakeholders) for efficiently measuring and monitoring the performance of their geospatial programs. Therefore this is one tool an SDI coordinator can use to assist in the identification of areas within the SDI, as well as stakeholders' entities that may need their support. The Online Self-Assessment tool can also be used as an incentive to stakeholders in that it provides them with the means of efficiently measuring and monitoring performance.

The key benefit of the Online Self-Assessment tool lays in its capability to allow the users to perform efficient, low cost, self-assessment of their programs. However, there are also other benefits which can be viewed from the perspective of the SDI coordinators/program managers and the stakeholders. These benefits are summarized in Table 1 below.

User	Benefits
SDI Coordinators	1. An on-demand method of measuring and monitoring the performance of individual components of the SDI or the SDI as a collection of components. 2. Customized presentation of results: SDI coordinators can view simplified results in a dashboard or view more detailed results of performance at both the SDI and stakeholder levels. 3. The Reporting Capabilities: Online templates provide various view format options for real time, easily read dashboard reports. Sponsors, executive managers and committee members can review program status from an ongoing performance perspective. Trends can also form discussion agenda for oversight committees. Reports provide a proactive means for management to monitor progress of an often not well-understood program. 4. Standardization: with standardized indicators and levels SDI coordinators can benchmark their performance against other SDIs, as well as benefit from best practices. 5. Cost savings: the online tool is expected to generate significant cost savings since it is less labor intensive and does not require the employment of performance management experts.
SDI Stakeholders	1. An indication of the ability to realize the potential benefits from investment in geospatial activities (Babinski, 2010). 2. Enterprise GIS managers, program sponsors, and executive management can analyze various performance data trends for agency specific purposes. 3. Provides a standard that can be used for benchmarking. That is, stakeholders can compare their performance trends to other similar organizations to gain different perspectives and learn what is working and what is not. Benchmarking can also be done against stipulated

	targets. 4. Support Best Practice Management System (i.e. A standardized list of indicators based on current formats e.g. the balanced score card, can easily fit into existing or planned corporate performance management systems and/or initiatives. 5. Customization: Currently, stakeholders can choose from over 200 indicators to measure and monitor performance. All or any combination can be used, as well as, a stakeholder can add indicators thus, enabling the creation of a 'customized' list of indicators specific to that stakeholder's objectives. 6. Item 2 above also applies for this category.

Table 1. Summary of the benefits of the online self-assessment tool by users

6. The Way Forward

The Online Self-Assessment tool has great potential in assessing the performance of an SDI and its stakeholders. This is evident from the demand in the SDI community for more simplified, cost-effective and less tedious methods of measuring and monitoring performance (KLD Consulting, 2009; and Giff and Crompvoets 2013). The potential of the tool is also highlighted in the benefits to be gained from its employment listed in Table 1. In addition the tool provides the geospatial community with a standardized set of performance indicators, as well as, a standard format for grading and reporting performance in the geospatial community. Of significance to SDI coordinators is the capability of the Online Self-Assessment tool to provide them with performance information on their stakeholders. This type of information is important as it provides SDI coordinators with a better understanding of the needs and interest of the stakeholders (KLD Consulting, 2009). It should be noted that the Online Self-Assessment tool is a dynamic tool that facilitates the addition of new indicators, as well as, the retirement of old indicators based on the feedback of the users.

The first version of the prototype of the Online Self-Assessment tool demonstrates great potential in the assessment of the Geospatial sector. However, there are still some shortcomings with the tool that needs to be addressed. The authors view these shortcomings as opportunities to improve on the performance of the next generation of the tool. In moving forward, these shortcomings will be systematically identified and addressed through stakeholders' workshops, prototype iteration, and pilot testing the prototype. That is, feedback from stakeholders' workshops and the implementation of the prototype, as well as, feedback from this chapter will be used to improve the tool.

The Online Self-Assessment tool is expected to evolve and become more efficient and precise with each iteration. The authors' long-term vision for the Online Self-Assessment tool is to have the tool integrated into the dashboards of SDI program funders, SDI coordinators, SDI program managers, the managers of stakeholder

entities and GIS managers providing different levels of performance information to support informed decision-making with respect to the future of Geospatial investment.

7. Conclusion

The chapter presented and analyzed the concept of the application of an online tool to be employed by the Geospatial community to self-assess the performance of their Geospatial activities and programs. In the analysis, the chapter justified the need for such a tool, reviewed the models employed by the tool and reviewed the benefits the Geospatial community will derived from the usage of the tool.

From the information presented in the chapter, it is clear that more work needs to be done on the development and prototyping of the Online Self-Assessment tool. However, the authors are of the opinion that the presentation of the tools' models and its possible application and benefits to the community will generate quality discussions and feedback that can be used to improve the tool in moving forward.

In concluding, the application of an Online Self-Assessment tool can prove a valuable contribution to SDI assessment; since a tool of this nature can produce cost-effective performance information capable of identify the areas within Geospatial programs that are performing and those that require additional support.

Acknowledgements

The authors would like to acknowledge the GeoConnections' Team of Paula McLeod, Simon Riopel and Rhian Evans for their contribution to the chapter, as well as their contributions to the advancement of SDI assessment methodologies.

References

ADSIC – Abu Dhabi Systems and Information Centre (2010), Abu Dhabi's GeoMaturity Assessment Framework. An ADSIC report, Abu Dhabi, UAE.

Babinski G., (2010), Managing GIS. ArcNews Winter 2010/2011. Available from: http://www.esri.com/news/arcnews/winter1011articles/urisa-proposes.html

Castelein, W. and Manso-Callejo, M., (2010), Monitoring of Spatial Data Infrastructures. I Jornadas Ibéricas de Infraestructuras de Datos Espaciales. Lisboa, 27-29 Octubre 2010 http://www2.topografia.upm.es/pdi/m.manso/publicaciones/JIDEE2010-Monitoring.pdf

Crompvoets, J., Bregt, A., Rajabifard, A. and Williamson, I.P., (2004), Assessing the Worldwide Developments of National Spatial Data Clearinghouses. *International Journal of Geographical Information Science*, 18(7): 665-689.

Crompvoets, J. and Bregt, A.K., (2007), National Spatial Data Clearinghouses, 2000-2005. In: Onsrud, H.J., (Ed.), Research and Theory in Advancing Spatial Data Infrastructure Concepts. ESRI Press, Redlands, pp. 141-154.

Crompvoets, J. and Bregt, A., (2008), Clearinghouse Suitability Index. In: A Multi-view Framework to Assess Spatial Data Infrastructures, ED: Crompvoets, J., Rajabifard, A., van Loenen, B., and Delgado Fernandez, T., Melbourne, Australia: The Melbourne University Press, pp. 135-144.

Crompvoets, J., Rajabifard, A., van Loenen, B. and Delgado Fernandez, T., (2008), A Multi-view Framework to Assess Spatial Data Infrastructures, Melbourne, Australia: The Melbourne University Press.

Delgado-Fernandez, T., Lance, K., Buck, M. and Onsrud, H. J., (2005), Assessing SDI Readiness Index. In Proceedings of the Eighth International Conference on Global Spatial Data Infrastructure, April 2005, Cairo, Egypt.

Delgado-Fernandez, T., Delgado-Fernandez, M. and Andrade, R., (2008), Spatial Data Infrastructure Readiness Model and its Worldwide Application. In: A Multi-view Framework to Assess Spatial Data Infrastructures, ED: Crompvoets, J., Rajabifard, A., van Loenen, B., and Delgado Fernandez, T., Melbourne, Australia: The Melbourne University Press, pp. 117-134.

GeoConnections (2011), GeoConnections Annual Report 2009-2010. A GeoConnections report, Ottawa, Ontario, Canada.

GeoConnections (2013), Canadian Geospatial Data Infrastructure Performance Project – Overview. A GeoConnections Report, Ottawa, Ontario, Canada. Available from: ftp://ftp2.cits.rncan.gc.ca/pub/geott/ess_pubs/292/292421/cgdi_ip_30e.pdf

Geodata (2010), Balanserat Styrkort Metodbeskrivning för hur resultat inom samordningsrollen följs upp. A Geodata Report, Gävle, Sweden.

GISCC- Georgia Geographic Information Systems Coordinating Committee (2008), The 2007- 2008 Georgia Geospatial Maturity Assessment. Available from: www.fgdc.gov/grants/2008CAP/Reports/008-08-3-GAStrategicPlan2009.pdf

Giff G.A., (2005), Conceptual Funding Models For Spatial Data Infrastructure Implementation. Ph.D. Thesis, University of New Brunswick, New Brunswick, Canada. Available from: http://dspace.hil.unb.ca:8080/xmlui/handle/1882/44359

Giff, G.A., (2008), A Framework for Designing Performance Indicators for Spatial Data Infrastructure Assessment In: A Multi-view Framework to Assess Spatial Data Infrastructures, ED: Crompvoets, J.; Rajabifard, A.; Loenen, B. van; Delgado Fernandez, T., Melbourne, Australia: The Melbourne University Press, pp. 211-230.

Giff, G.A. and Crompvoets, J., (2008), Performance Indicators a Tool to Support Spatial Data Infrastructure Assessment. *Computers, Environment and Urban Systems*, 32: 365-376.

Giff, G.A. and Crompvoets, J., (2013), Measuring and Monitoring Spatial Data Infrastructures. In Spatial Data Infrastructure Manual for the Americas, ED: Natural Resources Canada, Ottawa, Ontario, Canada. Available from: http://www.snit.cl/CPIDEA/index.php/component/jdownloads/finish/48-10-reunion-cp-idea/283-spatial-data-infrastructure-sdi-manual-for-the-americas?Itemid=0

Grus, L., Castelein W.T., Crompvoets, J., Overduin, T., Van Loenen, B., Van Groenestijn, A., Rajabifard, A. and Bregt, A., (2010), An assessment view to evaluate if Spatial Data Infrastructures meet their goals: In Assessment of Spatial Data infrastructures, PhD thesis Wageningen University, Wageningen, The Netherlands.

INSEAD (2004), The Networked Readiness Index 2003–2004: Overview and Analysis Framework. Available from: www.weforum.org/pdf/Gcr/GITR_2003_2004/ Framework_Chapter.pdf

Johnson, R. and Kline, C., (2009), MetroGIS 2008 Performance Measurement Report. A MetroGIS report, St. Paul, MN, USA.

KLD Consulting (2009), MetroGIS Performance Measurement Plan Update. A MetroGIS report, St. Paul, MN, USA.

KU Leuven (2011), AD-SDI Assessment Report. Report prepared for Abu Dhabi Systems and Information Centre, Abu Dhabi.

NSGIC – National States Geographic Information Council (2011), State Government Geospatial Maturity Assessment. Available from: www.fgdc.gov/participation/ coordination-group/meeting-minutes/2012/april/full-gma.pdf/view

Makela, J., (2012), Model for Assessing GIS Maturity of an Organization. In Spatially Enabling Government, Industry and Citizens, ED: Rajabifard, A. and Coleman, D., Needham, Massachusetts, USA. GSDI Association Press, pp. 143-165.

Mangan, M., (2009), Introducing a Maturity Model for Enterprise GIS. An Even Keel Strategies Whitepaper, Chicago IL, USA. Available from www.w4sight.com/uploads/1/4/3/8/14386328/w4paper_gismaturity.pdf.

Morera Amaya, C., (2011), Modelo Evaluación Costo – Beneficio de la Infraestructura Colombiana de Datos Espaciales – ICDE. Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de: Magister en Geomática, Universidad Nacional de Colombia Facultad de Agronomía, Bogotá, Colombia.

Paulk, M. C., Weber, C. V., Curtis, B. and Chrissis, M., (1994), The Capability Maturity Model: Guidelines for Improving the Software Process. Reading, Massachusetts: Addison-Wesley Professional.

Rajabifard, A. Escobar, F. and Williamson, I.P., (2000), Hierarchical Spatial Reasoning Applied to Spatial Data Infrastructures. *Cartography Journal*, 29(2).

Toomanian, A., Mansourian, A., Harrie, L. and Rydén, A., (2011), Using Balanced Scorecard for Evaluation of Spatial Data Infrastructures: A Swedish Case Study in Accordance with INSPIRE. *International Journal of Spatial Data Infrastructures Research*, 6: 311-343.

URISA – Urban and Regional Information Systems Association (2013), GIS Capability Maturity Model. Available from: www.urisa.org/files/GISCMM-PublicReviewDraft(201304rev).pdf

Vandenbroucke, D., Janssen, K. and Van Orshoven, J., (2008), INSPIRE State of Play: Generic approach to assess the status of NSDIs. In: A Multi-view Framework to Assess Spatial Data Infrastructures, ED: Crompvoets, J., Rajabifard, A., van Loenen, B., Delgado Fernandez, T., Melbourne, Australia: The Melbourne University Press, pp. 145-172.

Watts, H., (1989), Managing the Software Process Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA ©1989 ISBN:0-201-18095-2.

PART TWO

EMPOWERING COMMUNITIES – PARTICIPATORY APPLICATIONS FOR SDIS

CHAPTER 7

Public Participatory GIS, Spatial Data Infrastructure, and Citizen-Inclusive Collaborative Governance

Michael Sutherland[1], Titus Tienaah[2], Amit Seeram[1], Bheshem Ramlal[1] and Susan Nichols[2]

[1] Department of Geomatics Engineering and Land Management, The University of the West Indies, St. Augustine, Trinidad and Tobago

[2] Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada E3B 5A3

Michael.Sutherland@sta.uwi.edu; ttienaah@unb.ca; amitseeram@hotmail.com; Bheshem.Ramlal@sta.uwi.edu; Nichols@unb.ca

Abstract

Public Participatory Geographic Information Systems (PPGIS) are capable of affording community groups and members opportunities to use digital or analog spatial objects, described as Volunteered Geographic Information (VGI), to express views on phenomena affecting their communities. These expressions of local knowledge can augment, complement or verify governance decision-making processes. VGI inputs to PPGIS can also together be important components of a locally relevant Spatial Data Infrastructure (SDI), especially in developing countries where information may be incomplete or unreliable. This chapter describes a PPGIS prototype that processes VGI, and that can be part of SDI in a developing country. However, VGI ought to meet SDI standards as framework or infrastructure datasets. The prototype system is capable of combining both empirical data and VGI. The prototype is also capable of supporting citizen-inclusive collaborative governance through the combination of empirical and local knowledge VGI, which provides richer governance decision-making resources.

KEYWORDS: PPGIS, VGI, Governance, Spatial Data Infrastructure

1. Introduction

The terms Public Participatory Geographic Information Systems (PPGIS), Participatory Geographic Information Systems (PGIS), Participatory Geographic Information and Multimedia Systems (PGIMS) and other similar terms refer to systems that manage spatial objects in simple GIS technologies (Corbett and Keller, 2005). The terms are also used in relation to spatial models created through community-collaborative activities such as those creating 3-dimensional physical spatial models. This chapter uses the term PPGIS. All of these systems afford community groups and members opportunities to use digital or analog spatial objects, described as Volunteered Geographic Information (VGI), to express views on phenomena affecting their communities.

Ansell and Gash (2008, p. 544) define collaborative governance as a "governing arrangement where one or more public agencies directly engage non-state stakeholders in a collective decision-making process that is formal, consensus-oriented, and deliberative and that aims to make or implement public policy or manage public programs or assets". This chapter explores the role of PPGIS and VGI as vehicles for including community groups in collaborative governance affecting their communities. The chapter's discussions leverage the development of a prototype PPGIS, designed to collaboratively engage communities, researchers and government agencies in the development of appropriate mitigation and adaptation strategies for climate change threats. The PPGIS prototype facilitates the integration of empirical primary and secondary spatial data with moderated community-relevant VGI (i.e. points, lines and polygons) representing community anecdotal local knowledge. This combined data provides richer sources for decision makers (inclusive of communities) involved in developing and implementing policies and other governance strategies for participating communities.

The prototype PPGIS uses open source platforms, which can be beneficial to developing countries. It is based on research done by Tienaah (2011) at the University of New Brunswick, Canada. Collaboration with researchers from the Department of Geomatics Engineering and Land Management at the University of the West Indies (UWI), St. Augustine, Trinidad and Tobago produced a system further developed and enhanced for Caribbean coastal communities. Grande Riviere, Trinidad and Tobago is the first test site for the PPGIS because the country has begun serious discussions on Spatial Data Infrastructure (SDI) development and the PPGIS development is therefore timely.

The development of the PPGIS is important for a number of reasons. Good governance requires informed decisions and this is especially a challenge at the local level because information is often lacking at that level. Often, and especially in developing countries, there is not a lot of good data at the local level and therefore the importance of local knowledge needs to be recognized. Local knowledge can augment, complement, or verify what information is available for governance decision making at the community level. PPGIS is a way to make this happen; it is a way to get more local information and

engage local stakeholders. All of this can be part of a robust and locally relevant SDI, supporting governance at the local level and empowering realistic decision making where national databases may not be enough. In this chapter, a prototype PPGIS constructed with open-source tools exemplifies the need for these systems, which supports local knowledge inclusion as VGI, to be part of National SDI and hence citizen-inclusive collaborative governance. To support this perspective, the chapter first discusses the following links: SDI and collaborative governance; SDI and communities as spatial data suppliers; VGI and SDI; open source systems and SDI; PPGIS and SDI in the Caribbean.

2. SDI and Collaborative Governance

Governance is about decision making and steering in relation to defined jurisdictional, community or organizational objectives. The literature on governance point to the benefits of non-hierarchical governance models such as collaborative, cooperative, coordinative or integrative governance (Paquet, 1997; Rosell, 1999; Sutherland, 2005; Ansell and Gash, 2008). Discussions on the concept of SDI are more than twenty years old (Figure 1). Even then the benefits of institutional collaboration, cooperation or integration, especially with regard to the sharing of data, were recognized (McLaughlin, 1991; McLaughlin *et al*., 1993; Clinton, 1994; Coleman and McLaughlin, 1994; Tosta 1994).

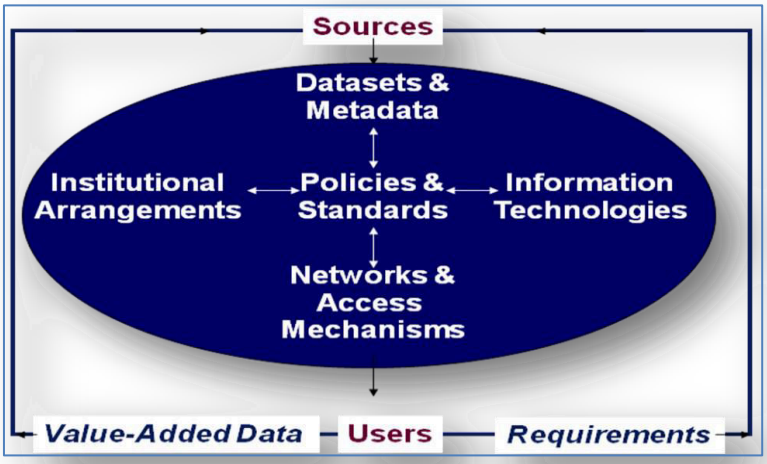

Figure 1. SDI, after McLaughlin and Nichols (1994)

The benefits of SDI are expounded to include *inter alia* (Vanderhaegen and Muro, 2005; International Hydrographic Organization, 2011):

- Improved decision making;

- Improved cost savings data management through reductions in duplication of effort related to data collection, processing and dissemination;
- More effective use of public funds;
- Greater cooperation/collaboration/integration with other spatial data providers;
- Increased opportunities for revenue generation.

These SDI benefits imply improved governance through collaboration of stakeholder data resources. The very concept of SDI is based on the recognition that non-hierarchical, non-silo models of governance produce better management of limited resources. To underscore this point, United Nations World Food Programme (2007, p. 2) specifically recommends that in the development of SDI, organizations should focus more on governance and sustainability, and the building of partnerships, than on "mere technology".

3. SDI and Communities as Spatial Data Suppliers

Definitions of SDI in the literature are similar in defining components which include spatial data and databases, attributes for spatial objects, metadata, information and communications infrastructures, institutions and institutional arrangements for sharing data, policies, standards, and spatial data suppliers and users (McLaughlin and Nichols, 1994; Chan and Williamson, 1999; Groot and McLaughlin, 2000; Global Spatial Data Infrastructure, 2004; International Hydrographic Organization, 2011). There is, however, a tendency to omit community groups and members as valid spatial data suppliers.

Global Spatial Data Infrastructure (2004) states:

> Anyone who is involved in a project of which spatial information forms an integral part and who intends leaving a legacy of spatial data or tools to exploit the data that lasts beyond the period of funding for the project is, by definition, participating in some of the fundamental elements required by an SDI. As coordination between such organizations expands, these projects very often lay the foundations on which initiatives formally dedicated to the establishment of SDI can then build.

An examination of this statement could lead to a deduction that "ordinary" citizens (very often being involved in projects where spatial information is an integral part thereof and who often are involved in project-participating community organizations contributing to the acquisition of local knowledge in the form of spatial data) are important data contributors to SDI. Yet much of the literature on SDI makes reference to cooperation among government and private-sector organizations. Generally, municipal-, state/provincial- and federal/national governments along with the private sector are accepted as the major suppliers of spatial data within SDI (Jacoby *et al*., 2002; Global Spatial Data Infrastructure, 2004; Bernard *et al*., 2005; DataBC, n.d.; GeoConnections, 2009; International Hydrographic Organization, 2011). Citizens are

mainly conceptualized as only users or recipients of the spatial data instead of as integral collaborative data providers (Budhathoki *et al*., 2008; Elwood, 2008).

Budhathok *et al*., (2008, p. 150) however, state that there is a need to "reconceptualize the notion of the SDI user from a passive recipient to an active information actor". The authors argue that in first generation SDIs, users were marginalized as passive participants and that second-generation SDIs shifted from the provision of data to services. A user-centric approach is foreseen in the next evolution of SDI (Rajabifard *et al*., 2006; Goodchild, 2008; McDougal, 2010). Budhathok *et al*., (2008) go on to describe a reconceptualization of the user with production functions extended from expert organizations to user organizations and individuals. The implication of this reconceptualization is stronger involvement of non-governmental- and community organizations in governance, especially in relation to their communities. Meltzer (2000) supports this perception of governance, referring to processes and traditions that give weight to the voices of citizens on public concerns, especially on how societies are directed and how decisions are made. Appropriate mechanisms of collaborative governance involving citizen groups and organizations providing VGI may be facilitated through 3rd generation evolutions of SDI. Trinidad and Tobago, which is contemplating SDI development, could leverage this evolution of SDI through the PPGIS described in Section 7 of this chapter. Other Caribbean jurisdictions could also derive this benefit.

4. Volunteered Geographic Information and SDI

VGI refers to spatial data and information voluntarily created by private citizens who are without formal training in geographic sciences, geomatics engineering, or geoinformatics (Goodchild, 2007a). The information produced in this manner may sometimes lack accuracies expected from empirical and other processes associated with the formal disciplines traditionally producing this type of information. However, VGI can often usefully complement more formally collected spatial data, in support of pertinent decision-making processes. Several authors have explored the connection between VGI and SDI (Craglia, 2007; Goodchild, 2007b; Budhathoki *et al*., 2008; Coleman, 2010; Miranda *et al*., 2011). The consensus is that VGI is under-utilized. The synergy between SDI and VGI has a potential to lead to a third-generation SDI in its development continuum (Rajabifard *et al*., 2006; Budhathoki *et al*., 2008).

Despite the major challenges of VGI to meet well-established standards of data quality, timeliness and completeness in an SDI, it does have great potential in collaborative governance. More positively, Ansel and Gash (2008) argue that trends toward collaboration also arise from the growth of knowledge and situational capacity. As knowledge becomes increasingly specialized and distributed and as institutional infrastructures become more complex and interdependent, the demand for collaboration increases. VGI data generated by the public can be used for such things as change detection and to identify public views and opinions on community phenomena in a collaborative governance environment.

5. Open-Source Systems and SDI

The emergence of open-source software provides a great opportunity for users to innovate and build atop existing platforms. Most open-source systems have flexible licensing models, which usually include the rights to use, modify, redistribute or redistribute modified copies. This "free" right makes it suitable for sharing source code, application and services in a SDI. Furthermore, most open-source licensing allows "free" (right and cost) for non-commercial purposes; this is cost effective for developing nations and local governments with limited budgets for software (Ouédraogo, 2005). Open-source systems can be useful for developing countries wanting to avoid expensive licensing costs associated with proprietary systems, and who are developing SDI.

6. Public Participatory GIS and SDI in the Caribbean

The term PPGIS covers systems that manage spatial models in simple GIS technologies (Corbett and Keller, 2005), or spatial models created from community-collaborative activities such those creating 3-dimensional physical spatial models. These systems afford community groups and members opportunities to use digital or analog spatial objects to express views on phenomena affecting their communities. Corbett and Keller (2005, p. 25) report that "…uses of Participatory Geographic Information Systems (PGISs) by disadvantaged groups can be empowering by enabling community groups and members to communicate local information and world views, using the commonly recognized language of Cartography in a way that might influence decision-making processes related to land use and planning". Considering the discussions in previous sections, this chapter views PPGIS as important potential sources of VGI. If a particular PPGIS can become a vehicle for the transmission of spatial objects representing moderated and approved community VGI, it can also be a data supply node in a SDI. From the same perspective, PPGIS can be a powerful tool for collaborative governance that empowers communities as active participants in decision-making processes.

In developing regions such as the Caribbean, many communities are not engaged in governance activities, and some of them wish to be. A PPGIS project, through the use of 3D modeling, was recently completed relevant to the island of Tobago, to offer such opportunities to community members. The project was part of an effort to prepare island stakeholders to develop adaptation strategies for climate change, to better cope with changes that are already occurring and further changes that are anticipated. The project was managed by The University of the West Indies (UWI) with funding from the United Nations Development Small Environmental Grant Fund, in collaboration with the Caribbean Natural Resources Institute with funding from the Centre for Agriculture and Technical Cooperation. The project entailed the building of a physical 3D model of Tobago (Figure 2) and the mapping of traditional and indigenous knowledge by community stakeholders and to identify possible impacts of climate change on their communities.

Figure 2. Physical 3D PPGIS model for Tobago

The 3D model was constructed using cardboard layers cut to match the contours of the island, and then stuck together to create a 3D representation of the island. Community stakeholders, involved in fishing, farming, tourism, hunting, environmental protection, forest management and resource management and who were previously briefed about the process, were then invited to populate the model with various types of features that they considered important. More than one hundred stakeholders participated in the mapping process. Most of the community stakeholders had the opportunity to describe the impacts that they were likely to encounter as a result of global climate change and the strategies that they would use in adapting to these changes. Most felt that they were better prepared to deal with the changes that were anticipated in the short and longer term. In the same manner, the PPGIS prototype that is the focus of this chapter could be used to gather community local knowledge inputs as part of a collaborative governance process. The inputs could be part of SDI because they are moderated and approved spatial objects in a simple GIS.

In Trinidad and Tobago too, there has been significant growth in the demand for digital spatial data products over the last decade. Several government agencies have embarked on the development of SDI independently to meet their own needs without consultation or collaboration with any of the other stakeholder agencies. This has led to duplication of resources and effort and has resulted in wastage, inefficiency and, in some cases, ineffectiveness. Recognizing the need to coordinate the efforts of SDI development in the country, the Government of the Republic of Trinidad and Tobago (GORTT) decided in November 2012 to appoint a committee to develop a framework for the implementation of a National Spatial Data Infrastructure (NSDI) for the country (Cabinet Minute 2860, November 1, 2012). The committee was expected to present its

findings and report to the Ministry of Planning and Sustainable Development with six months of its appointment. The terms of reference for the committee include (but not limited to):

(i) Conduct a comprehensive evaluation of the status of GIS and SDI within Government Agencies in Trinidad and Tobago;
(ii) Review existing international and regional structures of NSDIs and its applicability to Trinidad and Tobago; and
(iii) Conceptualize a framework for the development, implementation and governance of a NSDI.

The NSDI committee is also expected to engage with stakeholders to ensure that appropriate support and coordination for the initiative will be forthcoming. In this NSDI development scenario, GORTT has opportunity to leapfrog the historical development of SDI elsewhere by considering the inclusion of VGI from communities as part of a collaborative governance framework. The PPGIS prototype described in the next section provides one method of ensuring that communities are considered in NSDI development.

7. A PPGIS to Support Collaborative Governance and SDI

This section describes the PPGIS prototype developed to facilitate moderated community spatial data input as part of an envisioned collaborative governance framework. The prototype is an example of PPGIS in the form of functionally simple GIS technology. The PPGIS client and backend services were developed using open-source tools. The client interface consists of the mapping and graphics library: OpenLayers and ExtJS (GeoExt). The backend is scripted using Python. A model, view, and controller pattern is implemented using Django Web framework. The Web application is served using Apache as a Web server and GeoServer as a spatial data server. Spatial data uploaded are stored in a spatially enabled PostgreSQL database with PostGIS extension. Tools such as GDAL and Shapely are used for geometry construction and manipulation. Figure 3 describes the PPGIS system tiers and Figure 4 shows the user interface.

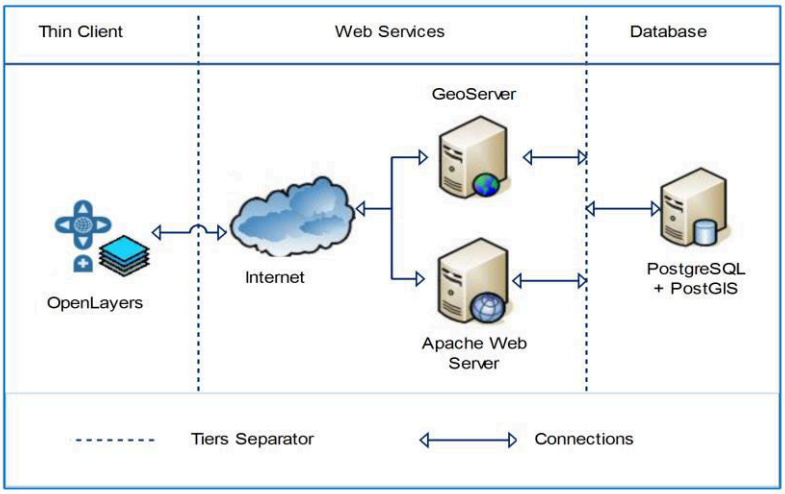

Figure 3. PPGIS system tiers (from Tienaah, 2011)

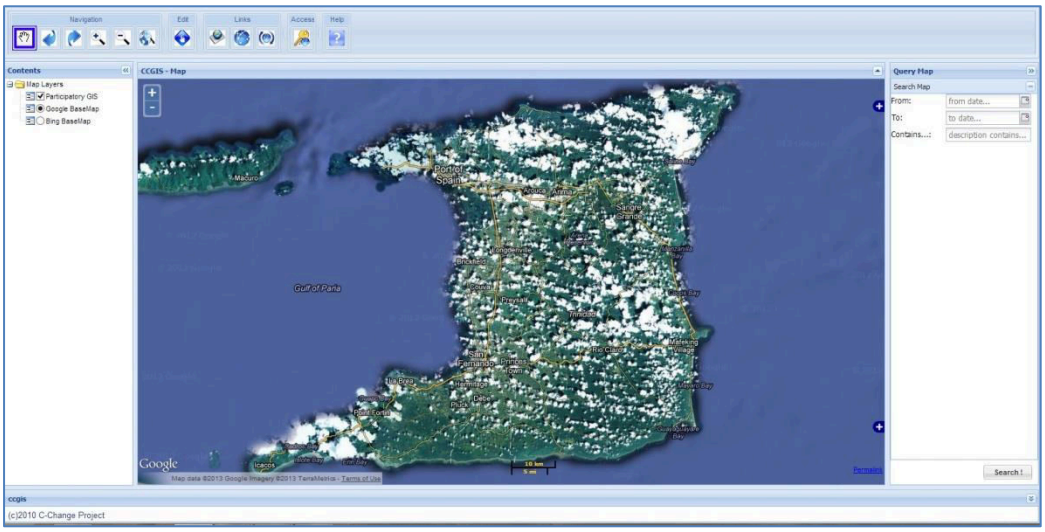

Figure 4. PPGIS user interface

The PPGIS uses Open Geospatial Consortium (OGC®) services through GeoServer. Through the RESTful interface of GeoServer, it is possible to manage services to be utilized by the PPGIS client using OpenLayers. By using tools and services conformant to OGC® specifications, the PPGIS can utilize and also serve spatial content within a SDI that is conformant with OGC® specifications. Users and decision makers can overlay multiple views or opinions as layers in a spatial context to enact participatory governance over defined spatial extents.

The PPGIS prototype is a public participatory platform to support community engagement in the process of developing adaptation and mitigation strategies to address the potential effects of sea level rise and storm events. Users become local sensors and voluntarily share their local knowledge and spatial experience in a Web 2.0 platform. In the broader picture of an SDI, the user or citizen group is an active participant in decision-making and in governance affecting their communities (i.e. communities become active “nodes” in collaborative governance).

The PPGIS facilitates moderated community local knowledge inputs in the form of spatial objects, i.e. points, lines or polygons (Figure 5). These objects may be integrated with empirical spatial input from government and other sources. This allows for spatial overlays of user-generated and authoritative contents. Each community is given a unique username and password that allows community members to access the system. Communities may view data for any global spatial extent but can only input data for their particular communities’ geographical extent. Additional meta-attributes such as text, pictures and videos may be linked to specific spatial objects.

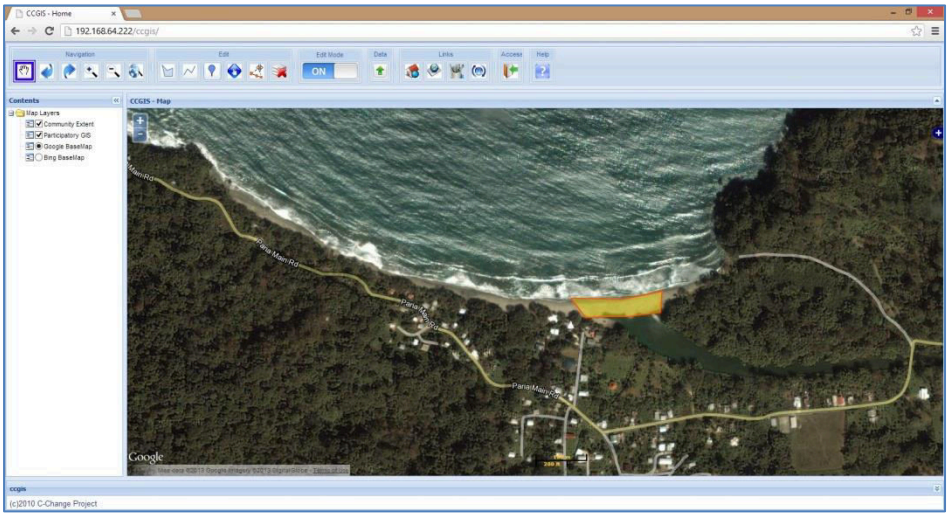

Figure 5. Sample community PPGIS input as a polygon (yellow)

The submitted spatial objects are checked by the moderator for acceptable content before being integrated into the system (Figure 6). The moderator is usually a system administrator at the host site.

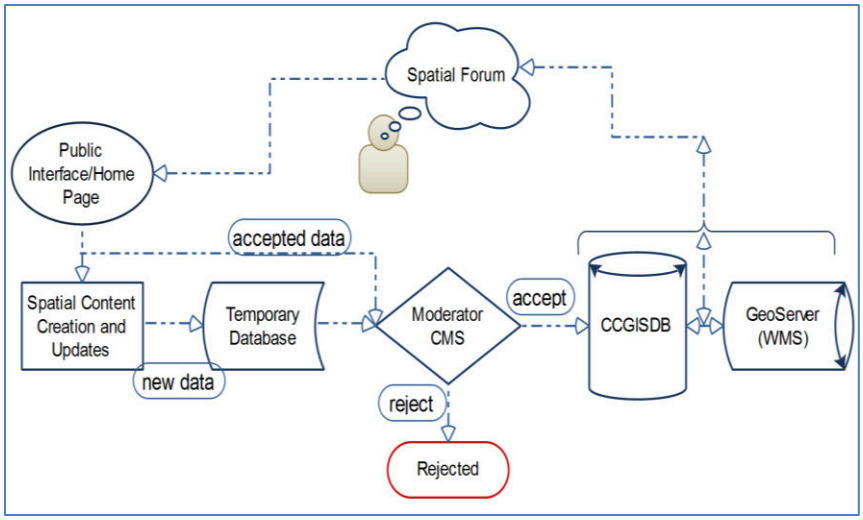

Figure 6. PPGIS moderated review process (from Tienaah, 2011)

An example of how the PPGIS may be used is in relation to sea level rise inundation models created under an International Community-University Research Alliance (ICURA) project for Grande Riviere using primary and secondary spatial data. The models show projected inundated areas based on Intergovernmental Panel on Climate Change (IPCC) sea level rise (SLR) projections, and were designed to aid decision makers in developing appropriate adaptation and mitigation strategies. During the field data collection, phase community members verbally volunteered environmental information related to flooding and inundation (i.e. local knowledge), pointing to spatial extents that used to be dry land or that were affected by flooding. Due to a lack of reliable historical spatial data, the models' outputs were unable to capture these local knowledge phenomena. In other words, the models are limited by a lack of data representing "community knowledge", a common occurrence in the Caribbean. Any decision maker basing decisions solely on the models' outputs may, in some circumstances, be misled. The PPGIS's ability to incorporate both empirical and community anecdotal local knowledge data provides richer materials for the governance decision-making process. Figure 7 depicts an SLR inundation model of Grande Riviere, and a community member pointing to structures built to mitigate flooding – data about the structure and its potential impact during floods or inundation events were not captured by the model.

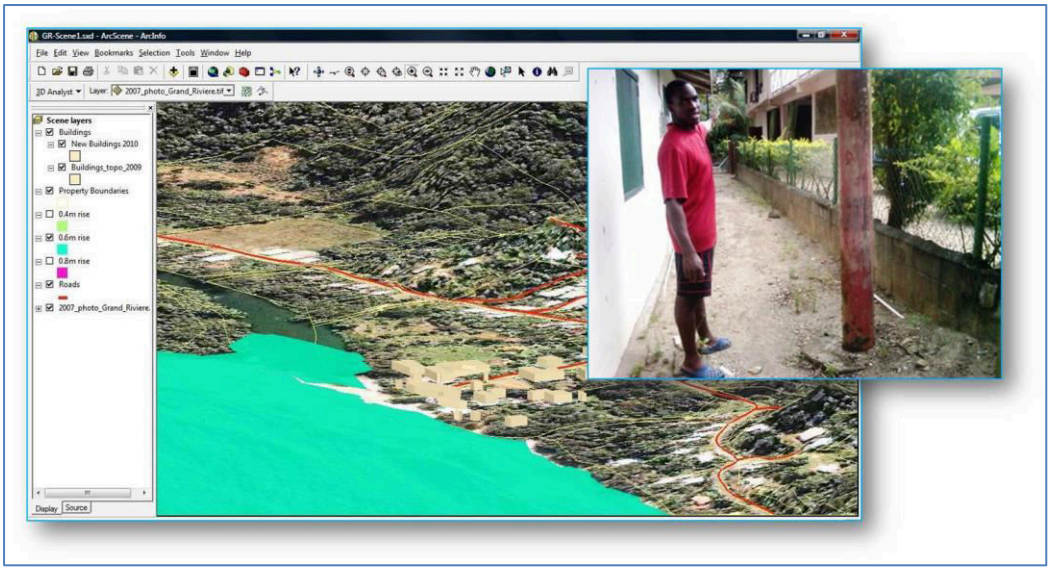

Figure 7. SLR inundation model, and Grande Riviere community member pointing to flood mitigating structures

This mode of collaborative governance is especially very useful in developing countries and regions such as the Caribbean where complete reliable data is not always at hand or easily obtainable. The PPGIS moderated system facilitates stakeholders (government, communities etc.) to collaborate through discussions/forums on the thematic and spatial meaning of community local knowledge input. In other words, the PPGIS supports collaborative governance and subsidiarity, and validly makes community VGI sources potentially part of SDIs. To make best use of the system, governments would have to develop and implement government-community collaborative governance frameworks and arrangements.

Currently, the PPGIS prototype is hosted in the Department of Geomatics Engineering and Land Management, UWI, St. Augustine, Trinidad. The system has been preliminarily introduced to members of the Grande Riviere community. A training workshop is planned to intensify the system's introduction to, and use by, the community. Feedback will facilitate further developments and assist when it is introduced to the other targeted communities.

8. Discussion

The PPGIS prototype uses an open-source model to facilitate community participation in collaborative governance. Though open- and closed-sourced software have different models on how users participate in source code development, both have important subtle overlaps. For most proprietary software, a team of developers are organized to produce a product to the end user. This model is similar to early SDIs where a set of

data providers (usually mapping organizations) provide data to meet the needs of users. Here the user is a passive participant in production. In the open-source model, users have access to the source code and therefore can modify or add functionality to meet their specialized needs. The open-source model in the application domain leads to fragmentation and organic growth because a modification is usually made to meet a particular users needs without necessarily considering the overall standards or the needs of others in the original project. Despite these varied use cases in the application domain, the core in the case of developing a huge infrastructure such the Linux kernel system is usually well organized to manage organic growth of the project. To apply the open-source model to VGI or PPGIS, and thus to use geographic data from the public in core infrastructure datasets, the data must pass through specific filters to meet overall infrastructure requirements and standards.

The intelligent use of external data sources and VGI could be introduced in a *graduated* manner that takes better advantage of local knowledge while not relinquishing control over contributor accountability and reliability of contributions (Tienaah *et al*., 2013). The benefits of using open-source tools in this project is to give communities and local governments the free rights to use, modify and redistribute copies for various projects. This liberates the participatory process from buying/renewing licenses with proprietary vendors. The PPGIS prototype described in this chapter was developed with climate change adaptation and mitigation strategies in mind. It was also developed based on the concept of community stakeholders being valid node contributors of local knowledge spatial data to SDI in the Caribbean. However, the functionalities of the PPGIS is generic enough to allow for its collaborative use by government, community and other stakeholders, relevant to any situation that would benefit from combined empirical spatial data- and community VGI inputs.

A major strength of the PPGIS also shows up as its major limitation: access to Internet connection. This is an important feature but has limitations in developing countries with slow Internet connection. Slow Internet connection can cause libraries and APIs to load slowly and can lead to slow application responses. Multimedia data and streaming of video also become a challenge. Another limitation is out of date satellite images as base layers to provide local spatial context. In less populated areas, satellite images from Google and Bing maps are usually of low resolution or may be out of date in rapid developing communities. The PPGIS application attempts to reduce this limitation by including both Google and Bing maps as optional base layers for instances where image have cloud cover, low resolution or are out of date.

Introducing the PPGIS into communities may also impose some social or economic cost factors for consideration. Socially, communities and other decision-making stakeholders will have to be convinced of the benefits to them of using the PPGIS. Both sets of stakeholders will first, however, have to share the perspective that collaboration among them will produce better decision-making. Varying levels of education within communities may also put certain community members in positions of advantage or disadvantage in terms of equity of access to the system, and the

systems implementation may create or exacerbate power relationships. There will definitely be a need for training to use the software. Use of the PPGIS will also require economic expenditures for at least computer hardware and Internet access, and some of these costs may have to be borne by the communities. Depending on the nature of the VGI and empirical data, issues of privacy may also have to be addressed. The foregoing issues are not definitive but give indication of social and economic expenditures that will have to be addressed, and which may pose impositions on target communities (Rambaldi and Callosa-Tarr, 2001; MacEachren *et al*., 2005; Elwood, 2006; Kim, 2008).

9. Conclusion

The use of PPGIS/VGI datasets may be used in the future to supplement authoritative datasets, as location-based devices improve in accuracy or tools are developed to validate user-generated content. Before VGI can be accepted into infrastructure datasets, it must meet SDI standards as a framework or infrastructure dataset. These standards provide a base for which other datasets in the user or application domain can be put in context. The prototype PPGIS described in this chapter can serve to include communities as valid SDI data contributors, especially in developing countries. The combined empirical and local knowledge VGI provides richer governance decision-making resources. The coastal community collaborator, be it a single user or a community action group, becomes an active user in the development of adaptation strategies through the submission of spatial objects representing community local knowledge. In this way, the PPGIS prototype facilitates citizen-inclusive collaborative governance, potentially giving communities voices in decision-making processes that affect them, under moderated and approved conditions.

Acknowledgement

The research and development of the prototype PPGIS was completed with joint financial support from the Association of Universities and Colleges of Canada and the International Development Research Centre, and also with funding from the Campus Research and Publication Grant, University of the West Indies, St. Augustine, Trinidad and Tobago. The authors also acknowledge the International Community-University Research Alliance project supported by Canada's Social Sciences and Humanities Research Council (SSHRC) and the International Development Research Centre (IDRC). The project is titled "Managing Adaptation to Coastal Environmental Change: Canada and the Caribbean" and was the source of funding for Tienaah (2011).

References

Ansell, C. and Gash, A., (2008), Collaborative Governance in Theory and Practice. *Journal of Public Administration Research and Theory*, 18(4): 543-571.

Bernard, L., Kanellopoulos, I., Annoni, A. and Smits, P., (2005), The European Geoportal – One Step Towards the Establishment of a European Spatial Data Infrastructure. *Computers, Environment and Urban Systems*, 29(1): 15-31.

Budhathoki, N.R., Bruce, B. and Nedovic-Budic, Z., (2008), Reconceptualizing the Role of the User of Spatial Data Infrastructure. *GeoJournal* (2008) 72: 149-160, DOI 10.1007/s10708-008-9189-x.

Chan, O.C. and Williamson, I. P., (1999), Spatial Data Infrastructure Management: Lessons from Corporate GIS Development. *Proceedings of The 27th Annual Conference of AURISA, Blue Mountains, New South Wales, 22-26 November 1999*.

Clinton, W.J., (1994), Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure. Presidential Executive Order 12906. Federal Register, Vol. 59: 71, pp. 17671-17674. April 13.

Coleman, D.J., (2010), Volunteered Geographic Information in Spatial Data Infrastructure: An Early Look at Opportunities and Constraints, Available from http://www.gsdi.org/gsdiconf/gsdi12/papers/905.pdf [Accessed 3 May 2013].

Coleman, D.J. and McLaughlin, J.D., (1994), Building a Global Spatial Data Infrastructure: Usage Paradigms and Market Influences. *Geomatica* 48(3): 225-236.

Corbett, J. M. and Keller, C. P., (2005), Empowerment and Participatory Geographic Information and Multimedia Systems: Observations from Two Communities in Indonesia. *Information Technologies and International Development*, 2(2): 25-44.

Craglia, M., (2007), Volunteered Geographic Information and Spatial Data Infrastructures: When do Parallel Lines Converge?, Available from www.ncgia.ucsb.edu/projects/vgi/docs/position/Craglia_paper.pdf. [Accessed 10 May 2013].

DataBC (n.d.). Spatial Data Infrastructure, Available from www.data.gov.bc.ca/dbc/about/spatial_data_infrastructure/index.page [Accessed 10 May 2013].

Elwood, S., (2008), Grassroots Groups as Stakeholders in Spatial Data Infrastructures: Challenges and Opportunities for Local Data Development and Sharing. *International Journal of Geographical Information Science*, 22(1): 71-90.

Elwood, S., (2006), Critical Issues in Participatory GIS: Deconstructions, Reconstructions, and New Research Directions. *Transactions in GIS,* 10(5): 693-708.

GeoConnections (2009), *2009 Status Report to the United Nations: Canadian Geospatial Data Infrastructure*. GeoConnections, Natural Resources Canada, 23 pages.

Global Spatial Data Infrastructure (2004), *Developing Spatial Data Infrastructures: The SDI Cookbook, Version 2.0*, Ed. Douglas D. Nebert, Technical Working Group Chair, GSDI, 150 pages.

Goodchild, M.F., (2007a), Citizens as Sensors: Web 2.0 and the Volunteering of Geographic Information. *GeoFocus* (Editorial), 7: 8-10.

Goodchild, M.F., (2007b), Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0*. International Journal of Spatial Data Infrastructures Research*, 2: 24-32.

Goodchild, M.F., (2008), Commentary: Whither VGI?. *GeoJournal,* 72(3): 239-244.

Groot, R., and McLaughlin, J., (2000), "Introduction" In R. Groot & J. McLaughlin (Eds.), *Geospatial Data Infrastructure: Concepts, Cases, and Good Practice*, UK: Cambridge University Press, pp. 1–12.

International Hydrographic Organization (2011), *Spatial Data Infrastructures: The Marine Dimension – Guidance for Hydrographic Offices, Edition 1.1*. International Hydrographic Office, Monaco, 36 pages.

Jacoby, S., Smith, J., Ting, L., and Williamson, I., (2002), Developing a Common Spatial Data Infrastructure between State and Local Government – an Australian Case Study. *International Journal of Geographical Information Science*, 16(4): 305-322.

Kim, S., (2008), Spatial Inequality and Economic Development: Theories, Facts and Policies. The International Bank for Reconstruction and Development / The World Bank Working Paper No. 16, 48 pages.

MacEachren, A. M., Cai, G., Sharma, R., Rauschert, I., Brewer, I., Bolell, L., Shaparenko, B., Fuhrmann, S. and Wang, H., (2005), Enabling Collaborative Geoinformation Access and Decision-making Through a Natural, Multimodal Interface. *International Journal of Geographical Information Science*, 19(3): 293-317.

McDougal, K., (2010), From Silos to Networks – Will Users Drive Spatial Data Infrastructures in the Future?. *Proceedings of International Federation of Surveyors Congress 2010, Sydney, Australia*.

McLaughlin, J., and Nichols, S., (1994), Developing a National Spatial Data Infrastructure. *Journal of Surveying Engineering*, 120(2): 62-76.

McLaughlin, J.D., (1991), Towards National Spatial Data Infrastructure. *Proceedings of the 1991 Canadian Conference on GIS, Ottawa, Canada, pp. 1-5. Canadian Institute of Geomatics, Ottawa, Canada. March*.

McLaughlin, J.D. and Nichols S.E., and Coleman, D.J., (1993), Building a National Spatial Data Infrastructure. *Proceedings of the 1993 Annual Conference of the Canadian Society of Civil Engineering,* Fredericton, New Brunswick, pp. 437-447.

Meltzer, E., (2000), Oceans Governance: A Paradigm Shift or Pipe Dream?. Paper Presented at Coastal Zone 2000 Conference, Saint John, NB, September.

Miranda, T. S., Lisboa-Filho, J., de Souza, W.D., da Silva, O.C. and Davis Jr, C.A., (2011), Volunteered Geographic Information in the Context of Local Spatial Data Infrastructures. *Proceedings Urban Data Management Society Conferences,* Delft , Netherlands, 2011, 16 pages.

Ouédraogo, L., (2005), *Policies of United Nations System Organizations Towards the Use of Open Source Software (OSS) for Development*. Prepared for the Joint Inspection Unit, United Nations, Geneva, 31 pages.

Paquet, G., (1997), Alternative Service Delivery: Transforming the Practices of Governance. In: R. Ford, & D. Zussman (Eds.), *Alternative Service Delivery: Sharing Governance in Canada*. KPMG, The Institute of Public Administration, University of Ottawa Libraries.

Rajabifard, A., Binns, A., Masser, I. and Williamson, I., (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.

Raymond, E., (1999), The Cathedral and the Bazaar. *Knowledge Technology and Policy*, 12(3): 23-49.

Rambaldi, G. and Callosa-Tarr, J., (2001), Participatory 3-D Modeling: Bridging the Gap between Communities and GIS Technology, Paper presented at the International Workshop on “Participatory Technology Development and Local Knowledge for Sustainable Land Use in Southeast Asia”, Chiang Mai, Thailand, 6-7 June 2001.

Rosell, S.A., (1999), Renewing Governance: Governing by Learning in the Information Age. Oxford University Press.

Sutherland, M. D., (2001), *Marine Boundaries in Support of Good Ocean Governance*. PhD Thesis, Technical Report No. 232, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada, 372 pages.

Tienaah, T., (2011), *Design and Implementation of a Costal Collaborative GIS to Support Sea Level Rise and Storm Surge Adaptation Strategies*. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 276, University of New Brunswick, Fredericton, New Brunswick, Canada, 187 pages.

Tienaah, T., Rak, A. and Coleman, D., (2013), *An Examination and Critical Comparison of Alternative Maintenance Models for the Nova Scotia Digital Topographic Database*. Contract Report of 2-year consulting study undertaken for the GeoNova Program Office, Service Nova Scotia and Municipal Relations, Province of Nova Scotia.

Tosta, N., (1994), The National Spatial Data Infrastructure: Where Are We Now?. *Geo Info Systems*, 4(1): 25-28.

United Nations World Food Programme (2007), *WFP Spatial Data Infrastructure (SDI) In support to a UN SDI: Preliminary proposal, Version 1.0, 18/05/2007*. United Nations World Food Programme, 13 pages.

Vanderhaegen, M. and Muro, E., (2005), Contribution of a European Spatial Data Infrastructure to the Effectiveness of EIA and SEA studies. *Environmental Impact Assessment Review*, 25(2): 123-142.

CHAPTER 8

Developing Spatial Information Sharing Strategies across Natural Resource Management Communities

Dev Raj Paudyal[1], Kevin McDougall[2] and Armando Apan[3]

[1] Faculty of Health, Engineering and Sciences, University of Southern Queensland, Queensland 4350, Australia

[2] School of Civil Engineering and Surveying, University of Southern Queensland, Queensland 4350, Australia

[3] Australian Centre for Sustainable Catchments, University of Southern Queensland, Queensland 4350, Australia

paudyal@usq.edu.au; mcdougak@usq.edu.au; apana@usq.edu.au

Abstract

Spatial information plays an important role in many social, environmental, economic and political decisions and is increasingly acknowledged as a national resource essential for wider societal benefits. Natural Resource Management (NRM) is one area where spatial information can be used for improved planning and decision-making. Traditionally, national mapping agencies and government organizations have been the main spatial data providers for the natural resource management sector. Recent developments in spatial and information communication technology have provided a new opportunity for the NRM community to collect and manage spatial information. With this new environment, the access and sharing of spatial information between NRM communities and government agencies is emerging as an important issue. The aim of this chapter is to identify the key factors which influence spatial information sharing between state government organizations and regional NRM bodies/catchment management authorities in Australia and formulate strategies to facilitate spatial information sharing and hence support spatial enablement initiatives. A mixed method research approach was utilized to collect both quantitative and qualitative data from regional NRM bodies. A questionnaire survey conducted across 56 regional NRM

bodies provided the current status of spatial information access and sharing and explored the SDI development activities in the NRM sector in Australia. A detailed case study explored the effectiveness of spatial information and knowledge-sharing arrangements between regional NRM bodies and state government organizations. Using the mixed method design framework, the key factors which influence spatial information sharing between state government organizations and regional NRM bodies/catchment management authorities were identified and classified into six major classes as governance, economic, policy, legal, cultural and technical. The study suggests that the adoption and implementation of strategies can facilitate spatial information sharing and hence advancing spatially enabled communities across the NRM sector.

KEYWORDS: Spatial data infrastructure, spatial information sharing, natural resource management, catchment management, spatial enablement

1. Introduction

The issues related to climate change, urbanization, land use change, environmental degradation and sustainable development, are of global concern. For environmental sustainability and sustainable development, many initiatives have been undertaken which range from global to local scales including the Brundtland Report (1987), UN Rio Earth Summit-Agenda 21 (1992), Bogor Declaration (1996), Bathurst Declaration (1999), Millennium Development Goals (2000), Johannesburg World Summit (2002) (Dalrymple, 2005). According to Brundtland Report (United Nations, 1987), sustainable development means “meeting the needs of the present without compromising the needs of the future”. The three dimensions of sustainable development include the economic, environmental and social dimensions which form the “triple bottom line” (Williamson *et al*., 2010). Sustainable development requires meaningful dialog between the economic, environmental and social aspects of life (Ting, 2002) and strong frameworks are required by which land and natural resources can be effectively managed. Reliable information infrastructure is needed to record the environmental, social and economic dimensions of natural resource management and to support appropriate decision making and conflict resolution (Paudyal *et al*., 2009).

Within the information infrastructure, spatial information may be considered a special type of information and is increasingly acknowledged as a national resource essential for sustainable development (Warnest, 2005). This speciality has resulted in the emergence of spatial data infrastructures (SDI) as part of, or independent of, information infrastructures (Van Loenen, 2006). SDIs are about the facilitation and coordination of the exchange and sharing of spatial data between stakeholders in the spatial data community. An SDI is a network-based solution which can provide convenient, consistent, and effective access to geographic information and services to improve decision-making in the real world in which we live and interact (Onsrud, 2011). The ultimate objectives of these initiatives, as summarized by Masser (1998),

are to promote economic development, to stimulate better government and to foster environmental sustainability. The principal objective of SDIs is to facilitate access to the geographic information assets that are held by a wide range of stakeholders with a view to maximizing their overall usage (Masser, 2011).

The spatial information sharing will increase the benefits to society through the reduction of duplication of effort in collecting and maintaining of spatial data. Further, the exposure of these data to a wider community of users may also result in improvements in the quality of the data. The sharing of spatial data is critical to the development of comprehensive and inclusive SDIs (McDougall, 2006). However, the sharing of spatial data between jurisdictions, and hence SDI development, continues to be problematic. Therefore, it is necessary to identify the key factors which influence spatial information sharing between organizations.

This chapter focuses on understanding the current mechanisms of spatial information sharing amongst regional NRM bodies/catchment management authorities (CMAs) and state government organizations for sustainable catchment management outcomes. Further, it identifies key factors which influence spatial information sharing between state government organizations and regional NRM bodies/CMAs within Australia, and formulates appropriate strategies to facilitate spatial information sharing and hence support SDI development.

2. Background

2.1 Catchment Management for Sustainable Development

Catchment management refers to the practice of managing natural resources using river catchment systems as the unit of management (Commonwealth of Australia, 2000). It involves integrating and managing ecological, economic and social aspects of land, water and biodiversity resources around an identified catchment system. Catchment management issues are characterized by multiple stakeholders and multiple goals which cut across traditional as well as administrative boundaries (Love *et al*., 2006). Catchment management requires an integrated management approach as different institutions and individuals need to work together towards sustainable catchment outcomes (Paudyal and McDougall, 2008). From theme perspectives, catchment management is about management of land, water, biodiversity, coast and marine theme (Paudyal *et al*., 2012). The term catchment management and watershed management are used interchangeably. In USA and Canada, the term watershed management is used, however in Australia and UK, the term catchment management is more widely accepted. Catchment management strategies need to support initiatives aimed at meeting the demands of our changing world particularly to serve sustainable development in the broader sense through environmental management. The four pillars of sustainable development are economic development, environmental management, social justice and good governance (Rajabifard *et al*., 2011).

2.2 Spatial Information for Natural Resource Management

Spatial information (also known as geographic information) is any information that can be geographically referenced, i.e. describing a location, or any information that can be linked to a location (ANZLIC, 2010). Spatial information is a key and integral component for the delivery of good governance, promoting efficiency in business and supporting sustainable development. It provides an enabling framework for modern societies and is recognized as fundamental for wealth creation and good decision making. As a result, policy makers and managers have begun to realize the value of spatial data to their business. They consider spatial data as a resource and also a part of fundamental infrastructure that needs to be coordinated and managed effectively (Ryttersgaard, 2001). Spatial information underpins decision-making for many disciplines (Clinton, 1994; Gore, 1998; Rajabifard *et al*., 2003) including natural resource management. Reliable, up to date and easy accessible spatial information is needed to support appropriate decision making and conflict resolution. Traditionally, government organizations and mapping agencies were the custodians of spatial information necessary for the catchment management whilst NRM community bodies were just the users of spatial information (Paudyal *et al*., 2011). The easily accessible and available spatial technologies and products like Google Earth, handheld navigation systems, Web 2.0 technologies, and social media can provide natural resource management communities with access to spatial data. However, with different organizations under different jurisdictions working towards natural resource management, the access, use and sharing of spatial information to support multi-stakeholder decision-making processes and policy development continues to be problematic.

2.3 Spatial Information Sharing: Research Gap

Calkins and Weatherbe (1995) defined spatial data sharing as “the (normally) electronic transfer of spatial data/information between two or more organizational units where there is independence between the holder of the data and the prospective user.” Omran (2007) defined it as “those transactions in which individuals, organizations or parts of organizations obtain access from other individuals, organizations or parts of organizations to spatial data.” McDougall (2006) clarified that the term “transaction” could be routine or non-routine, may be internal or external to the organization, but importantly it is an “arm’s-length exchange or transfer.”

Bregt (2011) reviewed the book “Building European Spatial Data Infrastructures” by Ian Masser (2010) and advised that the narrative anchor for SDI is “sharing spatial data”. Spatial data sharing is recognized as one of the important components in spatial data infrastructure design and development. There are many studies done by scholars for sharing spatial data (Kevany, 1995; McDougall, 2006; Omran, 2007; Onsrud and Rushton, 1995; Warnest, 2005; Wehn de Montalvo, 2003), however, the studies were mainly based on the spatial data provider’s point of view and do not recognize the power of spatial data users. Due to the advent of spatial technology and spatial

awareness, spatial information users are becoming more important for the spatial data infrastructure design and development and hence it is necessary to look from the users' perspectives.

Despite all these benefits, spatial data sharing is easier to advocate than to practice (Azad and Wiggins, 1995). There are many issues that hinder sharing spatial information between organizations. The issues can be categorized into organizational/institutional issues, technical and technological issues, economic issues, legal considerations and political issues (McDougall, 2006). McDougall (2006) undertook a critical analysis of the spatial information issues through a literature study and concluded that the growing importance of Internet connectivity, resourcing, trust and institutional frameworks (particularly policy), are key issues.

There has been limited previous research on spatial data infrastructure and data sharing in catchment management.

2.4 Motivations and Barriers for Spatial Information Sharing

The issues that impact on the sharing of spatial information are broad-ranging and include organizational/institutional issues, technical and technological issues, economic factors, legal considerations and political issues (McDougall, 2006). Nedovic-Budic and Pinto (2000) identified two factors that shape the processes involved in data-sharing activities and their outcomes: motivations for engaging in data sharing activities, and structural characteristics of the interaction mechanisms implemented by the data-sharing entities. Many researchers (Harvey, 2001; Harvey and Tulloch, 2006b; McDougall, 2006; Nedovic-Budic and Pinto, 2000; Nedovic-Budic *et al*., 2011; Omran, 2007; Onsrud and Rushton, 1995; Sebake and Coetzee, 2013; Wehn de Montalvo, 2003) tried to understand the spatial data-sharing issues and the benefits and constraints in spatial data sharing. McDougall (2006) categorized these issues into barriers (constraints) and the benefits (which will motivate). Table 1 summarizes the motivators and barriers for spatial data sharing (i.e. why organizations may or may not engage in spatial data sharing). These motivators and barriers for spatial information sharing were determined through the literature review.

Motivators
Cost saving through lack of duplication of data collection and maintenance efforts
Improved data availability and quality
Enhanced organizational relationships through promotion of cross organizational relationships
Reduction in risk if organizations are prepared to contribute to the costs or development time for a shared initiative
High returns on investment
Improved user satisfaction
Barriers
Cost recovery, copyrights and legal liability
Priorities of the organization, organizational disincentives and lack of support from management
Trust and unequal commitment from organizations
Insufficient staff, staff turnover and lack of technical resources
Networking costs; data confidentiality, liability and pricing
Differences in data quality
Lack of common data definitions, format and models
Conflicting priorities
Lack of leadership and coordination mechanism
Cultural (political and institutional)
Power disparities and differing risk perception

Table 1. Motivators and barriers for spatial information sharing (after Sebake and Coetzee, 2013)

2.5 Spatial Information Sharing Components

Various frameworks and components on data sharing are found in the literature. Amongst them are a generic model of the Mapping Science Committee of the National Research Council (National Research Council, 1993), taxonomy for research into spatial data sharing (Calkins and Weatherbe, 1995), antecedents and consequences of information sharing (Pinto and Onsrud, 1995), factors relevant to GIS data sharing (Kevany, 1995), a typology of six determinants of inter-organizational relationships (Oliver, 1990), typology based on inter-organizational relations and dynamics (Azad and Wiggins, 1995), an organizational data-sharing framework (Nedovic-Budic and Pinto, 1999) a model of willingness based on theory of planned behavior (Wehn de Montalvo, 2003), interaction between organizational behavior of spatial data sharing and social and cultural aspects (Omran, 2007), a collaboration model for national spatial data infrastructure (Warnest, 2005), local government data sharing (Harvey and Tulloch, 2006a; Tulloch and Harvey, 2008), the local-state data sharing partnership model (McDougall, 2006) and Geospatial one-stop (Goodchild *et al*., 2007). Most of these frameworks were based on the authors' experiences and have not been proven empirically except for Nedovic-Budic and Pinto's (1999), Wehn de Montalvo's (2003) Harvey and Tulloch's (2006a) and McDougall's (2006).

Australian Government Information Management Office (2009) has proposed some nine conditions for information sharing. They include provision of leadership, demonstrate value, act collaboratively, establish clear governance, establish custodianship guidelines, build for interoperability, use standards-based information, promote information re-use and ensure privacy and security. Pinto and Onsrud (1995) argued the factors to facilitate spatial information sharing between two or more GIS using organizations are superordinate goals, bureaucratization rules and procedures, incentives, accessibility, quality of relationships and resource scarcity. They demonstrated how these antecedent variables influenced the efficiency, effectiveness and enhanced decision-making ability of organization. This approach is based on organizational theory. The Office of the Director of National Intelligence (2008) has proposed a range of issues for information sharing that span governance, policy, technology, culture, and economic facets. Based on these three literatures five areas and their attributes are identified for spatial information sharing through collaborative networks. Table 2 describes these five key areas and their main attributes for spatial information sharing to improve NRM planning and decision-making process.

Components	Attributes
Governance (Sharing environment)	mission, goal, objectives, stakeholders (data producers and users), leadership, custodianship, roles and responsibilities, rights and restrictions, governance methods
Policy (Rules for sharing)	laws, rules and regulations, policies and procedures, protocols, accessibility, privacy, liability, copyrights, IPRs
Technology (Capacity to enable sharing)	data model, standards, software, security, tools/mechanism, data quality, metadata, resource, interoperability
Culture (Willingness to share)	trust, motivation, communication, adaptation during circumstances changes, reciprocity, relationship
Economics (Value of sharing)	funding, incentives, pricing, cost recovery, transaction cost

Table 2. Spatial information sharing components (Paudyal *et al*., 2010)

3. Methodology

3.1 Study Area Description

As catchment management issues are characterized by multiple stakeholders and multi-level governance cutting across traditional as well as administrative boundaries, the Australian case has been considered suitable for this study. Catchment management arrangements in Australia are implemented through the partnerships of government, community groups, private sector and academia. Under the Australian Constitution, the States are responsible for land and water management within their boundaries (Marshall, 2001). All states/territories have some form of catchment management authorities or natural resource management groups under their jurisdiction. There are both top-down and bottom-up approaches exist for catchment

management. Government organizations are leading from a top-down approach and the activities of regional NRM bodies /community organizations are bottom-up.

Regional NRM bodies/catchment management authorities (CMAs) have been established to address complex catchment management issues that involve many community groups and government agencies. There are 56 regional NRM bodies which are responsible for catchment management in Australia. The regional NRM bodies vary in their name, corporate structure, catchment management philosophy, and relationship to the state government organization. They are termed catchment management authorities in New South Wales and Victoria, catchment councils in Western Australia, NRM boards in South Australia, regional NRM groups in Queensland and Regional committees in Tasmania. CMAs comprise representatives of the major sectors of the community and government which are involved in, or influenced by, the management of land and water resources in the catchment. Their major role is to provide a forum for community input and discussion, prioritize the issues, and develop and promote the adoption of catchment management strategies. Figure 1 shows the location of case study area and boundary of 56 NRM regions.

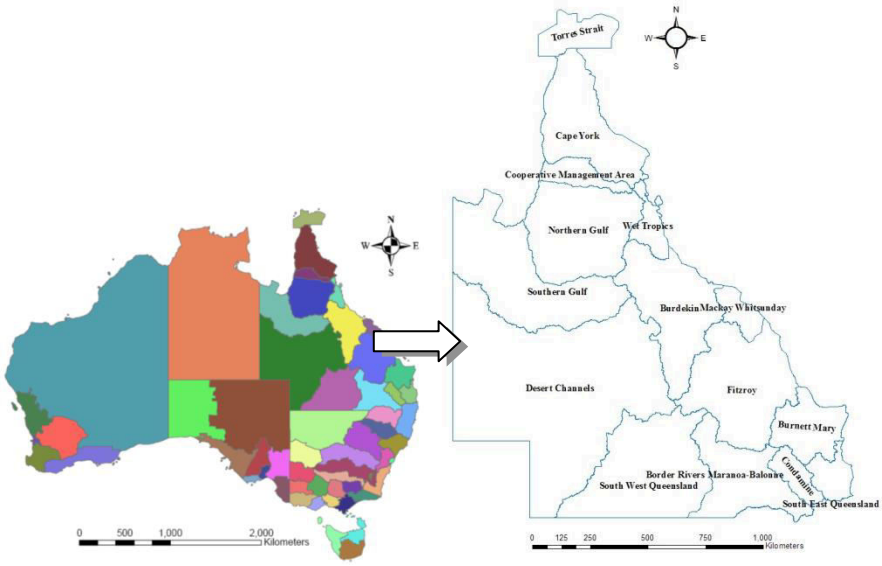

Figure 1. Location map of study areas

3.2 Research Method

This research has utilized mixed method strategy which involves collecting and analyzing both qualitative and quantitative data in a research study and mixing them. It has been argued by a number of researchers that the selection and use of appropriate data collection and analysis techniques are very important to the success of research (de Vaus, 2001; Marshall, 2006; Yin, 2009). Survey and case study were

considered to be the most appropriate method for data collection and analysis. The survey and case study data were collected and analyzed sequentially. Using the mixed method design framework as suggested by Creswell and Plano Clark (2011), the key factors which influence spatial information sharing between state government organizations and regional NRM bodies/catchment management authorities were identified and classified into six major classes as governance, economic, policy, legal, cultural and technical.

The survey was conducted with all 56 regional NRM bodies responsible for catchment management in Australia. The survey was undertaken from 15 June 2010 to 9 September 2010. A total of 56 valid responses were received to the online questionnaire giving an overall response rate of 100%. The questionnaire survey was distributed in two stages. Initially, the questionnaires were distributed to regional NRM bodies which belong to the Murray Darling Basin Authority (MDBA) and later to the remaining NRM bodies around Australia. The feedback and experience from the first distribution assisted in the second stage of the survey and assisted in achieving the high response rate. The online questionnaire was designed such that the data from questionnaire was automatically collected into an Excel spread sheet via a Web server. The raw data were reviewed and cleaned up before inputting into the statistical software. The statistical analysis was performed using SPSS statistics package. The profile of respondents is tabulated in Figure 2, with the largest group of respondents being GIS officers, with other respondents including staff who were directly or indirectly involved with spatial information management or the GIS operations of that regional NRM body. The majority of respondents were full-time staff.

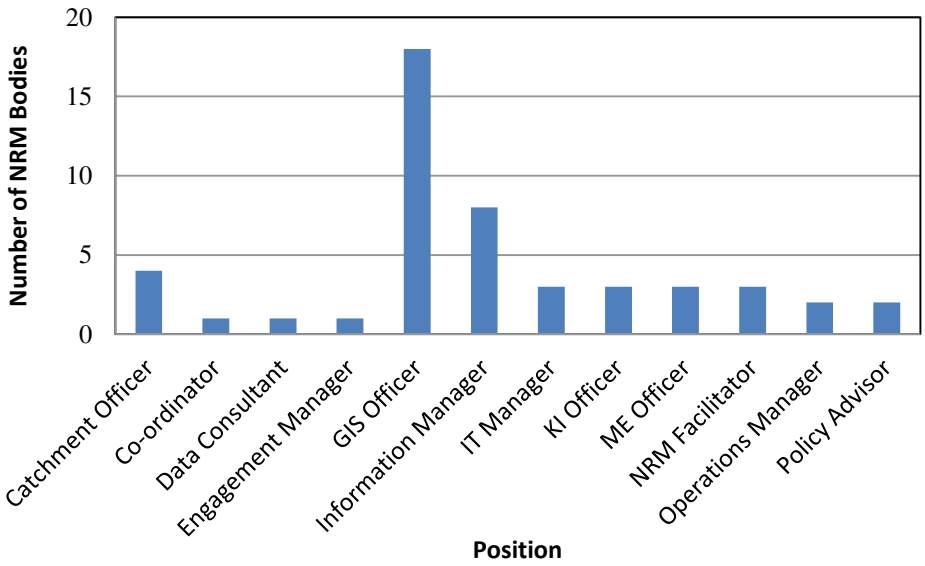

Figure 2. Profile of respondent (by position)

The case study approach was considered to be the most suitable approach for developing a deeper understanding about the motivation factors and constraints for spatial information sharing between regional NRM bodies and state government organizations, and confirm the issues related to spatial information management which were identified during the survey. The Knowledge and Information Network (KIN) project in Queensland was selected as a representative case to investigate spatial information and knowledge sharing process for catchment management. Queensland has 14 regional natural resource management (NRM) bodies and 74 local authorities spread from the far-northern region of Torres Strait to the New South Wales (NSW) border at southern end. These groups develop regional NRM plans and deliver sustainable catchment outcomes at grass-roots level.

The aim of case study was to determine the motivation factors and constraints for collaborating in the KIN project. Semi-structured interviews with all 14 regional NRM bodies, state government representatives and Queensland Regional NRM Groups Collective (RGC) were undertaken. The staff who were experienced in spatial and knowledge management activities were targeted for interview. A total of 15 staff from regional NRM bodies, two staff from RGC and three staff (both executive level and operational level staff) from state government agencies were interviewed. The responses were transcribed, analyzed and the main factors were determined.

4. Results

4.1 Results from Survey

4.1.1 Catchment Management Issues and Role of Spatial Information

There are disparities among regional NRM bodies regarding the catchment management issues on which they focus. However, we tried to explore the main catchment issues at national scale. Table 3 shows the top ten catchment management issues at the national level in Australia. The highest priorities include healthy habitat & biodiversity conservation, pest animal & weed management, community capacity building & indigenous engagement, disaster management, and water resource management. The grazing land & property management and Aboriginal NRM & cultural heritage are the less focused issue at national scale. This finding may assist federal and state government organizations for prioritizing funding and planning.

Rank	Catchment Management Issues	Frequency
1	Healthy Habitat and Biodiversity Conservation	38
2	Pest Animal and Weed Management	29
3	Community Capacity Building and Indigenous Engagement	27
4	Disaster Management (Fire Mapping, Floodplain, Land erosion, etc)	24
5	Water Resource Management	23
6	Land Use Planning and Soil Conservation	19
7	Climate Change	7
8	Coastal and Marine Management (estuarine and near shore)	5
9	Grazing Land and Property Management	4
10	Aboriginal NRM and Cultural Heritage	3

Table 3. Main catchment management issues

When asked to identify the role that spatial information can play in addressing the catchment management issues listed in Table 3, it was interesting to observe that approximately 60% of the regional NRM bodies responded that spatial information can play a very significant role, with the remaining 40% of the organizations responding that it can play a significant role. Not a single organization responded that it was not aware of the role of spatial information in addressing catchment management issues. This response indicates the importance of spatial information in supporting better catchment outcomes at the regional level (catchment level).

4.1.2 Spatial Data Providers and Identification of Spatial Information Requirement

The main spatial information providers to regional NRM bodies are the state government organizations. The majority (86%) of regional NRM bodies rated state government organizations as of high importance, whilst only 28% of regional NRM bodies rated commonwealth government organizations (e.g. Geoscience Australia, Bureau of Rural Sciences, etc) as of high importance. Local government organizations and private industries were identified as being of limited importance as a source of data. As spatial information is a critical component for improved catchment decision, the identification of the spatial information requirements is fundamental. Table 4 ranks the importance of spatial information for catchment management activities as identified by the NRM bodies.

Rank	Spatial information
1	Vegetation data
2	Cadastral data
3	Watershed/catchment boundary data
4	Land use/land cover data
5	Topography data
6	Aerial Photography and DEM
7	Satellite Imagery and LIDAR
8	Administrative boundary data
9	Infrastructure and utilities data (building, transportation etc)
10	Locally gathered data (GPS mainly) and Landholder data
11	Spatial project specific data
12	Geology and soil data
13	Open source data (Google Maps, OpenStreetMap, WikiMapia etc)
14	Mineral resources

Table 4. Spatial information needs for catchment management

Table 4 identifies that vegetation, cadastral and catchment boundary/watershed boundary, and land use/land cover data are the highest priority spatial data for catchment decisions. The regional NRM bodies were less concerned with geology and soil data, open source data or mineral resources data.

4.1.3 Spatial Information Sharing, Collaboration and Networking

The collaborative arrangements of regional NRM bodies with other organizations with respect to the exchange of resources, skills and technology were examined. The majority (83%) of the regional NRM bodies advised that they have a collaborative arrangement with other organizations. After investigation, it was found that data sharing and spatial information management were the main areas of collaboration. However, it was identified that the majority of regional NRM bodies had a silo approach to the spatial information management which did not encourage to spatial information sharing. The next most important area of collaboration was knowledge transfer (as illustrated in Figure 3).

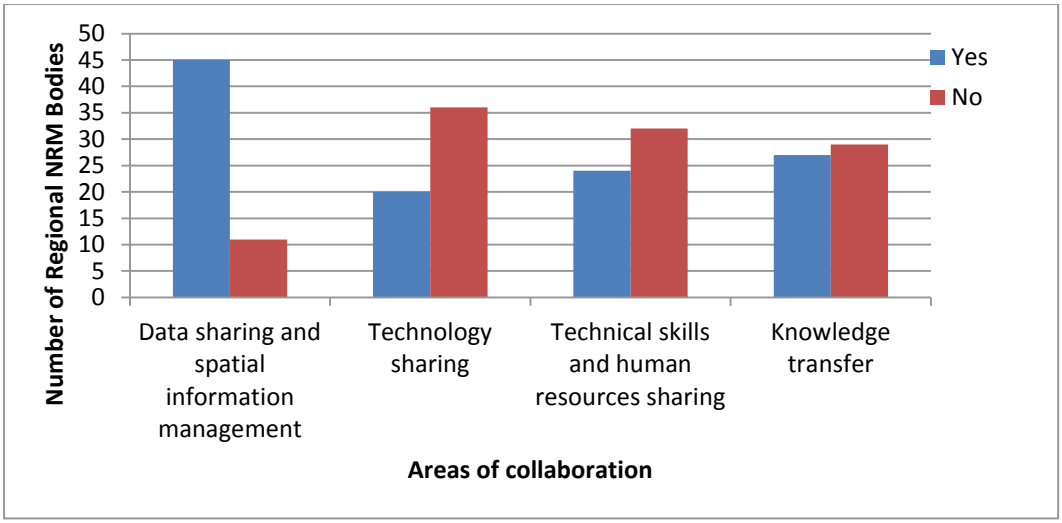

Figure 3. Areas of Collaboration

The main partners for these collaboration and networking activities were state government organizations with community organizations, including other regional NRM bodies, the next most common.

Spatial information sharing factors were identified and their importance in facilitating information sharing with other organizations was examined. Having a formal agreement, organizational attitude to sharing, individual attitude, ability and willingness to share, and leadership were found most important. Table 5 lists the spatial information sharing factors and their importance as rated by regional NRM bodies.

Spatial Information Sharing Factors	Importance
Formal agreement	Very High
Organizational attitude to sharing	Very high
Individual attitude, ability and willingness	Very High
Leadership	Very High
Networking and contacts	High
IT system and technical tools	High

Table 5. Spatial information sharing factors and their importance

4.1.4 Key Factors that Influence Data Sharing across Natural Resource Management Areas

A total of 21 factors were identified and classified into five broad groups: sharing environment (governance), rules for sharing (policy), capacity to enable sharing

(technology), willingness to share (culture) and value of sharing (economic). The five broad groups were identified during literature review (Section 2.5). The factors which were rated above 70% importance were classified as high, 50-70% are medium and less than 50% are low. The factors are shown in Table 6. The factors indicating the spatial capacity of the organization, spatial information policies and data sharing arrangements, spatial data requirements, access mechanisms, collaborative arrangements and willingness to provide data were the main factors which impacted on spatial information sharing between the regional NRM bodies and government agencies. The sharing environment, rules for sharing and willingness to share were the most important conditions for spatial information sharing.

Individual Factors	Conditions for sharing	Importance
Organization type	Sharing environment (Governance)	High
Spatial information use by staff	Sharing environment (Governance)	High
GIS maturity	Sharing environment (Governance)	High
Organizational capacity	Sharing environment (Governance)	High
Volunteer activities	Willingness to share (Cultural)	Low
Scale of spatial data	Sharing environment (Governance)	Low
Spatial information policy	Rules for sharing (Policy)	High
Funding sources	Value of sharing (Economic)	Medium
Spatial data requirements	Value of sharing (Economic)	High
Spatial data access medium	Rules for sharing (Policy)	Medium
Importance of spatial data providers	Sharing environment (Governance)	High
Ease of access to spatial data	Rules for sharing (Policy)	High
Frequency of supply	Capacity to enable sharing (Technical)	Low
Spatial data receiving medium	Capacity to enable sharing (Technical)	Medium
Restrictions on spatial data	Rules for sharing (Policy)	Medium
Integration issues	Capacity to enable sharing (Technical)	Low
Pricing of spatial data	Value of sharing (Economic)	Low
Collaborative arrangements	Sharing environment (Governance)	High
Data sharing agreement	Rules for sharing (Policy)	High
Social media, Web 2.0 technology	Capacity to enable sharing (Technical)	Medium
Willingness to provide spatial data	Willingness to share (Cultural)	High

Table 6. Factors that influence spatial information sharing

4.2 Results from Case Study

4.2.1 Motivational Factors for Collaborating and Data Sharing

The motivational factors for collaborating in the KIN project were determined through a semi-structured interview with all 14 regional NRM bodies, state government representatives and Queensland Regional NRM Groups Collective (RGC).

The motivation for collaborating in the KIN project was to better organize information and knowledge, to reduce cost, avoid duplication, and to enhance better collaboration and networking. However, the motivational factors varied between stakeholders. Basically, three types of organizations were involved in the KIN project and the motivations for these organizations are shown in Table 7.

Motivational Factors	
Regional NRM Bodies	State-wide project
	To enhance collaboration and networking
	To better organize knowledge and information
	To create an improved information portal
	To reduce cost, avoid duplication and optimize the use of resources
State Government Organization	To maximize the use of spatial information
	To improve collaboration and networking
	To achieve better regional NRM outcomes
Regional NRM Groups Collective	To avoid duplication
	To reduce cost and resources
	To encourage collaboration and networking
	The project was aligned with the organizational mandate and strategic goal

Table 7. Motivation factors for collaborating and data sharing

The main motivational factors for collaborating in the KIN project were to organize information and knowledge better, to reduce cost, avoid duplication, and to enhance better collaboration and networking. These motivational factors are also supported by previous research (Harvey, 2001; Harvey and Tulloch, 2006; McDougall, 2006; Nedovic-Budic and Pinto, 2000; Nedovic-Budic *et al.,* 2011; Omran, 2007; Onsrud and Rushton, 1995; Wehn de Montalvo, 2003).

4.2.2 Constraints Managing KIN Project and Spatial Information Sharing

There were a number of constraints in managing the KIN project and the spatial information sharing. The constraints were categorized into five broad areas as policy issues, organizational/governance issues, cultural issues, economic issues and

technical issues. The main organizational issues included concern about losing authority, and data sharing not being an organizational priority. The policy issues included the lack of spatial policy, pricing issues, and the lack of policies to return the data to the state repository. The legal issues included the licensing arrangements and privacy/confidentiality. The continuity of funding and incentives for sharing were identified as the key economic issues, whilst lack of trust and confidentiality were identified as cultural issues. Finally, lack of metadata and no single gateway to access spatial data were the main technical issues. From case study, it has been identified that the non-technical issues such as policy, governance, cultural and economic issues were found to be more significant for the success of the KIN project in comparison with the technical issues. The constraints managing KIN project are shown in Table 8.

Constraints	
No state government policy to include the spatial information back to the state repository	Policy
Spatial data has different scales, contents, qualities and standards and does not match with state government standards	Policy
Access policy, pricing and licensing arrangements	Policy
Lack of common standards or specification during data collection	Technical
National standard developed by Geoscience Australia is not suitable for catchment level data	Technical
Lack of single gateway to access NRM related spatial information	Technical
Data integration difficulties	Technical
People in state government organizations concerned about lose their power and control	Governance
Privacy issue with landholders' information	Cultural
Lack of trust and fear of data misuse	Cultural
Funding	Economic

Table 8. Constraints managing KIN Project and spatial information sharing

The KIN study identified the importance of improving the institutional and cultural component of the data sharing mechanism.

4.3 Integration of Survey and Case Study Results

This research followed the embedded mixed method design. In the embedded mixed method design, different datasets are connected within the methodology framed by other datasets at design phase to help in interpretation of the results (Creswell and Plano Clark, 2011). The case study results provided a supportive role and enhanced the findings from the national survey. A summary of the spatial information sharing issues identified during the survey and case study are presented in Table 9. Table 2 was used to classify the factors into five broad groups. The factors which were identified during survey or case study were indicated by (√).

Spatial Information Sharing Factors	Survey	Case study	Factor's group/Class
Organization type	√		Governance
Spatial information use by staff	√		Governance
GIS maturity	√		Governance
Organizational capacity	√		Governance
Spatial information policy	√	√	Policy
Data custodianship	√	√	Governance
Funding	√	√	Economic
Incentives	√	√	Economic
Spatial data requirements	√		Governance
Spatial information access medium	√		Technical
Importance of spatial information providers	√		Governance
Ease of access spatial information	√		Policy
Spatial information receiving medium	√		Technical
Restrictions on spatial information	√		Legal
Collaborative arrangements	√	√	Governance
Data sharing agreement		√	Legal
Licensing		√	Legal
Social media, web 2.0/3.0 technology	√	√	Technical
Willingness to provide spatial data	√	√	Governance
Trust		√	Cultural
Willingness to share spatial data	√	√	Cultural
Data integration	√	√	Technical
Data portal	√		Technical
Networking/contact	√	√	Governance
Leadership/champion	√		Governance

Table 9. Factors that influence spatial information sharing

4.4 Relationship between NRM Sectors and Identified Factors

The common findings from survey and case study were interpreted and the conditions which influence data sharing across catchment were categorized into six groups, namely governance (sharing environment), policy (rules for sharing), technical (capacity to enable sharing), cultural (will to share), legal and economic (value of sharing). The governance, policy/legal, cultural and economic factors were the most important conditions for spatial information sharing. The technological capacity to share spatial information was available, however, the governance, policy, cultural and economic issues need to be addressed to improve spatial information sharing. This

research identified that non-technical factors were more important than technical factors, which was also supported by previous research (de Man, 2011; McDougall, 2006; Mohammadi, 2008; Nedovic-Budic and Pinto, 2000).

The six main governance factors that influence the spatial information sharing between regional NRM bodies and state government organizations include leadership/champion, collaboration arrangement, organizational capacity, networking/contact, organizational mandate and willingness to provide spatial data. Spatial information policy, data custodianship and ease of access were the three main policy factors. There were no or limited policies/guidelines in regional NRM bodies to manage spatial information. Specifically, there was no policy to return the spatial information collected by regional NRM bodies to the state repositories or to utilize that spatial information for updating statewide NRM databases. Spatial information sharing was not considered a part of the organizational mandate and was always considered a lower priority. The continuity of funding and incentives for spatial information sharing activities were the two main economic factors, whilst the data sharing agreements, licensing and restrictions were identified as the legal factors. Regional NRM bodies were used to multiple licensing arrangements with state government organizations and showed interest in sharing data under the Creative Commons Framework. Trust, willingness to share and attitude were cultural factors. The landholders' data contained information that was considered private and they feared that the information could be used against them by government. The data portal, standards and data integration and the lack of a single gateway to access NRM related spatial information, were identified as technical factors.

5. Developing Spatial Information Sharing Strategies

The strategies were developed to address the spatial information sharing factors. The adoption and implementation of these strategies can assist to improve spatial data sharing. Further, these strategies can accelerate the progress in the development of catchment SDI initiatives. Each strategy has been presented in Figure 4 and discussed in more detail below.

Figure 4. Spatial information sharing strategies

5.1 Collaboration and Networking

Collaboration and networking was identified as an important strategy to improve spatial information sharing. A particular issue that was identified was the poor relationship between regional NRM bodies and state government organizations in the provision of data. Various regional collaboration and networking activities already exist for natural resource management and lessons from their development can be gleaned and transplanted for spatial information sharing.

5.2 Promote knowledge sharing

Knowledge sharing is one activity where community organizations such as Landcare, Watercare, Bushcare, and Coastcare are achieving better natural resource management outcomes. The current focus of regional NRM bodies is for spatial data

and information sharing. The raw spatial data can be translated into meaningful knowledge resources for the wider benefits of society using spatial technology and web tools. Therefore, knowledge sharing is an emerging area to be considered when developing spatial data infrastructure (SDI).

5.3 Place people at the front and empower brokering

There are many technical solutions in place and it was found that a technology-based approach was not likely to make a significant difference for spatial information access and sharing. The real need was to place people at the front. The people part of SDI was found to be critical for sharing spatial information. It was found from the case study that the role of the classic librarian should be formalized and placed at the front within the institutional framework either as a knowledge broker or a focal person. The role of librarian will provide both energy and focus to enable better cataloging, indexing, interpretation and publication of NRM information. It was also found from the case study that the function of the librarian should not be housed in any regional NRM body but should be independent.

5.4 Prioritize Spatial Data Sharing as an Organizational Activity

Spatial information sharing is not an organizational mandate for regional NRM bodies. The organizational mandate should be revised and spatial data sharing should be included as a priority area.

5.5 Create Awareness

There is a need to create awareness regarding spatial data sharing. Awareness is not simply the knowledge about spatial information sharing benefits; it also involves the appreciation, recognition and engagement of regional NRM bodies and other community organizations for spatial information management. The organizational attitudes and individual willingness to share data can be improved through improved awareness.

5.6 Make Foundation Data Free

There is growing pressure for state government organizations to make foundation data free. Seventy five per cent of regional NRM bodies argue that foundation data should be made free as it is a public good and paid for by the public through their taxes. This will also maximize the use of spatial information. Additionally, private organizations such as Google Earth and OpenStreetMaps have already placed their spatial products free in the market place. In this competitive market, there is pressure on state government organizations and mapping agencies to make foundation data free. The Commonwealth Government and the Victorian Government have already recognized the benefits of improved access and availability of public sector information (PSI). The

findings from case study showed that making foundation data free will also encourage regional NRM bodies to utilize foundation data and to better organize their data.

5.7 Establish and Harmonize Information Policy

It was found that there was a lack of information policy in regional NRM bodies and so it is important to establish an appropriate information policy in these bodies. The main areas for the preparation of spatial information policy include spatial information access, pricing, data custodianship, licensing arrangements, utilization of open-source information and social media, and should include an arrangement for the spatial information collected by regional NRM bodies to be returned to the state repositories.

5.8 Continuous Funding and Provide Incentives for Information Sharing

One of the major constraints for spatial information sharing and SDI development for catchment management activities was funding. The key funding sources for regional NRM bodies are the commonwealth government, state government, landowner's "in-kind" contribution and local government. There is a need for more reliable and continuing funding for spatial information management area for NRM bodies.

Spatial information sharing is not the core business of regional NRM bodies. There is little motivation for regional NRM bodies to share spatial information as they are busy with their core business. Incentives should be put in place to encourage further sharing of spatial information. The incentives could be economic incentives or some form of acknowledgment, recognition or appreciation so that the individual's willingness to share spatial information will be increased.

5.9 Improved Licensing Arrangements

It is recommended that regional NRM bodies use a single licensing arrangement rather than multiple licensing with state government organizations. The Queensland licensing framework used by the RGC when sharing spatial information between regional NRM bodies and state government organizations is a useful model to follow for other states. This could be facilitated through utilizing the Creative Commons licensing framework or the Australian Government Open and Access Licensing (AusGOAL) framework. Creative Commons licenses are designed to facilitate and encourage greater flexibility in copyright. A single licensing arrangement will improve efficiency in accessing and sharing of spatial information between regional NRM bodies and government agencies.

5. 10 Respect Privacy and Build Trust

The data which is collected by Landcare groups and landholders often have privacy/confidentiality issues. It is necessary to respect the privacy of spatial information during data sharing. The community groups and farmers should be

assured that the collected data regarding their properties will not be misused. This will also help to build trust and enhance collaboration in the future.

5.11 Promote Volunteerism

The volunteer participation and engagement of community groups and citizens for natural resource management has a long history in Australia. These community volunteer activities have been successful in achieving improved environmental outcomes and are acknowledged by government agencies. The local environmental knowledge of these groups can also be used for spatial information collection and management. Recent developments in ICT tools and spatial technology have provided community groups with a new opportunity to collect and manage the spatial data and facilitate spatial information access, sharing and SDI development.

5.12 Utilizing a Single Gateway for Access

Many IT solutions and spatial portals exist; however, NRM bodies are confused about where to go and how to access the data they need. It was identified by regional NRM bodies that a single gateway (access point) for natural resource information would improve discovery and access to spatial data.

5.13 Use of Open Standards

A continuing technical difficulty for spatial information sharing and spatial data infrastructure development at sub-national level is interoperability. The spatial information collected or generated by regional NRM bodies are generally local and have various standards and formats. Because it is very difficult to integrate and utilize spatial data gathered from different sources, spatial portals need to be built using open source and OGC standards to encourage interoperability. If open standards are embraced, the integration, access and sharing of spatial data can be improved.

5.14 An Enterprise Approach

The regional NRM bodies have a silo approach to spatial data management. The silo approach does not encourage the sharing of spatial data. The enterprise approach is more reliable and stable. It consolidates ‘silos’ of information, standardizes existing technologies, and minimizes the duplication of information services. As catchment management issues cross the administrative boundaries the adoption of an enterprise approach for data management is recommended.

6. Conclusions

This chapter has contributed to the current body of knowledge by exploring the spatial information sharing arrangements in natural resource management areas and

formulating strategies to facilitate spatial information sharing between NRM communities and government agencies. Natural resource management in Australia are implemented through the partnerships of government, community groups, private sector and academia. The national survey provides a unique nationwide perspective on the spatial information access and sharing for catchment management. The output from the survey will help to identify priority catchment management issues, national NRM datasets and information infrastructure in Australia. Though there are disparities among regional NRM bodies regarding the catchment management issues on which they focus, we identified the top ten catchment management issues at national level. This may assist federal and state government organizations for prioritizing funding and planning. The main catchment management issues at national scale were healthy habitat & biodiversity conservation, pest animal & weed management, community capacity building & indigenous engagement, disaster management, and water resource management. Spatial information plays a significant role in addressing these catchment management issues and majority of regional NRM bodies agreed this statement. Vegetation, cadastral, catchment boundary and land use information were the highly used spatial data by regional NRM bodies for catchment decisions. Spatial information and knowledge sharing were identified as the main areas of collaboration with the main collaboration partners being state government agencies and community organizations.

The main motivational factors for collaboration were to better organize information and knowledge, to reduce cost/resources, to avoid duplication, to maximize the use of spatial information and to achieve better regional NRM outcomes. These motivational factors are also supported by previous research. Lack of spatial policy, lack of trust, privacy/confidentiality, and continuity of funding were identified as KIN framework implementation issues.

The critical factors for improving data sharing across catchment management authorities were identified through triangulating the findings from the literature review, the results of the national survey of regional NRM bodies and the KIN project case study. Eighteen issues were identified as being highly significant and classified into the six major classes of organizational, policy, economic, legal, cultural and technical. The non-technical factors (organizational, policy, economic, legal and cultural) were found to be more significant in comparison with the technical factor. Based on these findings, information-sharing strategies were developed. Fourteen major strategies were formulated and suggested that the adoption and implementation of strategies can assist in overcoming the spatial information sharing issues and will contribute to the development of catchment SDI. The findings and strategies from this research have the potential to improve spatial information sharing between regional NRM bodies and government organizations to support better catchment management decisions.

References

ANZLIC (2010), *Spatial Information Related Terms*. Available from: www.anzlic.org.au/glossary_terms.html, [Accessed 10 January 2010].

Australian Government Information Management Office (2009), *National Government Information Sharing Strategy: Unlocking Government Information Assets to Benefit the Broader Community*. Australian Government Information Management Office (AGIMO). Available from: www.finance.gov.au/publications/national-government-information-sharing-strategy/docs/ngiss.pdf, [Accessed 20 October 2012].

Azad, B. and Wiggins, L.L., (1995), Dynamics of Inter-organisational Data Sharing: A Conceptual Framework for Research, In *Sharing Geographic Information* (Eds, Onsrud, HJ and Rushton, G), Centre for Urban Policy Research, New Brunswick, pp. 22-43.

Bregt, A.K., (2011), Building European Spatial Data Infrastructures by Ian Masser. *International Journal of Geographical Information Science*, 25(11).

Calkins, H.W. and Weatherbe, R., (1995), Taxonomy of Spatial Data Sharing, In *Sharing Geographic Information* (Eds, Onsrud, HJ and Rushton, G), Center for Urban Policy Research, New Brunswick, New Jersey, pp. 65-100.

Clinton, W.J., (1994), *Coordinating Geographic Data Acquisition and Access to the National Spatial Data Infrastructure. Executive Order 12096*. Available from: http://govinfo.library.unt.edu/npr/library/direct/orders/20fa.html [Accessed 15 September 2011].

Commonwealth of Australia (2000), *Co-ordinating Catchment Management: Report of the Inquiry into Catchment Management*, House of Representatives Standing Committee on Environment and Heritage, Canberra.

Creswell, J.W. and Plano Clark, V.L., (2011), *Designing and Conducting Mixed Methods Research,* 2nd edn*,* Sage Publications, LA.

Dalrymple, K., (2005), *Expanding Rural Land Tenures to Alleviate Poverty,* Doctor of Philosophy, The University of Melbourne, Australia.

de Man, W.H.E., (2011), Spatial Data Infrastructuring: Praxis between Dilemmas. *International Journal of Spatial Data Infrastructures Research (IJSDIR)*, 6: 261-289.

de Vaus, D.A., (2001), *Research Design in Social Research,* Sage Publications, London.

Goodchild, M.F., Fu, P. and Rich, P., (2007), Sharing Geographic Information: An Assessment of the Geospatial One-Stop. *Annals of the Association of American Geographers*, 97(2): 249-265.

Gore, A., (1998), *The Digital Earth: Understanding our planet in the 21st century*. Digital Earth, Wikipedia. Available from: www.isde5.org/al_gore_speech.htm# [Accessed 15 September 2011].

Harvey, F., (2001), *Constructing GIS: Actor Networks of Collaboration*. Available from: www.urisa.org/files/HarveyVol13no1-3.pdf [Accessed 4 October 2011].

Harvey, F. and Tulloch, D., (2006a), Local government data sharing: Evaluating the foundations of spatial data infrastructures. *International Journal of Geographical Information Science*, 20(7): 743-768.

Harvey, F., and Tulloch, D., (2006b), Local Government Data Sharing: Evaluating the Foundations of Spatial Data Infrastructures. *International Journal of Geographical Information Science*, 20(7): 743-768.

Kevany, M.J., (1995), A Proposed Structure for Observing Data Sharing, In *Sharing Geographic Information* (Eds, Onsrud, H.J. and Rushton, G), Center for Urban Policy Research, New Brunswick, pp. 76-100.

Love, S., Paine, M., Melland, A. and Gourley, C., (2006), *Research or Extension? Scientists Participating in Collaborative Catchment Management,* Paper presented at the APEN International Conference, 6-8 March 2006, Beechworth, Victoria, Australia.

Marshall, C., (2006), *Designing Qualitative Research,* 4th edn*,* Sage Publications, Thousands Oaks Calif.

Marshall, G.R., (2001), *Crafting Cooperation in the Commons: An Economic Analysis of Prospects for Collaborative Environmental Governance,* Doctor of Philosophy, University of New England, Armidale, Australia.

Masser, I., (1998), *The First Generation of National Geographic Information Strategies,* Paper presented at the The 3rd Global Spatial Data Infrastructure Conference, 17-19 November 1998, Canberra, Australia.

Masser, I., (2010), *Building European Spatial Data Infrastructures,* Esri Press.

Masser, I., (2011), Emerging Frameworks in the Information Age: The Spatial Data Infrastructure Phenomenon, In *The SAGE Handbook of GIS and Society* (Eds, Nyerges, TL, Couclelis, H and McMaster, R), Sage Publications, London.

McDougall, K., (2006), *A Local-State Government Spatial Data Sharing Partnership Model to Facilitate SDI Development,* Doctor of Philosophy, The University of Melbourne, Australia.

Mohammadi, H., (2008), *The Integration of Multi-source Spatial Datasets in the Context of SDI Initiatives,* Doctor of Philosophy, The University of Melbourne, Australia.

National Research Council (1993), *Toward a Coordinated Spatial Data Infrastructure for the Nation*, Mapping Science Committee, National Academy Press, Washington DC.

Nedovic-Budic, Z. and Pinto, J.K., (1999), Understanding Interorganisational GIS Activities: A Conceptual Framework. *Journal of Urban and Regional Information Systems Association*, 11(1): 53-64.

Nedovic-Budic, Z. and Pinto, J.K., (2000), Information Sharing in an Interorganisational GIS Environment. *Environment and Planning B: Planning and Design*, 27(3): 455-474.

Nedovic-Budic, Z., Pinto, J.K. and Warnecke, L., (2011), GIS Database Development and Exchange: Interaction Mechanisms and Motivations, In *Spatial Data Infrastructures in Context: North and South* (Eds, Nedovic-Budic, Z, Crompvoets, J and Georgiadou, Y), CRS Press, Taylor & Francis Group, Boca Raton, London, New York.

Office of the Director of National Intelligence (2008), *Information Sharing Strategy*, Director of National Intelligence, Washington DC, USA.

Oliver, C., (1990), Determinants of Inter-organizational Relationships: Integration and Future Directions. *Academy of Management Review*, 15(2): 241-65.

Omran, E.E., (2007), *Spatial Data Sharing: From Theory to Practice,* Doctor of Philosophy, Wageningen University, The Netherlands.

Onsrud, H.J., (2011), 'Afterword' In *Spatial Data Infrastructure in Context: North and South* (Eds, Nedovic-Budic, Z, Crompvoets, J and Georgiadou, Y), Taylor & Francis, CRC Press, London.

Onsrud, H.J. and Rushton, G., (1995), *Sharing Geographic Information,* Centre for Urban Policy Research, Urban Policy Research, New Brunswick, New Jersey.

Paudyal, D.R. and McDougall, K., (2008), *Building Spatial Data Infrastructure to Support Sustainable Catchment Management,* Paper presented at the Queensland Spatial Conference 2008, 17-19 July 2008, Gold Coast, Queensland.

Paudyal, D.R., McDougall, K. and Apan, A., (2009), Building SDI Bridges for Catchment Management, In *Spatial Data Infrastructure Convergence: Research, Emerging Trends, and Critical Assessment* (Eds, Van Loenen, B, Besemer, JWJ and Zevenbergen, JA), Netherland Geodetic Commission (NGC), Delft, The Netherlands, pp. 265-279.

Paudyal, D.R., McDougall, K. and Apan, A., (2010), *A Regional Collaborative Network to Improve Spatial Information Sharing in Australia,* Paper presented at the 12th Global Spatial Data Infrastructure Association World Conference (GSDI 12), 19-22 October, 2010, Singapore.

Paudyal, D.R., McDougall, K. and Apan, A., (2011), Spatial Data Infrastructure Convergence: Building Spatial Data Infrastructure Bridges to Address Climate Change, In *Geoinformatics for Climate Change Studies* (Eds, Joshi, PK and Singh, TP), The Energy and Resourcecs Institute (TERI), New Delhi, India, pp. 377-392.

Paudyal, D.R., McDougall, K. and Apan, A., (2012), Exploring the Application of Volunteered Geographic Information to Catchment Management: A Survey Approach, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-4: 275-280.

Pinto, J.K. and Onsrud, H.J., (1995), Sharing Geographic Information Across Organisational Boundaries: A Research Framework, In *Sharing Geographic Information* (Eds, Onsrud, HJ and Rushton, G), Centre for Urban Policy Research, New Brunswick, New Jersey, pp. 45-64.

Rajabifard, A., Feeney, M-E.F. and Williamson, I.P., (2003), Spatial Data Infrastructures: Concept, Nature and SDI Hierarchy, In *Developing Spatial Data Infrastructures: From Concept to Reality* (Eds, Williamson, I.P., Rajabifard, A. and Feeney, M-E.F.), Taylor & Francis, London, UK, pp. 17-40.

Rajabifard, A., Williamson, I.P., Wallace, J. and Bennett, R., (2011), *Spatially Enabled Society,* Paper presented at the FIG Working Week 2011, 18-22 May 2011, Marrakech, Morocco.

Ryttersgaard, J., (2001), *Spatial Data Infrastructure-Developing Trends and Challenges,* Paper presented at the CODI-2, 4-7 September 2001, Addis Ababa, Ethiopia.

Sebake, M.D. and Coetzee, S., (2013), Address Data Sharing: Organizational Motivators and Barriers and their Implications for the South African Spatial Data Infrastructure, *International Journal of Spatial Data Infrastructure (IJSDIR)*, vol 8.

Ting, L., (2002), *Principles for an Integrated Land Administration System to Support Sustainable Development*, The University of Melbourne, Australia.

Tulloch, D.L. and Harvey, F., (2008), When Data Sharing Becomes Institutionalized: Best Practices in Local Government Geographic Information Relationships. *URISA Journal*, 19(2): 51-59.

United Nations (1987), *Our Common Future: From One Earth to One World,* University Press, Oxford.

Van Loenen, B., (2006), *Developing Geographic Information Infrastructures: The Role of Information Policies,* Doctor of Philosophy, Delft University of Technology, The Netherlands.

Warnest, M., (2005), *A Collaboration Model for National Spatial Data Infrastructure in Federated Countries,* Doctor of Philosophy, The University of Melbourne, Australia.

Wehn de Montalvo, U., (2003), *Mapping the Determinants of Spatial Data Sharing,* Ashgate Publishing Ltd, Aldershot, UK.

Williamson, I.P., Enemark, S., Wallace, J. and Rajabifard, A., (2010), *Land Administration for Sustainable Development,* ESRI Press California.

Yin, R.K., (2009), How to Do Better Case Studies, In *The SAGE Handbook of Applied Social Research Methods* (Eds, Bickman, L and Rog, DJ), SAGE, Los Angeles, pp. 254-282.

CHAPTER 9

Application of Multi-criteria Decision Analysis and GIS Techniques in Vulnerability Assessment of Coastal Inhabitants in Nigeria to Crude Oil Production and Transportation Activities

Omoleomo Olutoyin Omo-Irabor[1] and Samuel Bamidele Olobaniyi[2]

[1] Department of Earth Sciences, Federal University of Petroleum Resources, Effurun, Nigeria

[2] Department of Geosciences, University of Lagos, Lagos, Nigeria

omoomoirabor@yahoo.co.uk; sbolobaniyi@yahoo.com

Abstract

A framework integrating environmental, social and economic criteria for vulnerability assessment of coastal inhabitants is necessary to reliably assess the impact of industrial activities in such areas without bias. This chapter reports a procedure involving the integration of multi-criteria decision analysis (MCDA), remote sensing data and GIS techniques to evaluate the impact of crude oil production and transportation activities in the western Niger Delta region. First, potentially vulnerable areas were established using factors based on the threat, exposure and sensitivity of the environment to crude oil activities. Adaptive capacity criteria were applied to gauge the socio-economic ability of inhabitants of host communities to cope with problems arising from exploitation, exploration and transportation of crude oil and its related products. Various stakeholders interest (operators, regulators, community members and other major stakeholders) were directly incorporated into the approach to improve decision-making processes. Scores based on the adaptive capacity (AC) show that host communities have poor to moderate socio-economic development and the component weights indicate that economic wealth (AC1) and access to information/services (AC3) represent the most and least important factors respectively. The threat factor was adjudged the most important of the potential impact assessment (PIA) criteria. This is due to the age of the facilities used in oil exploration and production activities in the area and the frequency of vandalization of such facilities. The integration of the two main factors AC and PIA show that 70% of

the local government areas (LGA) investigated require greater capacity building for the inhabitants to be equipped in dealing with threats posed by oil pollution. The result also shows that 20% of the LGAs require rehabilitation. This implies that the producing companies need to embark of environmental degradation reversal strategies and urgently address the issue of facilities maintenance and the socio-economic wellbeing of the affected communities.

KEYWORDS: Multi-criteria decision analysis, GIS, Vulnerability assessment, adaptive capacity, Nigeria

1. Introduction

Coastal regions are under serious threat of environmental degradation due to their endowment with natural resources. Despite policies that have been put in place to protect the environment, increased destruction of forest resources, contamination of surface water, groundwater and soil and changes in the life pattern of rural dwellers mainly in developing countries by industrialization have adversely affected this fragile region.

Industrial activities are usually expected to contribute to sustainable development by improving on the growth and wellbeing of neighboring communities of operation. According to Agenda 21 of the Rio's declaration, "human beings are the centre of concern and an integral part of the development process for sustainable development and are therefore entitled to a healthy and productive life in harmony with nature" (Epps, 1997), In reality this is not the case as rural inhabitants and their sources of sustenance are particularly vulnerable as a result of their close proximity to industrial facilities such as crude oil wells and the aged pipelines that are the major transporting medium. Similarly, other methods of transportation such as oil tank trucks, barges and ships; likewise contribute to the problems posed on the environment and inhabitants. Factors responsible for oil-spill disasters include social, historical, political and environmental variables (Aprioku, 2003). Oil spills caused by sabotage, human error and equipment failure have directly affected the local inhabitants leading to social unrests in some regions, thus prompting the need to develop a framework to aid in the establishment of the degree of vulnerability of coastal inhabitants to environmental changes resulting from industrialization. This would aid in environmental planning, development of management strategies and unbiased distribution of limited resources to affected communities.

Impact of oil spills on the environment has been carried out by a number of authors (Oyeike *et al*., 2002; Olajire *et al*., 2005; Okereke *et al*., 2007; Osuji and Nwoye, 2007). Risk assessment models have also been developed to establish the probability of oil spill occurring in the marine environment (Roberts and Crawford, 2004). These and other works have not taken into consideration the vulnerability of local inhabitants living within the vicinity of industrial activities. This chapter thus provides a means of assessing the vulnerability of such persons with the aid of Multi-criteria decision

analysis (MCDA) and GIS techniques. MCDA provides the ability of coupling expert judgment and stakeholder values in the sustainable management of the environment. The Niger Delta region of Nigeria was used to test the applicability of the framework. In problems involving multiple choices with conflicting objectives, Multi-criteria Analysis (MCA) has been employed due to its ability to rank alternative options according to stakeholder preferences. A number of authors have applied MCA protected area zoning (Geneletti and van Duren, 2008), selection of appropriate technology for contaminated land cleanup and remediation, (Accorsi *et al*., 1999; Balasubramaniam *et al*., 2007), and land suitability analysis (Delgado *et al*., 2008). In this work, the Weighted Summation Method (WSM) was used for the evaluation of scores and weights for the selected criteria and the determination of different management options. WSM is the most popular evaluation method, mainly because of its simplicity. Whilst other models have a stronger theoretical basis they are rarely used because they are complicated and time consuming (von Winterfeldt and Edwards, 1986). WSM assumes additive aggregation of criterion values, which are normalized to make them comparable by means of value functions (Giupponi, 2007).

2. Conceptualization of Vulnerability

Vulnerability is the degree to which a system, sub-system, or system component is likely to experience harm due to exposure to hazard, either a perturbation or stress/stressor (Turner *et al*., 2003). Vulnerability in the context of this work is approached from the integrated model which views the impact of hazard as a function of exposure, sensitivity and adaptive capacity of the exposed target. According to Hinkel (2011), vulnerability assessments can be carried out to (i) identify mitigation targets; (ii) identify vulnerable entities; (iii) raise awareness; (iv) allocate adaptation funds; (v) monitor adaptation policy; and (vi) conduct scientific research. The criteria selected for the vulnerability assessment can be expressed as a function of exposure, sensitivity and adaptive capacity (Metzger *et al*., 2006).

$$Vulnerability = f\left[Exposure, Sensitivity, Adaptive\ Capacity\right] \quad (1)$$

2.1 Potential Impact Assessment

According to Mertzger *et al*., (2006), potential impact is a function of exposure and sensitivity as shown in Equation 2.

$$Potential\ \ Impact = f\left[Exposure, Sensitivity\right] \quad (2)$$

In terms of oil exploration and production activities a third function was included by the authors to accommodate the threat posed by oil facilities. Therefore, PIA (equation 3) can be estimated from three main criteria; (1) the threat posed by oil facilities, (2) exposure of ecosystem and rural populace and (3) sensitivity of the landscape to pollution.

$$\text{Potential Impact Assessment} = f\left[\text{Threat}, \text{Exposure}, \text{Sensitivity}\right] \quad (3)$$

2.2 Adaptive Capacity

Adaptive capacity is the ability of households to anticipate and respond to changes in coastal ecosystems and to minimize, cope with, and recover from the consequences (Smit and Wandel, 2006). The concept of adaptive capacity was introduced in the IPCC TAR (Intergovernmental Panel on Climate Change, 2001), according to which the factors that determine adaptive capacity to climate change include economic wealth, technology and infrastructure, information, knowledge and skills, institutions, equity and social capital (Metzger *et al*., 2006).

Adaptive capacity is context-specific and varies from country to country, from community to community, among social groups and individuals, and over time (Smit and Wandel, 2006). In order to ensure sustainable development and within the context of this research, the adaptive capacity was established by utilizing eight socio-economic indicators namely; (1) economic wealth, (2) education & knowledge, (3) access to information/services, (4) alternative sources of livelihood, (5) health care, (6) political will, (7) response agencies, (8) kinship network.

2.3 Human vulnerability

Vulnerability can be expressed as a function of potential impacts and adaptive capacity:

$$Vulnerability = f\left[Potential\ \ Impact, Adaptive\ Capacity\right] \quad (4)$$

The most vulnerable individuals or groups are those that (1) experience the most exposure to perturbations or stresses, (2) are the most sensitive to perturbations or stresses (i.e. most likely to suffer from exposure), and (3) have the weakest capacity to respond and ability to recover (Research and Assessment Systems for Sustainability Program, 2001).

Human vulnerability can be established by quantifying and plotting the potentially impacted variables against their adaptive capacity. Figure 1 presents a novel framework developed by the authors to integrate these two considerations. The framework provides options which could aid environmental planners/decision makers in the selection of recovery approaches resulting from the impact of industrial activities.

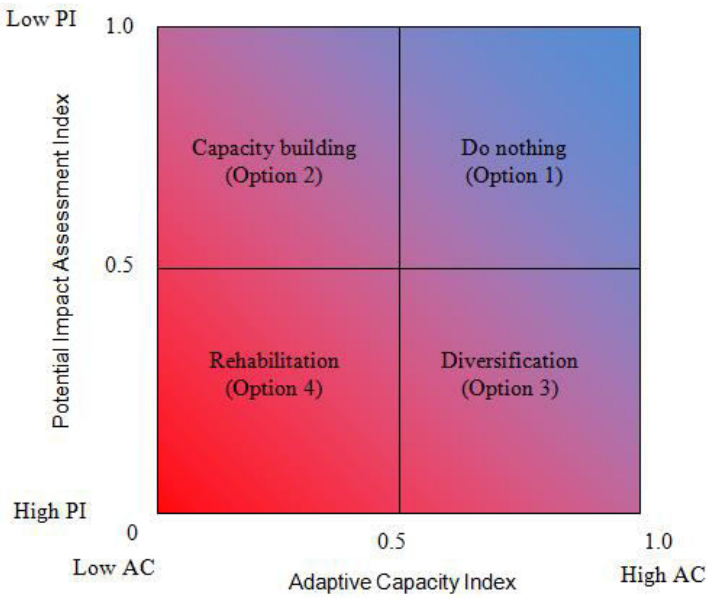

Figure 1. Human vulnerability assessment derived from potential impacts and adaptive capacity

From Figure 1, the following options can be deduced:

Option 1, the "Do nothing" scenario is the best option which indicates that potential impact of oil activities is low while the adaptive capacity is high. This scenario requires that the status quo remains although protected area management could also be appropriate. Option 2 involves capacity building implying low potential impact and adaptive. This option indicates capacity development through investments in poverty alleviation, infrastructure, social capital and alternative incomes. Option 3 has the likelihood for socio-economic change and diversification as both potential impact and adaptive capacity are high. Diversification involves reduction in oil activities and seeking alternative means of revenue generation. Option 4 (the worst scenario) has high potential impact and low adaptive capacity. This indicates that the environment has suffered significant degradation and requires remediation. Affected inhabitants may not have the resources/ability to adapt, therefore rehabilitation/resettlement is necessary.

3. Study Area

Delta State of Nigeria was selected for this study. The study area covers approximately 160 km of the State's coastline and a landmass of about 18,050km^2 of which less than 30% comprises of water bodies (Figure 2). The State produces about 30% of the total crude oil and natural gas output of Nigeria. The selection of the region for this study was based on the fact that oil and gas exploration and production activities are prominent making the petroleum industry the major source of revenue for Delta state

and Nigeria at large. The region also contains sensitive eco-systems with high marine biodiversity and critical habitats, particularly the mangrove. Therefore these conditions are ideal for developing a decision-support framework to identify the most suitable location where rehabilitation and/or developmental efforts should commence.

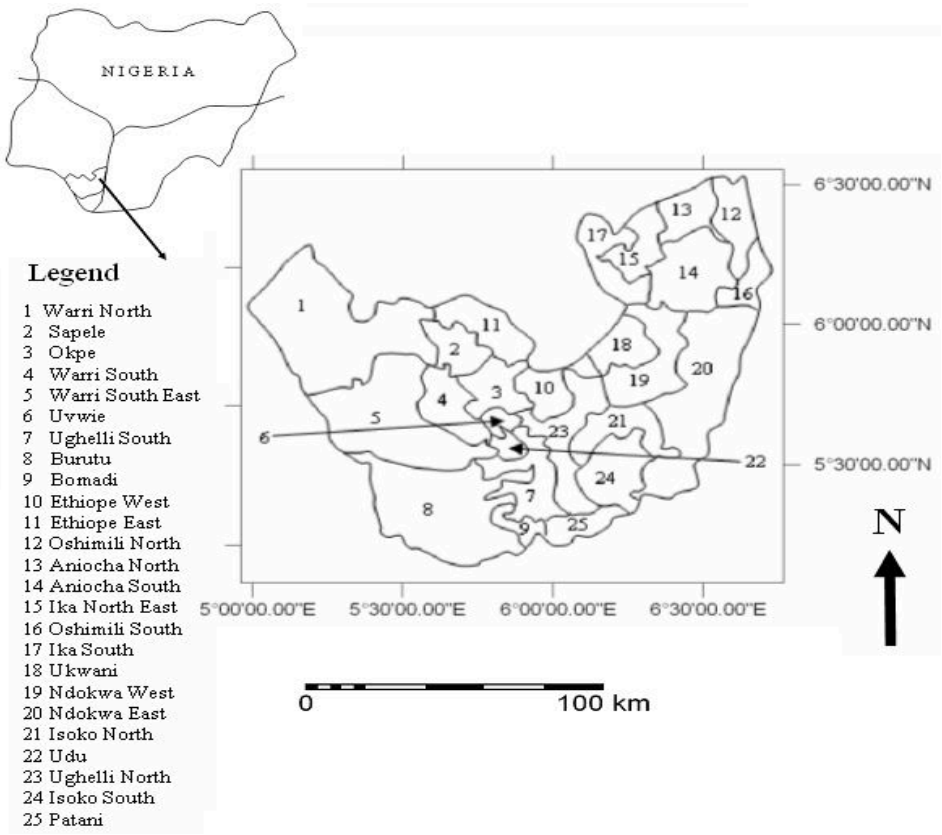

Figure 2. Map of Nigeria with Delta State and 25 local government areas (LGAs)

4. Methodology

The procedure developed and executed in this work is displayed in Figure 3.

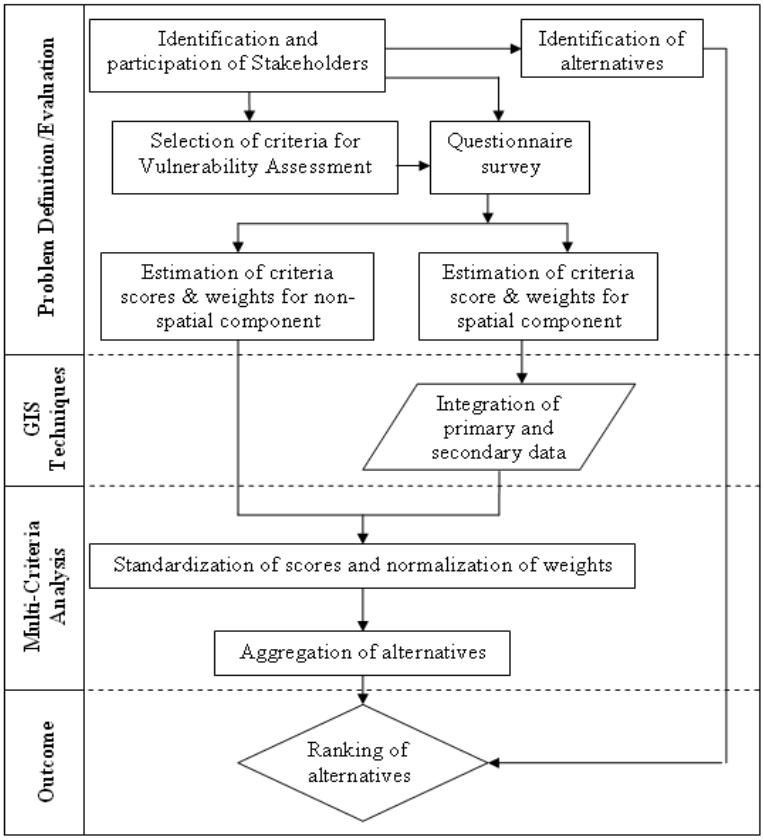

Figure 3. Components of multi-criteria decision analysis (MCDA) for vulnerability assessment

4.1 Problem Definition/Evaluation

4.1.1 Identification and Participation of Stakeholders

No matter the context, stakeholder involvement is increasingly recognized as being an essential element of successful environmental decision-making (Linkov *et al*., 2006). Participation of stakeholders in this study was divided along two main groups. The institutional/academic members (experts) provided expert opinion for assigning weights to the different criteria while the local residents (host community members) were solely for the purpose of obtaining information on adaptive capacity. This approach was adopted due to lack of socio-economic data.

4.1.2 Selection of Criteria and Identification of Alternatives

(i) Criteria for Potential Impact Assessment

Table 1 summarizes the sub-criteria for the main criteria based on expert opinion. These were derived from long-term field study by the authors and interaction with experts in environmental related fields.

Criteria	Sub-criteria	Classes	Score
Threat	Type of facility	Oil well	1
		Pipeline/flowline	2
		Flowstation	3
		Refinery	3
	Age of facility	Less than 5 years	1
		5 to less than 10 years	2
		10 to less than 20 years	2
		Greater than 20 years	3
	Type of spill/emission	Gas	5
		Petrol	4
		Kerosene	3
		Diesel	2
		Crude oil	1
	Volume of spill	Less than 2500 m^3	1
		Greater or equal to 2500 m^3	2
	Area coverage	Less than 1600 m^2	1
		Greater or equal to 1600 m^2	2
Exposure	Distance to rural settlements	0 to less than 100 m	4
		Greater or equal to 100 m	2
	Distance to agricultural lands	0 to less than 250 m	3
		Greater or equal to 250 m	1
	Distance to surface water bodies	0 to less than 250 m	3
		Greater or equal to 250 m	1
	Distance to forest	0 to less than 500 m	2
		Greater or equal to 500 m	1
Sensitivity	Size of population affected	Less than 1000	2
		1000 to 5000	3
		Greater than 5000	4
	Topography (slope)	Less than 8 %	1
		Greater or equal to 8 %	2
	Soil type	Gravel	4
		Sand	3
		Silt	2
		Clay	1
	Depth to water table	0 to less than 5 m	3
		5 to less than 10 m	2
		Greater or equal to 10 m	1
	Geology/Geomorphology	Coastal Plain sands	4
		Sombreiro-Warri plains	3
		Floodplain	2
		Mangrove swamp	2

Table 1. Criteria and sub-criteria and corresponding scores (raw) for potential impact assessment

(ii) Criteria for Adaptive Capacity

Eight criteria for determining the adaptive capacity are presented in Figure 4.

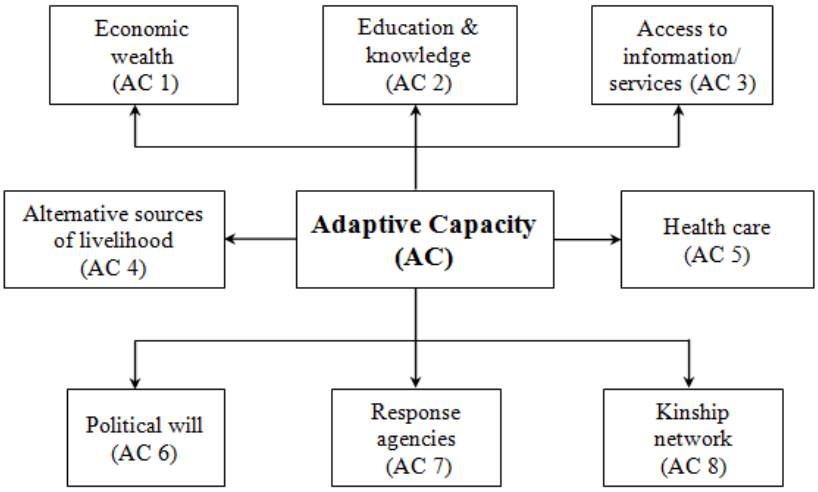

Figure 4. Criteria for adaptive capacity

4.1.3 Stakeholder Questionnaire Survey

The questionnaire was prepared carefully taking into consideration input from different stakeholders. The initial questionnaire was sampled by these groups to ensure that the questions were well understood and the meaningful within the context of the research. Two different sets of questionnaire were disseminated; the first questionnaire was used to obtain scores from community members for analyzing the quality of life through adaptive capacity while the second set of questionnaires was used to elicit relative importance of selected criteria. A total of 100 participants were involved in the eliciting of scores for selected criteria. These set of participants are expected to benefit from the developed framework in terms of the decision outcome.

The second set of questionnaire was used to obtain criteria weights from experts. 36 participants were drawn from oil companies (operators), government organizations (developers and regulators) and lecturers from higher institutions. Expert opinion can be considered a very important tool, as it provides flexibility without requiring detailed information or data for the problem under consideration. This process is performed by using experience and theoretical knowledge of the expert (Ercanoglu *et al.*, 2006).

4.2 Estimation of Scores and Weights (Relative Importance) of Criteria

4.2.1 Spatial Component (Potential Impact Assessment)

Score assignment for PIA was achieved through participatory group discussion involving experts made up of lecturers in one of the tertiary institution located in the study area. The researcher who headed the group discussion requested input on the scores to be assigned for each of the classes for the sub-criteria of PIA. The scores were assigned in such as manner that all sub-criteria were benefits to the overall objective (i.e. a high value implies a high impact). A high impact in the case of oil activities on the environment implies a negative effect. The scores which varied from 1 – 5 that were finally agreed upon are displayed in Table 1.

The scores assigned to each criterion were dependent on their overall effect on the vulnerability of the rural populace and their sources of sustenance (e.g. farming and fishing). Although soil type and geology/geomorphology appear to be similar, the soil type refers to the characteristics within the immediate vicinity of the area being assessed, if for instance clay is the soil type, it would be easier to contain any spill compared to if it was sand or gravel. The geology on the other hand is the characteristics on a regional scale with focus on the impact on the sub-surface features (e.g. aquifers) being of more significance.

4.2.2 Non-spatial Component (Adaptive Capacity)

The central tendency values e.g. median, grouped median, and mean of scores and weights of the criteria were computed using the SPSS Statistical package version 15.0 for Windows. Scores for all criteria were obtained on a similar scale of 1 to 3. A value of 1 indicates the worst outcome (i.e. the area has the highest negative impact) for each criterion, while 3 indicates the best outcome. The weights were obtained on a 5-point scale, where 1 indicates least important and 5 most important criteria.

4.3 Application of GIS Techniques

The GIS – Integrated Land and Water Information System (ILWIS) – Spatial Multi-criteria Evaluation (SMCE) was selected for the evaluation of PIA. SMCE window in ILWIS is an application that assists and guides a user in doing Multi-Criteria Evaluation (MCE) in a spatial way. ILWIS-SMCE consists of three phases – problem analysis, design of alternatives and decision making from alternative options. SMCE, method was used for determining relative importance of conditions affecting the rural populace and their immediate environment.

4.3.1 Preparation of Data Layers and Tables

The first step in PIA was to prepare the input data in compatible formats. The Landsat 7 TM orthorectified satellite images (row 56, paths 189 and 190) obtained from Global

Land Cover Facility (2005) provided the coordinate system and geo-reference that was applied to all the maps created. A common co-ordinate system (for vector data) and/or geo-reference (for raster data) is required for merging of data layers. The boundary of the study area was created by on-screen digitization of the satellite image. All spatial maps were converted to raster while others were stored in columns of attribute tables that were linked to one of the raster maps. Raster format is made of pixels (picture elements) of a certain size, e.g. 30m x 30m spatial resolution. In the case of this work a 30m pixel size was selected. This was done in order for the other maps to correspond to the classified landcover map derived from Landsat satellite images (Figure 5a).

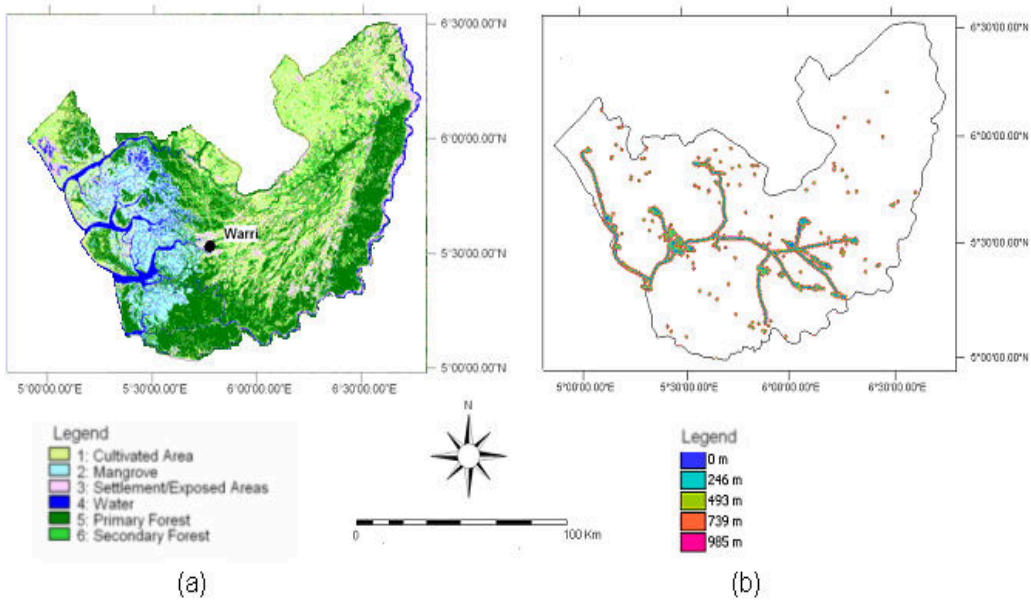

Figure 5. PIA layers (a) Landcover classes derived from Landsat TM satellite image (1986 and 1987) (b) Distance map of oil facilities

4.3.2 Construction of Criteria Tree in GIS Environment

After all the datasets required for PIA have been assembled, the next step was the construction of criteria /sub-criteria tree in the SMCE mode in the ILWIS software. The criteria tree is a tree whose root is the main goal defined by the researcher. In this case, the main goal is the potential impacts of oil activities on rural inhabitants. The leaves of the tree are the criteria required to evaluate the performance of the main goal while the branches divide the main goal into partial goals, namely threat, exposure and sensitivity (Table 2). A criterion can be a constraint or a factor. The constraints identified in this study were urban areas. Since the focus of vulnerability assessment was on the rural populace, urban areas were excluded from further

analysis by the assignment of a zero value. This implies that there is no compensation for other information for urban areas. The factors identified included threat, exposure and sensitivity.

The main factors or criteria were further divided into sub-criteria for ease of analysis and then assigned scores. It is essential to note that a criterion can be a cost or benefit to the overall assessment. Cost (C) and benefit (B) as used in this research does not imply financial loss or gain, but that an increase in the criteria will lead to an increase in the potential impact and vice versa.

4.4 Multi-criteria Analysis

Multi-criteria analysis involved a three-step approach; standardization of scores, normalization of weights and the ranking of alternatives.

4.4.1 Standardization of Criteria Scores

(i) Potential Impact Assessment

The criteria for PIA were measured using different measurement units therefore they had to be standardized to the same scale. The standardization process was executed using the ILWIS GIS software package. Standardization procedures are slightly different for constraints and factors. The standardized output values for constraints which are Boolean are either 0 (false) or 1 (true).

Standardized value for factors range between 0 and 1, such that low or poor performance of one criterion can be compensated by good performance in another criterion.

A distance map was created from available data of oil facilities as shown in Figure 5b. These included oil wells, pipelines, flow-stations and refinery.

This was then followed by the standardization of the sub-criteria using the raw scores displayed in Table 2. Depending on the type of class the combination method or direct method was applied. The combination method was used when dealing with a range of values. For example, the age of oil pipelines varied from 0-42 years. Pipelines less than 5 years old were considered to be of good quality, hence they were standardized to 0, those from 5 to 10 years had increased threat with age, while above 10 years were assigned a value of 1.

The direct method was applied to criteria that had qualitative information, for instance the spill type. The threat each posed increased with their volatility.

Topographical slope information for this study was extracted from the US Shuttle Radar Topography Mission (SRTM) (N05E05 - N06E06) because similar information was in lacking from locally available maps.

(ii) Adaptive Capacity

Standardization is essential when the unit of measure of the selected criteria differ. Since all criteria were on a 1-3 scale, it was not necessary to carry out any further standardization procedure.

4.4.2 Normalization of Criteria Weights

In order for the weight values to be combined, the process of normalization was carried by dividing each weight by the sum of the weights such that their total sum equals unity. A normalization of weights for AC and PI was accomplished using the formula

$$z = y_i \Big/ \sum_{i=1}^{n} y_i \tag{5}$$

Where z is the normalized weight value for the ith class, y_i is the raw weight.

4.4.3 Aggregation of Alternatives using Weighted Summation Method

The Weighted Summation Method (WSM) was applied for the aggregation of weights and scores. Potential impact (PI) was determined by

$$PI = \sum_{i=1}^{n} w_i x_i \prod_{j=1}^{m} c_j \tag{6}$$

where PI is the potential impact index, w_i is the weight of factor i, x_i is criterion score of factor i, n is the number of factors, c_j, is the criterion score (1 or 0) of constraints j and m is the number of constraints. In other words, Boolean images are created to represent each constraint, where the Boolean image has a value 1 for reclassified cells that satisfies the constraint and 0 otherwise.

The adaptive capacity (AC) was calculated using Equation 7.

$$AC = \sum_{i=1}^{n} w_i x_i \tag{7}$$

While the overall human vulnerability index (HVI) was calculated using Equation 8

$$HVI = f\left[Potential\ \ Impact, Adaptive\ Capacity\right] \tag{8}$$

5. Results

Table 2 presents a summary of the scores for adaptive capacity (AC) obtained from community members directly affected by oil activities. This result indicates that the host communities have poor to moderate adaptive capacities to risks associated with oil production and transportation activities.

LGAs	N	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8
Aniocha North	7	1.83	2.00	1.83	2.43	1.83	1.71	1.57	1.29
Aniocha South	3	2.00	2.00	2.33	1.67	2.33	2.67	2.00	1.33
Burutu	3	1.33	1.33	2.00	2.00	1.50	1.67	2.00	1.67
Ethiope East	31	1.53	1.88	1.59	2.21	1.93	2.15	1.76	1.52
Ethiope West	4	1.67	2.00	1.75	2.25	1.50	2.00	1.75	1.75
Ika North-East	7	1.57	1.50	2.00	2.00	2.14	2.29	2.00	1.33
Ika South	4	1.67	2.00	2.00	2.25	1.75	2.33	2.00	1.50
Isoko North	5	1.80	2.20	1.20	2.60	2.00	2.40	1.60	2.00
Isoko South	2	2.00	1.00	2.00	1.50	1.00	1.00	1.50	2.00
Ndokwa East	1	2.00	1.00	1.00	2.00	2.00	2.00	2.00	1.00
Ndokwa West	3	2.00	2.00	1.67	2.67	1.67	2.33	2.00	1.67
Okpe	5	1.50	1.60	1.60	1.75	1.40	1.80	2.00	1.50
Oshimili North	4	1.25	2.00	1.50	1.75	1.50	1.75	2.00	1.50
Udu	4	1.50	1.67	2.00	2.00	2.00	2.00	1.25	1.67
Ughelli North	4	1.75	2.50	1.75	2.25	2.00	2.00	1.75	2.00
Ughelli South	6	1.40	2.17	1.83	2.00	2.00	1.67	1.50	1.40
Ukwani	2	1.00	1.00	1.50	2.00	2.00	1.00	1.50	1.00
Uvwie	2	1.50	1.50	1.00	2.00	2.00	2.00	1.00	1.50
Warri South-West	2	1.50	1.50	1.50	2.00	3.00	2.50	1.50	2.50
Warri South	1	2.00	3.00	2.00	3.00	3.00	2.00	2.00	2.00

Table 2. Summary of score (grouped median) for adaptive capacity (AC) criteria from stakeholders (community members from local government areas in Delta State)
AC1-economic wealth, AC2-education & knowledge, AC3-access to information/services, AC4-alternative sources of livelihood, AC5-health care, AC6-political will, AC7-response agencies, AC8-kinship network

Weights were elicited from two groups of experts. The first group of experts were used to determine weights for the AC criteria. A summary of the weights and normalized values are presented in Table 3, including mean and standard deviation to show the variation among experts. This result indicates that economic wealth (AC1) constitutes the most important factor, while access to information/services (AC3) is the least important.

ACs	Experts			Grouped Median	Mean	SD	Normalized Weights
	Operators	Developers	Planners				
AC1	4.83	4.00	4.33	4.40	4.28	0.90	0.139
AC2	4.00	4.00	4.50	4.15	4.06	0.94	0.131
AC3	3.50	3.00	3.67	3.40	3.37	1.21	0.107
AC4	4.00	4.00	3.67	4.00	3.68	0.89	0.126
AC5	3.67	4.00	4.00	4.09	3.79	1.23	0.129
AC6	4.67	3.50	3.25	4.08	3.78	1.35	0.129
AC7	4.40	3.17	4.50	3.90	3.79	1.27	0.123
AC8	3.67	3.60	4.00	3.71	3.74	1.20	0.117

Table 3. Weight statistics and normalized values for adaptive capacity (AC) criteria from experts (petroleum related companies)

The second set of experts determined the weights for and PIA criteria and sub-criteria. For the weights (relative importance) of the main criteria and sub-criteria and the normalized weights for PIA the grouped median values of the experts were used (Tables 4 and 5). The threat factor was adjudged the most important.

This is due to the age of the facilities involved in oil exploration and production activities in the area and the frequency of vandalization of such facilities.

Criteria	Experts			Grouped median	Mean	SD	Normalized Weights
	Geology	Microbiology	Chemistry				
Threat	4.43	4.75	4.00	4.43	4.39	0.38	0.368
Exposure	4.20	3.75	3.50	3.88	3.82	0.35	0.322
Sensitivity	3.43	4.50	3.50	3.73	3.81	0.60	0.312

Table 4. Summary of weights statistics and normalized values for PI criteria from experts (lecturers)

Criteria	Sub-criteria	Experts			Grouped median	Mean	SD	Normalized Weights
		Geology	Microbiology	Chemistry				
Threat	T1	3.60	4.00	3.50	3.78	3.86	0.31	0.075
	T2	3.50	4.00	3.33	3.67	3.79	0.41	0.073
	T3	3.83	3.33	4.00	3.75	3.89	0.29	0.075
	T4	3.67	3.50	3.67	3.60	3.63	0.12	0.072
	T5	3.71	3.50	3.50	3.58	3.90	0.38	0.071
Exposure	E1	4.43	4.50	3.80	4.33	3.98	0.38	0.084
	E2	4.33	4.25	4.00	4.33	4.28	0.26	0.084
	E3	4.43	4.50	3.33	4.23	4.14	0.51	0.082
	E4	3.67	3.50	3.33	3.50	3.33	0.28	0.068
Sensitivity	S1	3.50	3.67	4.00	3.67	3.80	0.40	0.067
	S2	3.00	3.50	3.00	3.14	2.97	0.33	0.057
	S3	3.50	3.33	2.75	3.22	3.10	0.34	0.059
	S4	3.60	4.50	3.33	3.80	3.78	0.45	0.070
	S5	3.60	4.33	2.75	3.50	3.79	0.72	0.064

Table 5. Summary of weights statistics and normalized values for sub-criteria from experts (lecturers) for potential impact
T1-Type of facility, T2-Age of facility, T3-Type of spill/emission, T4-Volume of spill, T5-Area coverage, E1-Distance to rural settlements, E2-Distance to agricultural lands, E3-Distanc to surface water bodies, E4-Distance to forest, S1-Size of population affected, S2-Topography/slope, S3-Soil type, S4-Depth to water table, S5-Geology/geomorphology

By 'modeling' these Boolean images representing the constraints, only those cells that satisfy all constraints (non-zero) will be considered in the allocation. Those cells that have at least one zero value (because of at least one constraint not being satisfied), will have a zero multiplicative value, and hence, it is assigned a zero suitability (Mendoza, 1997), or as in this study zero vulnerability.

This information was calculated spatially with PIA values ranging from 0-1. Results of PI for the Local Government Areas assessed are displayed in Figure 6.

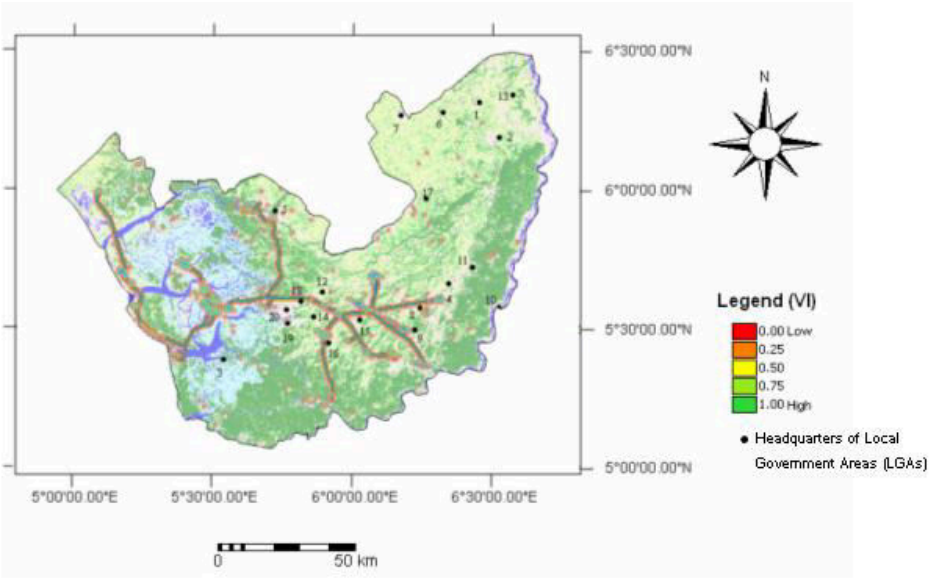

Figure 6. Aggregated spatial information for estimation of potential impact assessment

Table 6 displays the PIA and AC values for the inhabitants in the local government areas (LGAs) investigated, while the resulting vulnerability of inhabitant as a result of their close proximity to oil production and transportation facilities is presented in Figure 7.

S/N	LGAs	Headquarters	PIA	AC
1	Aniocha North	Issele-Uku	0.80	0.42
2	Aniocha South	Ogwashi-Uku	0.80	0.37
3	Burutu	Burutu	0.20	0.49
4	Ethiope East	Isiokolo	0.60	0.43
5	Ethiope West	Oghara	0.90	0.37
6	Ika North-East	Owa Oyibo	0.80	0.39
7	Ika South	Agbor	0.60	0.40
8	Isoko North	Ozoro	0.60	0.35
9	Isoko South	Oleh	0.60	0.49
10	Ndokwa East	Aboh	0.20	0.47
11	Ndokwa West	Kwale	0.80	0.38
12	Okpe	Orerokpe	0.60	0.43
13	Oshimili North	Akwukwu-Igbo	0.60	0.44
14	Udu	Asaba	1.00	0.44
15	Ughelli North	Ughelli	1.00	0.36
16	Ughelli South	Otu-Jeremi	0.30	0.41
17	Ukwani	Obiaruku	0.50	0.57
18	Uvwie	Effurun	0.90	0.51
19	Warri South-West	Ogbe-Ijoh	0.20	0.34
20	Warri South	Warri	1.00	0.25

Table 6. Calculated potential impact assessment (PIA) and adaptive capacity (AC) values for local government areas (LGAs) in Delta State

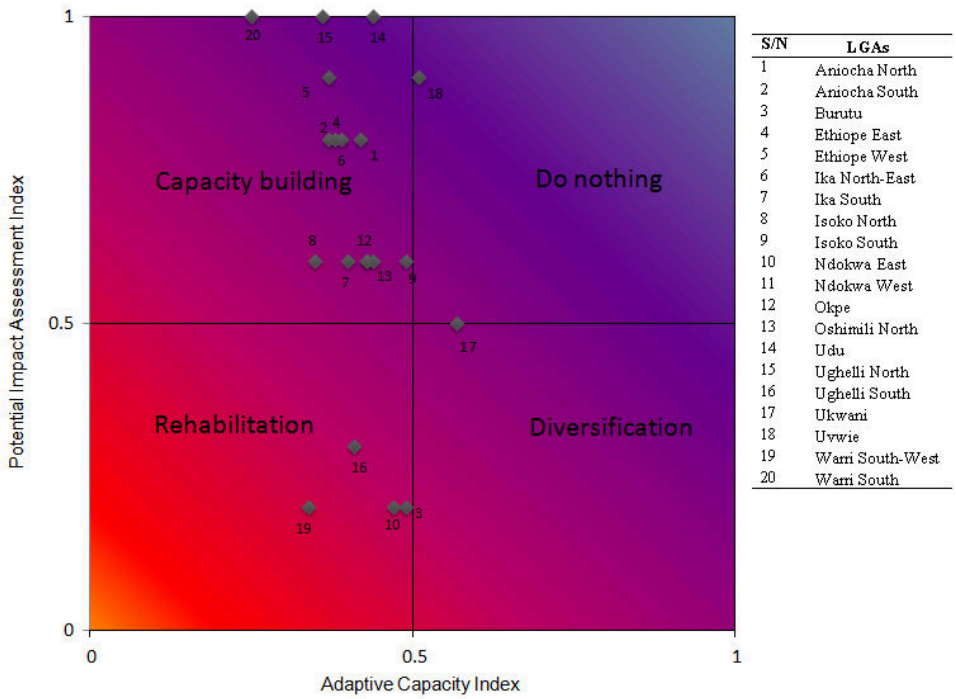

Figure 7. Options for human vulnerability assessment of inhabitants of an oil-producing region in Nigeria

The results in Figure 7 show that majority of the investigated areas fell within the region requiring capacity building and secondarily, rehabilitation.

6. Discussion

The results obtained from this study is in consonance with other studies which are based on geochemistry of groundwater (Olobaniyi *et al*., 2007; Omo-Irabor *et al*., 2008), ecological degradation (Twumasi and Merem, 2006; Omo-Irabor and Oduyemi, 2007) and anthropogenic impact (Chokor, 2004). In the region, the effect of anthropogenic pollution from hydrocarbon was quantified as 37.3% for surface water (Omo-Irabor *et al*., 2008). Contamination from hydrocarbon sources constitutes a significant portion of this fraction and arises from oil spills. According to Egberongbe *et al*., (2006), Nigeria recorded 9,107 spill incidences which led to leakage of about 3,121,910 barrels of oil into the environment between 1976 to 2005. Of all this, 50% of the spills have been attributed to oil facility corrosion, as a result of aging. 28% and 21% were adduced to sabotage and oil production operations respectively (Nwilo and Badejo, 2006). These spills unleash untold hardship on the local inhabitants and degrade the ecosystem. Consequently, it fuels agitation in the region which in turn limits the operating capacities of the producing companies.

This study indicates that two corrective measures firstly, capacity building (which accounts for 70%) among the local populace is immediately imperative to stem the tide of the environmental degradation currently experienced as a consequence of oil exploration and transportation activities in the area. Capacity building factors would include the provision of good quality education that can adequately enhance the knowledge of the people, and economic empowerment that will provide the wherewithal to respond to disaster. Others include the provision of good health care facilities and access to early and accurate information about oil spills and related disasters.

The second corrective measure, namely rehabilitation will involve remediation and clean-up of polluted lands. This will involve the active participation of the oil producing multinational companies and the relevant government institutions that are the response agencies. Recent United Nation's report suggests that several billion dollars would be needed to remediate polluted lands in the Ogoni district of Niger Delta (UNEP, 2011). This remediation exercise has not commenced, but it is expected to be executed by the oil firms responsible for operations in the area. This will require good governance and the much needed political will to enforce.

This work has revealed the significance and capability of GIS and MCDA in assessing the impact of man's activities on the environment. Although this study has focused on coastal inhabitants in an oil-producing environment, the framework developed is broadly applicable to other activities that involve interactions among host communities, the environment and natural resources extractive industries.

The result of this study indicates that the area of research has been substantially impacted by oil exploration activities and therefore requires intensive rehabilitation. Other options of less importance are diversification and capacity building which require alternative means of revenue generation base and empowerment of rural inhabitants. These are subject to government policy and regulations.

7. Conclusion

A novel approach for vulnerability assessment of rural inhabitants to the impact of industrial activities such as oil exploration has been presented in this study. The assessment technique utilizes a framework that combines GIS and MCDA to identify four options that could be applied by environmental planners and decision makers. The use of indicators/criteria for sustainable development was applied for the vulnerability assessment of inhabitants. This gives the framework a multi-disciplinary approach to decision making. The integration of different analytical tools and techniques such as GIS not only exposes the importance for the integration of different types of data, it also adds a spatial dimension to the vulnerability assessment. The incorporation of SMCA into the GIS made the analysis of spatial information feasible for vulnerability assessment.

The result of this study indicates that the area of research has been substantially impacted by oil exploration activities and therefore requires intensive capacity building for local inhabitant and rehabilitation of polluted sites. Other option of less importance is diversification which requires alternative means of revenue generation base and empowerment of rural inhabitants. These are subject to government policy and regulations.

References

Accorsi, R., Zio, E. and Apostolakis, G.E., (1999), Developing Utility Functions for Environmental Decision Making. *Progress in Nuclear Energy,* 34: 387-411.

Aprioku, I.M., (2003), Oil-spill Disasters and the Rural Hazardscape of Eastern Nigeria. *Geoforum,* 34(1): 99-112.

Balasubramaniam, A., Boyle, A.R. and Voulvoulis, N., (2007), Improving Petroleum Contaminated Land Remediation Decision-making Through the MCA Weighting Process. *Chemosphere,* 66(5): 791-798.

Chokor, B.A., (2004), Perception and Response to the Challenge of Poverty and Environmental Resource Degradation in rural Nigeria: Case Study from the Niger Delta. *Journal of Environmental Psychology,* 24: 305-318.

Delgado, O.B., Mendoza, M., Granados, E.L. and Geneletti, D., (2008), Analysis of Land Suitability for the Siting of Inter-municipal Landfills in the Cuitzeo Lake Basin, Mexico. *Waste Management,* 28(7): 1137-1146.

Egberongbe, F.O.A., Nwilo, P.C. and Badejo, O.T., (2006), Oil Spill Disaster Monitoring along Nigerian Coastline. Available from www.fig.net/pub/fig2006/papers/ ts16/ts16_02_egberongbe_etal_0223.pdf [Accessed 4 September 2007].

Epps, J.M., (1997), The Social Agenda in Mine Development. *Industry and Environment,* 20(4): 32-35.

Ercanoglu, M., Weber, K. T., Langille, J. and Neves, R., (2006), Modeling Wild Land Fire Susceptibility Using Fuzzy Systems. *GIScience and Remote Sensing,* 43: 268-282.

Geneletti, D. and van Duren, I., (2008), Protected Area Zoning for Conservation and Use: A Combination of Spatial Multi-criteria and Multi-objective Evaluation *Landscape and Urban Planning,* 85(2): 97-110.

Giupponi, C., (2007), Decision Support Systems for implementing the European Water Framework Directive: The MULINO approach. *Environmental Modeling & Software,* 22(2): 248-258.

Hinkel, J., (2011), Indicators of Vulnerability and Adaptive Capacity: Towards a Clarification of the Science-policy Interface. *Global Environmental Change,* 21: 198-208.

Intergovernmental Panel on Climate Change (2001), Technical Summary: Climate Change 2001: Impacts, Adaptation, and Vulnerability. A Report of Working Group II of the Intergovernmental Panel on Climate Change.

Linkov, I., F.K., S., Kiker, G., T.P., S., Bridges, T., Gardner, K.H., Rogers, S.H., Belluck, D. A. and Meyer, A., (2006), Multi-criteria Decision Analysis: A Comprehensive Decision Approach for Management of Contaminated Sediments. *Risk Analysis,* 26(1): 61-78.

Mendoza, G. A., (1997), Multi-criteria Approaches to Land Suitability Assessment and Allocation. *In Proceedings of the Seventh International Symposium on Systems Analysis in Forest Resources*, Traverse City MI.

Metzger, M.J., Rounsevell, M.D.A., Acosta-Michlik, L., Leemans, R. and Schroter, D., (2006), The Vulnerability of Ecosystem Services to Land Use Change. *Agriculture, Ecosystems & Environment,* 114(1): 69-85.

Nwilo, P.C. and Badejo, O.T., (2006), Impacts and Management of Oil Spill Pollution along the Nigerian Coastal Areas. Administering Marine Spaces: International Issues. A publication of FIG Commissions 4 & 7 Working Group 4.3, Available from www.fig.net/pub/figpub/pub36/chapters/chapter_8.pdf [Accessed 14 March 2007].

Okereke, J. N., Obiekezie, S.O. and Obasi, K.O., (2007), Microbial Flora of Oil-spilled Sites in Egbema, Imo State, Nigeria. *African Journal of Biotechnology,* 6(8): 991-993.

Olajire, A.A., Altenburger, R., Küster, E. and Brack, W., (2005), Chemical and Ecotoxicological Assessment of Polycyclic Aromatic Hydrocarbon-contaminated Sediments of the Niger Delta, Southern Nigeria. *Science of the Total Environment,* 340: 123-136.

Olobaniyi, S., Ogala, J. and Nfor, N., (2007), Hydrogeochemical and Bacteriological Investigation of Groundwater in Agbor Area, Southern Nigeria. *Journal of Mining and Geology,* 43(1): 79-89.

Omo-Irabor, O.O. and Oduyemi, K., (2007), A Hybrid Image Classification Approach for the Systematic Analysis of Land Cover (LC) Changes in the Niger Delta Region. *5th International Symposium on Spatial Data Quality (ISSDQ), 13-15 June, 2007, ITC, Enschede, Netherlands*.

Omo-Irabor, O.O., Olobaniyi, S.B., Oduyemi, K. and Akunna, J., (2008), Surface and Ground Water Quality Assessment using Multi-variate Analytical Methods – A Case Study of the Western Niger Delta. *Journal of Physics and Chemistry of the Earth,* 33(8-13): 666-673.

Osuji, L.C. and Nwoye, I., (2007), An Appraisal of the Impact of Petroleum Hydrocarbons on Soil Fertility: The Owaza Experience. *African Journal of Agricultural Research,* 2(7): 318-324.

Oyeike, E.N., Ogbuja, S.I. and Nwinuka, N.M., (2002), Inorganic Ion Levels of Soils and Streams in Some Areas of Ogoniland, Nigeria as Affected by Crude Oil Spillage. *Environ. Monitoring Assessment,* 73(2): 191-204.

Research and Assessment Systems for Sustainability Program (2001), Vulnerability and Resilience for Coupled Human-Environment Systems: Report for the Research and Assessment Systems for Sustainability Program 2001 Summer Study. Airlie House, Warrenton, Virginia, Cambridge, MA: Environment and Natural Resources Program, Belfer Center for Science and International Affairs, Kennedy School of Government, Harvard University.

Roberts, J. and Crawford, D., (2004), Developing a Framework for Assessing Oil Spill Consequences: The Application of Oil Spill Sensitivity Analysis in New Zealand. *Spillcon*.

Smit, B. and Wandel, J., (2006), Adaptation, Adaptive Capacity and Vulnerability. *Global Environmental Change,* 16(3): 282-292.

Turner, B.L., Kasperson, R.E., Matson, P.A., McCarthy, J.J., Corell, R.W., Christensen, L., Eckley, N., Kasperson, J.X., Luers, A., Martello, M.L., Polsky, C., Pulsipher, A. and Schiller, A., (2003), A Framework for Vulnerability Analysis in Sustainability Science. *Proceedings, National Academy of Sciences,* 100(14): 8074-8079.

Twumasi, Y.A. and Merem, E.C., (2006), GIS and Remote Sensing Applications in the Assessment of Change within a Coastal Environment in the Niger Delta Region of Nigeria. *Int. J. Environ. Res. Public Health,* 3(1): 98-106.

UNEP (2011), Environmental assessment of Ogoniland. United Nations Environmental Programme. Available from www.unep.org/nigeria [Accessed 15 August 2013].

von Winterfeldt, D. and Edwards, W., (1986), Decision Analysis and Behavioral Research. Cambridge, UK, Cambridge University Press.

CHAPTER 10

Resource-Constrained Agriculture in Developing Countries and Where Geo-ICT Can Help

Clarisse Kagoyire and Rolf A. de By

Faculty of Geo-Information Science and Earth Observation, University of Twente, The Netherlands

c.kagoyire@utwente.nl; r.a.deby@utwente.nl

Abstract

In many African countries, agriculture is constrained by resource scarcity such as in soil nutrients, seeds or plantlets, pesticides, fertilizers, water, weather conditions, labor, transport logistics, arable land, or access to credit. Each agricultural production stage connects with a set of prominent resource constraints, which are commonly location-specific, varying from one farm to another. In this chapter, we discuss the resource constraints encountered with on-farm agricultural activities, from the choice of seeds to the time of harvest. Our aim is to identify which Geospatial information must be put in action to improve farm production, in quantity or quality. To determine the predominant resources that limit a particular farm income, we propose a framework for constraint-based farm-plot profiles, in the context of limited access to ICT services. Exemplified by a case study of coffee farming in Rwanda, we specify a number of agricultural resources that potentially constrain coffee production. Besides traditional sources of agricultural information, our framework taps into another source that is often ignored: the farmers themselves.

KEYWORDS: Resource-constrained farming, geo-ICT, information flow management, computer-supported farm cooperative

1. Introduction

Much of the agricultural sector in Africa is *resource-constrained*, meaning that farming activities suffer directly from a lack of labor force, land parameters, machines, materials, or finance. Agriculture is an important industry for Africa, as it directly

supports a high percentage of the population, and that in times when urban populations, over the world, are on the rise. The sector therefore has the potential of growth and more sustainable income for a substantial portion of the population. In the country of Ethiopia, with its population of 91 million, 85% of the national GDP is earned in agriculture, and it boosts a labor force of approximately 30,000 agro-extension workers. Also here, much of the farming is resource-constrained, and Ethiopia could increase its farm production considerably if those constraints were understood and could subsequently be relaxed. But there is no single solution that fits all situations, and we need to grow our understanding of location as a parameter to such solutions.

Many of the resource constraints that are in play exhibit correlation with location, and hence can (no: must) often be looked at as spatial parameters. The data to inform about such spatial dependencies is available from various sources, or can be made available for exploitation with moderate effort. Information is itself also a resource, which happens to be often constrained at the farm, community or cooperative level. A lack of information must be addressed just like any other resource constraint. Our perspective is that of farming as a production process requiring inputs, and delivering produce, while depending on logistical processes also, as illustrated in Figure 1.

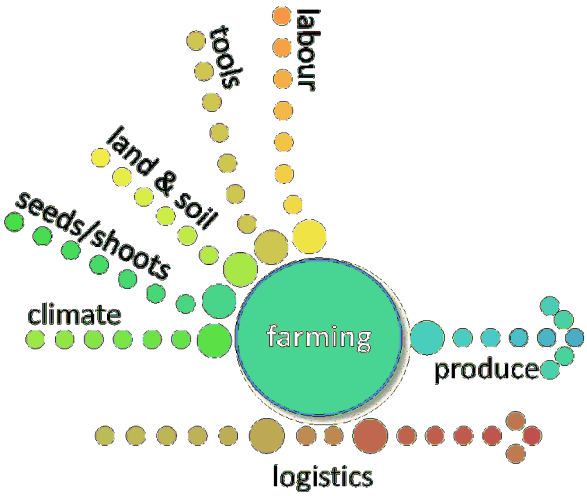

Figure 1. Agricultural supply chain

Farm production processes show high resource variability in space and time. At the farm plot scale, topography, soil and climatic conditions are important yet variable, at the household scale labor and practices, land ownership, access to machines and finance differ considerably, while at the community or cooperative scale access to logistics such as seed and produce stores, processing stations, and transportation varies widely. At national and international levels, policy and financial parameters play important roles, and differ from one place to the next. As a consequence, no single farm production process is like the next in practice. In a landscape of physical and

economic dynamics, farmers need all the inventiveness they can muster, and all the supportive information that can be availed to them. This is especially important in times when those parameters change in orders of magnitude that these communities have not witnessed before. Climate change, urbanization, globalizing markets, growing consumer awareness are such fundamental processes that carry the potential to display large changes, to which farmer communities must respond timely and with wisdom.

Whilst the physical and financial parameters in farming are naturally the most prominent and important ones to consider, the longer-term perspective on sustainable farming practices needs to take into account for each farmer community what is its pathway into a more secure economic position. It is generally believed that such a pathway has three important stepping stages:

1. *informal production process*, in which the farmer community has few, if any, external connections and produces crop primarily for subsistence and, secondarily, for local market sales;
2. *production process in a formalized supply chain*, in which agreements exist for delivery of goods, whether farm inputs or products;
3. *production process in a formalized value chain*, in which typically multi-year agreements exists for delivery of goods, with stated quality and quantity parameters.

A key difference between these stages is the role that information plays in support of the production process and the sources for that information.

2. Geo-ICT Potential for Sustainable Agriculture in Developing Context

Increased availability of ICT globally, also in developing economies, has led to the development of various information services, catering for a wide range of users. In agricultural sector, ICT is progressively adopted to allow better exploitation of resources and improve farm management. When it comes to Geospatial information, as derived most prominently from remote sensors and expansive farm surveys, the information mostly informs governance, and little of it reaches the hands of extension workers, let alone farmers or farmer communities. Where such information does reach the governance stakeholders, it often comes late, and after the fact, having lost some potential impact already. Moreover, occasionally information is hard to interpret well, due to a lack of ground truthing.

Adequate information provision to farm communities and information exchange amongst them within understood time limits is an important requirement to sustainable agriculture, especially in developing countries. Agriculture is sustainable if the required inputs to the farming process can continuously be replenished, and the produced outputs generate enough income. This often translates into an environmentally friendly production process, which satisfies the socio-economic needs

of the community. Given the environmental variability (climatic and topographic) from one location to another and the socio-economic disparities (like market access, market prices, transport), location is the key.

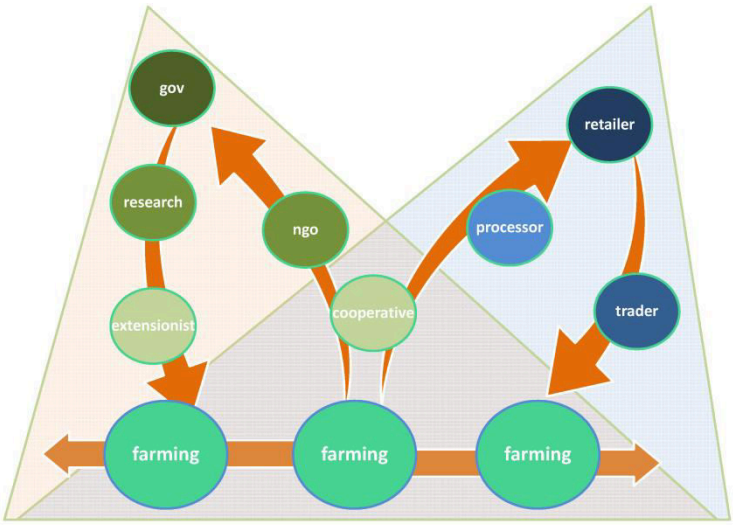

Figure 2. The governance and commercial pyramid in farming

We argue thus that location is an important factor in agriculture processes, and call for the exchange of relevant geo-information. The options are far from exhausted: much can be gained in the on-going network-and-mobile revolution, in terms of outreach and timeliness of information delivery, especially, of *downstreaming* information to extension workers, cooperatives and farmers themselves. More and more, we are seeing examples where *upstreaming* of (geo)information, collected by these three types of stakeholder, is built into community networks, allowing timely informing of stakeholders elsewhere in the governance pyramid, the financial process chain, and/or in the commercial value chain (Figure 2).

To address these information needs, geo-ICT (geo-information communication and technology) must be adopted (Bill *et al*., 2012), as it is the principal technology by which the important parameters can be brought together. Numerous previous studies have demonstrated its use to improve agricultural practices, while preserving the natural environment (Rao *et al*., 2000; Ahmad and Rai, 2002; Obade and Lal, 2013; Forkuor *et al*., 2013), and approaches have been proposed to support farmers in their decision making, through the use of geo-information (Sudharsan *et al*., 2009; Tian-en *et al*., 2009; Xie *et al*., 2012; Lan, 2012; Venus *et al*., 2013).

2.1 Agricultural Resource Constraints and Their Spatial Variability

Paradoxically, of those people globally suffering from famine, 70% is a farmer. If their productivity does not even meet their own food demands, how can they contribute to feeding increasing urban populations? An important challenge for smallholder farmers is the proper management and use of the limited agricultural resources, amongst others by reduction of input cost, while maintaining and improving production.

We typify the important agricultural constraints in the following scheme:

Physical constraints involve the soil and its characteristics, the seeds, shoots and plantlets from which the growth starts, materials added to the soil or plant for growth sustenance, such as fertilizers, mulch and pesticides, and further physical inputs including climatic conditions, especially sunshine and precipitation. Many of these directly vary with location, while others are location-dependent because of correlation with those direct locational parameters. For instance, the use of fertilizer depends on the soil type. Direct locational parameters are already mapped, and spatial models are needed to describe the indirect locational dependencies.

Support constraints involve the labor of people on the land, or as the season progresses, away from it, following the harvested crop. This should explicitly bring into the game, machine availability and skilled operators. A special case is transport logistics, both when feeding inputs to the farm and when transporting harvest to market. Labor and machine availability is a spatiotemporal phenomenon that sometimes forms an important obstacle to improvement, and which then should be mapped out.

Economic constraints involve parameters of ownership, permits and agreements, access to financial and insurance support and access to markets. Some of these parameters are spatial but at less local scale than the parameters discussed earlier. This actually makes them easier to map, but we have not seen much effort on this in the context of optimizing agricultural production in the global South. At the same time, tremendous results have been achieved with provision of market information to farmers.

Knowledge constraints affect directly the know-how and skills of the farmer community. Historically, know-how is passed on within communities by a heritage system, and the governance pyramid has with different levels of success also worked to bring novel knowledge to these communities. For the latter, typical condensed knowledge is found in textbooks for the respective branches of agriculture (Wintgens, 2004). Substantial disparities between (often unwritten) community knowledge and that textbook knowledge exist, however, one cannot be judged better than the other. Community knowledge often is highly location-specific, and is a greatly underexploited resource.

For maximal resource exploitation, farmers need various types of information, at different stages. For instance, to choose a crop variety, information on low-cost seed acquisition, taking into account transport costs, and compatibility of seed variety with local soil and weather. The same holds for acquisition of fertilizers and pesticides. To lower the transport cost, a farmer needs location intelligence on agricultural input, in relation to her farm fields.

2.2 ICT Challenges in the Agricultural Sector

ICT can help improve rural farmer income, by facilitating agricultural information dissemination as recognized in previous studies (Thysen, 2000; Niederhauser *et al*., 2008). However, its adoption in the agricultural sector is slow: Kulhlmann (1999) argues that farmers are reluctant to adopt information technology to reduce investment costs.

Beside financial costs, a low level of education and limited infrastructure also constrain technological penetration in developing countries. A number of information channels is used in agriculture, notably through extension workers, using face-to-face communication, booklets, radio broadcasts, television shows (dedicated to inform farmer communities), the Internet or cell phones, farmer-to-farmer communication (word-of-mouth), or by accessing existing traditional agricultural data repositories. Though such technology uptake may appear slow in places, there is no stopping it, and it will find its way even to the remotest of farm locations. It will equip society with means to inform farmers, be informed by them, and it will allow them to inform each other.

2.3 Information Chains in Agriculture

Various information chains can be recognized in a technologically equipped farming society. Under the mantra that relevant information empowers every agricultural stakeholder, one needs to identify the most feasible and useful information flows. Farmer communities can benefit from more mature *downstreaming* mechanisms of geo-information on the constraint types described in Section 2.1. This is the classical application of SDI, with a clear clientele of farmer communities and affiliated extension workers, who must be well-trained in interpreting such information.

Once farmers evolve their farm business, information chains should start to more formally parallel the production chain. This is needed for trustworthy stakeholdership in supply-chain mechanisms, where agreements become more formal, along with the quality/quantity/consistency of crop production. The agreements force information into existence, and farmer-representatives will find out that local data collection helps towards more consistent decisions. Soon, such data will find its way *upstreaming* the various channels: to the cooperative, the extension worker, the district or national policy agency or NGO, but also to the bank and the insurance company. Local information at that stage has become valuable: better finance or insurance conditions

can be negotiated for the next season on that very basis, and better prices can be obtained for the crop.

Within communities there commonly is no competition between producers, and farmers are often organized in cooperatives. These are ideal conditions for a third information stream, which we call *sidestreaming,* sometimes also coined as 'farmer-to-farmer'. Technological trends allow us to build systems that offer sidestreaming, which in itself is an exciting opportunity to improve exploitation of community knowledge (see Figure 2). This information is not directly spatially explicit, however, the farm or farmer as information origin commonly provides enough locational understanding.

All agricultural stakeholders, involved from the crop choice stage to the harvest and trade stage, hold essential information to improve production quality and quantity. That information can be exploited: farmers in the global South need access to environmental information to sync farm activities with the environment, with advices on farming practices, and to optimally exploit the limited resources. On the other hand, agricultural experts need information about on-farm activities, to provide better crop models, and adequate advisory services to farmers, to develop suitable agricultural techniques or improved seeds, based on local environmental parameters. The consumers of the agricultural products also need to be insured of the quality, origin and fair prices of the products, through documentation of the food production process.

3. Case Study: Coffee Farming in Rwanda

Agriculture is an important sector for Rwanda's economy, as it contributes 43% of GDP and engages about 80% of the labor force. Coffee is one of the main agricultural exports in Rwanda, 98% of the production is of type Arabica. About 500,000 households grow coffee as their main income crop, working together in agricultural cooperatives. Rwandan produce is currently recognized as a specialty coffee and has gained interest in international markets. According to a World Bank report, the average unit price of Rwandan coffee increased by 51%, between 2006 and 2010 (World Bank, 2011). Given Rwanda's topology, high population density, and low average household income, however, coffee farming is still constrained by a lack of resources. We propose a framework for constraint-based farm-plot profiles (Figure 3), to determine those resources that limit particular farm household incomes most. In Sections 3.1 and 3.2, we look at which resources constrain coffee farming in Rwanda, and demonstrate how geo-ICT can be used to support farmers.

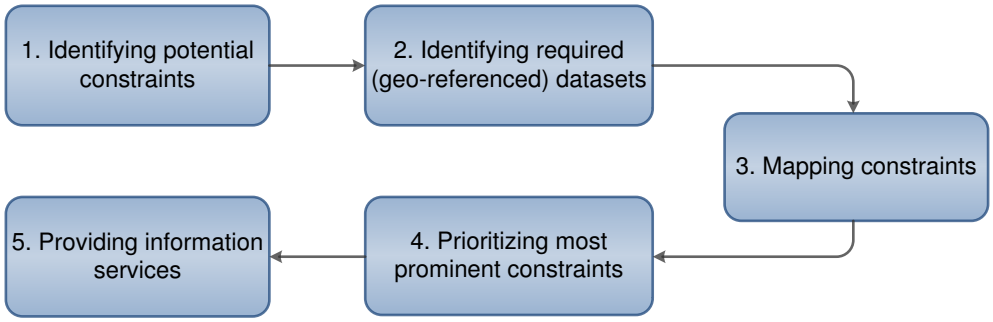

Figure 3. Framework for prioritizing location-specific constraints

3.1 Agricultural Resource Constraints in Rwandan Coffee Farming

Farm production depends on biophysical, financial and human resources. From literature review and fieldwork in Rwanda, we elicited a list of constraints to coffee farming. We subsequently identified the (geo-referenced) data required to inform about these limitations, given that most of the identified constraints are region-specific or vary between households.

a. Soil nutrients

Soil nutrients constrain coffee farming, fundamentally. Rwanda's mountainous topology causes 40% of the arable land to be at high erosion risk, i.e. with chances of soil degradation and nutrients loss. Due to a growing population, coffee plots are more scattered and fragmented. This compels farmers to over-cultivate lands by new intercropping systems, shortening fallow time and causing soil degradation. Steeper slopes are being cultivated, increasing soil erosion.

b. Coffee seedlings

Seedling production is carried out by cooperatives with the support of extension workers. The seedlings are subsequently distributed to members. Usually, seedling acquisition is not a constraint, though sometimes the distance between nursery plots and farmstead is a concern.

c. Pesticides and fertilizers

The access to and utilization of these chemicals is a financial resource and expertise constraint, given limited knowledge on impact on plants, and lack of information on how and where to purchase at a fair price (including transport costs).

d. Weather conditions

Coffee is sensitive to extreme weather, and changes in temperature, hours of sunshine, atmospheric humidity, rainfall, and wind may constrain plant growth. The lack of adequate weather information hinders farmers' decision making.

e. Transportation

Limited or inadequate transport infrastructure is an important constraint to coffee farming, especially during pre- and post-harvest activities. Post-harvest activities include the transport of cherries from plot to the coffee washing station (CWS), and then from there to dry mill, from which the produce is transported to traders.

f. Labor

The labor force demand varies with the size of plots and their location (in relation to farmstead or cooperative headquarters), and the farming activity being carried out.

g. Land

Limitations in available arable land result in small-sized coffee plots, entailing over-cultivation, cultivation on steep slopes, thus increasing the risk of soil erosion.

h. Financial resources

Limited financial resources are an important drawback for coffee farming, which requires investments. Moreover, low levels of education and lack of existing microfinance information cause only few farmers to benefit from microfinance.

i. Information on coffee farming practices

Lack of adequate information at household level, cooperative level or extension worker level, impedes the process of coffee production. The farmer needs to know the status of her plots and their environment, to make sensible decisions during required agricultural activities, and how those activities should be carried out. Moreover, cooperative managers and extension workers also need to stay updated on farmer concerns and farming activities.

3.2 Identifying Data Sources

A number of agencies provide geo-referenced data for different themes. Notably, the world data center for soils ISRIC, through the AfSIS project (Africa Soil Information Service) offers a soil profile database (Leenaars, 2013) and the Harmonized World Soil Database (FAO *et al*., 2009). The road network datasets used for our case study were acquired through CGIS-NUR (Centre for GIS and Remote Sensing of the National University of Rwanda) (Akinyemi and Kagoyire, 2010). Weather, elevation and land cover data help to determine the vulnerability to soil nutrient loss. These data are acquired from the EOS Data and Information System at NASA (National Aeronautics and Space Administration), GeonetCast and the eModis products generated by the Earth Resources Observation and Science at USGS (NASA, 2013; GeonetCast, 2013; Jenkerson and Schmidt, 2010).

Seeds, pesticides and fertilizers are mostly constrained by acquisition cost, transportation infrastructure, timing of acquisition in relation to crop calendar, and lack of usage monitoring. The World Bank data on GDP provides information that is used to estimate financial capacity of farmers compared to the acquisition cost of agri-inputs. Though this information is at country or regional level, and financial capacity

varies between households, it can be used as proxy. For our study, data on acquisition timing and usage monitoring, come from farmers themselves. Our application (described in Section 4) allows farmers to share information on the use of agri-inputs per farm plot.

To evaluate the impact of the transportation on agricultural production, a roads dataset from CGIS-NUR is used. Transportation cost is also dependent on farmstead and farm plot location. Once again, we take advantage of farmer community to obtain this information. Furthermore, to determine whether labor or financial resources are limiting factors, the data on population demography, employment or GDP are required. Absence of higher resolution data at household scale, forces us to apply global data on population distribution and other socio-economic parameters, such as in LandScan (Bright *et al*., 2012) or AfriPop (AfriPop, 2013).

3.3 Mapping Agricultural Resource Constraints

Some of the above agricultural resources present direct location dependency, notably soil nutrients, weather, transportation, labor and land. The remaining resources have more indirect location dependency. Soil properties and weather conditions have the strongest spatial dependencies as they vary with topography. Volcanic soils with a minimum depth of one meter are most suitable for coffee farming, while steep slopes are most vulnerable to soil nutrients loss (Verdoodt and van Ranst, 2003). Physical suitability map of soil nutrients for coffee farming can be generated from soil profiles and slope gradient (Figures 4 and 5). Furthermore, the amount of pesticides/fertilizers applied per coffee plot is recorded by farmers. The resulting data is later used to assess their impact on soil.

Climate parameter maps help determine plots vulnerable to weather conditions. Here, we need to verify that annual rainfall is between 1500-1800 mm, and temperature is between 20-25^{0}C. Spatial variability of transportation constraints is derived from road datasets, CWS location and coffee plots. GIS-buffering around road and CWS objects helps determine proximity to these objects. Transportation is a possible constraint for plots away from buffer areas (Figure 6); the location of farmstead – if available – is also taken into account.

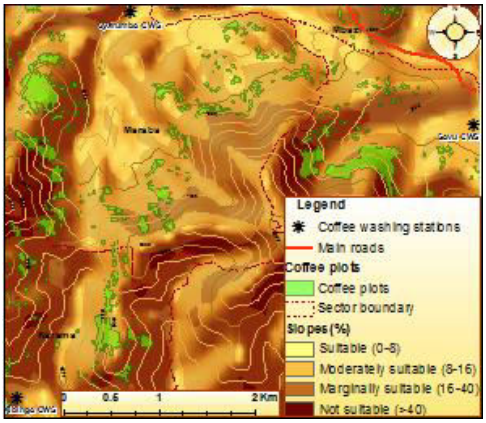

Figure 1. Topography suitability of coffee farming

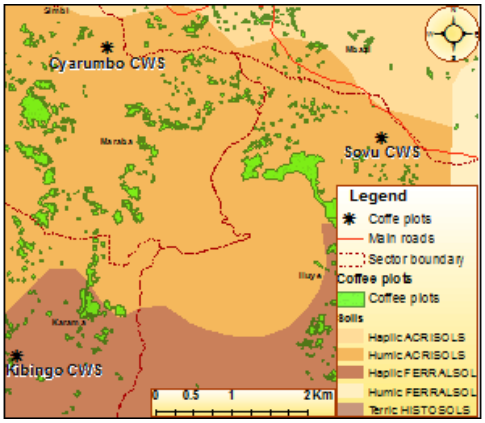

Figure 5. Mapping soil types

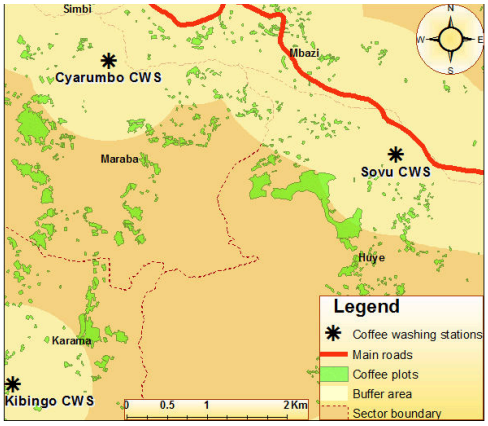

Figure 2 . Mapping the proximity of coffee plots to roads and CWSs

3.4 Prioritizing Most Prominent Location-specific Constraints

Location-specific problems require location intelligence. The first step of our prioritization process is to identify a short list of nine coffee farming constraints (Section 3.1), based on our fieldwork and literature review (Ndiaye and Sofranko, 1994; Byiringiro and Reardon, 1996; Verdoodt and van Ranst, 2006). Soil fertility and erosion are the main constraints for Rwandan agriculture, particularly for coffee (Andre and Platteau, 1998; Ngabitsinze *et al*., 2011). The focus group discussions with Southern Rwandan coffee farmers and discussions with agricultural stakeholders provided insight into important issues in this sector, at the country and community levels. Furthermore, the identified geospatial data are analyzed to determine prominent constraints at household, and community levels.

4. Information Service Provision

Adopting ICT for development helps alleviate poverty in rural areas, particularly in the agriculture sector (Munyua *et al*., 2008; Kiiza and Pederson, 2012; World Bank *et al*., 2012; Buhigiro, 2012). Hellström (2010) reviews the use of phones in pro-poor mobile applications in East Africa for information delivery services, and provides a list of existing mobile applications in different sectors. Numerous web- or phone-based approaches were designed for the benefit of smallholders in rural areas (Talukder and Das, 2010; TATA, 2013) and review studies discuss the challenges encountered in agricultural information services (Niyongabo, 2011; Brugger, 2011; Balraj and Pavalam, 2012; World Bank, 2012).

To provide information services to coffee farmers, we implemented a web-based application to facilitate access to and exchange of information. With a simplified interface, the *Sakaza Muhinzi* (which means ‘disseminate-farmer’, in Rwanda's native language) application allows farmers to locate and identify their coffee plots, record on-farm activities, and provides them with environmental plot information (like soil properties, weather parameters, and topography). The information provided by farmers is used, along with other data sources, to generate constraint-based plot profiles.

5. Conclusion

In this study, we demonstrate the use of geo-ICT to enable farmers to actively participate in producing location-specific information on agriculture, and addressing the challenges encountered in farming such as scarcity of resources. The proposed approach taps into a new source of information that is often underexploited, particularly the local farming knowledge. This is facilitated by the downstreaming, upstreaming and sidestreaming, information flow approaches, adopted in *Sakaza Muhinzi* web application. The participation of people without GIS formal training, in geo-information provision, is recognized in previous studies as an alternative or

complementary data source (Coleman *et al*., 2009; Goodchild, 2009; Elwood *et al*., 2012).

The spatial analysis of geo-referenced data, carried out during the process of mapping resource constraints, shows that 67% of mapped plots are located in areas prone to erosion, making the soil nutrients resource the most prominent constraint to Rwandan coffee farming. However, at the household level, the soil nutrients resource is not always the most prominent constraint.

References

AfriPop (2013), *AfriPop demography*. Available from: www.afripop.org/.

Ahmad, T. and Rai, A., (2002), An approach to GIS based spatial sampling procedures for environmental studies in agriculture. *American Journal of Mathematical and Management Sciences,* 22 (3,4): 251–260.

Akinyemi Olufunmilayo, F. and Kagoyire, C., (2010), The Rwanda Metadata Portal: A Web Catalogue Service. *International Journal of Spatial Data Infrastructures Research*: 382-401.

Andre, C. and Platteau, J.-P. (1998). Land Relations under Unbearable Stress: Rwanda Caught in the Malthusian Trap. *Journal of Economic Behavior and Organization*, 34(1): 1–47.

Balraj, P. L. and Pavalam, S., (2012), Integrating ICT in Agriculture for Knowledge-Based Economy. *Rwanda Journal, Series C: Mathematical Sciences, Engineering and Technology, 27*.

Bill, R., Nash, E. and Grenzdörffer, A.G., (2012), GIS in Agriculture. In W. Kresse and D. M. Danko (Eds.), *Springer Handbook of Geographic Information* (pp. 461-476).

Bright, E.A., Coleman, P.R., Rose, A. N. and Urban, M.L., (2012), LandScan 2011. *Oak Ridge National Laboratory*. Oak Ridge, TN.

Brugger, F., (2011), *Mobile applications in agriculture.* Basel, Switzerland: Syngenta Foundation.

Buhigiro, S., (2012), *The role of telecentres in promoting socio-economic development in Rwanda.* Johannesburg, South Africa: University of the Witwatersrand.

Byiringiro, F. and Reardon, T., (1996), Farm Productivity in Rwanda: Effects of Farm Size, Erosion, and Soil Conservation Investments. *Agricultural Economics*, 127-136.

Coleman, D.J., Georgiadou, Y. and Labonte, J., (2009), Volunteered Geographic Information: The Nature and Motivation of Producers. *International Journal of Spatial Data Infrastructures Research*, 332-358.

Elwood, S., Goodchild, M. F. and Sui, D.Z., (2012), Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers,* 102(3): 571-590.

FAO/IIASA/ISRIC/ISSCAS/JRC (2009), *Harmonized World Soil Database (version 1.1).* FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Forkuor, G., Pavelic, P., Asare, E. and Obuobie, E., (2013), Modelling Potential Areas of Groundwater Development for Agriculture in Northern Ghana using GIS/RS. *Hydrological Sciences Journal,* 58(2): 437-451.

GeonetCast (2013), *Delivering Operational Services to Users*. Available at: www.earthobservations.org/geonetcast.shtml.

Goodchild, M., (2009), NeoGeography and the nature of geographic expertise. *Journal of Location Based Services,* 3(2): 82-96.

Hellström, J., (2010), *The Innovative Use of Mobile Applications in East Africa.* (P.-E. Tröften, Ed.) Stockholm, Sweden.

Jenkerson, C. B. and Schmidt, G., (2010), *eMODIS: A user-friendly data source: U.S. Geological Survey Open-File Report 2010–1055.* U.S. Geological Survey.

Kiiza, B. and Pederson, G., (2012), ICT-based Market Information and Adoption of Agricultural Seed Technologies: Insights from Uganda. *Telecommunications Policy,* 36(5): 253-259.

Kuhlmann, F., (1999), Farm Level Information Systems: Developments and Perspectives. In R. H. G Scheifer (Ed.), *Perspectives of Modern Information and Communication Systems in Agriculture, Food Production and Environmental Control*, (pp. 7–17). Bonn, Germany.

Lan, B., (2012), The Establishment of Agriculture Information System Based on GIS and GPS. In X. Qu and Y. Yang (Eds.), *Information and Business Intelligence* (Vol. 268, pp. 506-511).

Leenaars, J.G., (2013), *Africa Soil Profiles Database, Version 1.1. A compilation of geo-referenced and standardized legacy soil profile data for Sub Saharan Africa (with dataset). ISRIC report 2013/03.* Africa Soil Information Service (AfSIS) project , ISRIC – World Soil Information, Wageningen, the Netherlands.

Munyua, H., Adera, E. and Jensen, M., (2008), Emerging ICTs and Their Potential in Revitalizing Small-Scale Agriculture in Africa. *Proceedings World Conference on Agricultural Information and IT, IAALD AFITA WCCA 2008*, (pp. 707-717). Tokyo, Japan.

NASA (2013), *Earth Observing System Data and Information System*. Available from: http://earthdata.nasa.gov/.

Ndiaye, S. M. and Sofranko, A. J., (1994), Farmers' perceptions of resource problems and adoption of conservation practices in a densely populated area. *Agriculture, Ecosystems & Environment*, 35-47.

Ngabitsinze, J., Mukashema, A., Ikirezi, M. and Niyitanga, F., (2011), *Planning and costing adaptation of perennial crop farming systems to climate change: Coffee and banana in Rwanda.* London, UK: International Institute for Environment and Development (IIED).

Niederhauser, N., Oberthür, T., Kattnig, S. and Cock, J., (2008), Information and its management for differentiation of agricultural products: The example of specialty coffee. *Computers and Electronics in Agriculture,* 61(2): 241-253.

Niyongabo, R., (2011), ICT in Development: The Search for Hidden eSoko among Rural Small Scale Farmers of Rwanda. The Hague, The Netherlands: International Institute of Social Studies.

Obade, V.D. and Lal, R., (2013), Assessing land cover and soil quality by remote sensing and geographical information systems (GIS). *CATENA,* 104(0): 77-92.

Rao, M.N., Waits, D. A. and Neilsen, M. L., (2000), A GIS-based modeling approach for implementation of sustainable farm management practices. *Environmental Modelling & Software,* 15(8): 745-753.

Sudharsan, D., Adinarayana, J. and Tripathy, A., (2009), Geo-information Services to Rural Extension Community for Rural Development Planning : A Framework. *Proceedings International Conference on Advanced Geographic Information Systems Web Services* (pp. 54–59). IEEE Computer Society, Washington, DC, USA.

Talukder, A.K. and Das, D., (2010), Mobile Web for Under-privileged in Developing Countries. *Telematics and Informatics*, 350-359.

TATA Consultancy Services. (2013). *mKRISHI: Rural Service Delivery Platform*. Available from: hwww.tcs.com/offerings/technology-products/mKRISHI/Pages/default.aspx [Accessed 15 April 2013].

Thysen, I., (2000), Agriculture in the Information Society. *Journal of Agricultural Engineering Research,* 76(3): 297-303.

Tian-en, C., Li-ping, C., Yunbin, G. and Yanji, W., (2009), Spatial Decision Support System for Precision Farming Based on GIS Web Service. In Q. Zhou (Ed.), *Proceedings International Forum on Information Technology and Applications (IFITA). 2*, pp. 372-376. IEEE Computer Society.

Venus, V., Asare-Kyei, D., Tijskens, L., Weir, M., de Bie, C., Ouedraogo, S., *et al*., (2013), Development and validation of a model to estimate postharvest losses during transport of tomatoes in West Africa. *Computers and Electronics in Agriculture,* 92(0): 32-47.

Verdoodt, A. and van Ranst, E., (2003), *Land evaluation for agricultural production in the tropics: A large-scale land suitability classification for Rwanda.*

Verdoodt, A. and van Ranst, E., (2006), Environmental assessment tools for multi-scale land resources information systems – A case study of Rwanda. *Agriculture, Ecosystems & Environment*, 170-184.

Wintgens, J.N. (Ed.), (2004, October), *Coffee: Growing, Processing, Sustainable Production: A Guidebook for Growers, Processors, Traders, and Researchers* (1 ed.). Wiley-VCH.

World Bank (2011), *Rwanda Economic Update: Seeds of Higher Growth.*

World Bank (2012), *Information and Communications for Development 2012: Maximizing Mobile.* Washington, DC.

World Bank, African Development Bank and African Union (2012), ICTs for Agriculture in Africa. In *The Transformational Use of Information and Communication Technologies in Africa.*

Xie, Y.W., Yang, J.Y., Du, S.L., Zhao, J., Li, Y. and Huffman, E.C., (2012), A GIS-based fertilizer decision support system for farmers in Northeast China: a case study at Tong-le village. *Nutrient Cycling in Agroecosystems,* 93(3): 323-336.

PART THREE

SPATIAL ENABLEMENT IN SUPPORT OF DEVELOPMENT

CHAPTER 11

GIS-based Land Suitability Assessment for Optimum Allocation of Land to Foster Sustainable Development: the Case of the Special Zone of Oromia Regional State around Addis Ababa City, Ethiopia

Dessalegn Gurmessa[1] and Sileshi Nemomissa[2]

[1] United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA), Addis Ababa, Ethiopia
[2] Addis Ababa University, Addis Ababa, Ethiopia

gurmessad@gmail.com; nemomissa@yahoo.co.uk

Abstract

Land use allocation involves the process of designing an optimal mix of land uses based on their estimated suitability. Whereas land use suitability is a generic term associating a combination of factors and their impacts with respect to potential land use. The study area surrounds the city of Addis Ababa and endowed with suitable environmental variables that enhanced demand for a piece of land. The apparent rapid urbanization is straining various land uses such as crop and livestock production, forestry, wildlife conservation etc. Focused researches on suitability based land use allocation are, therefore, vital to disentangle and understand the intricacies of land management in the study area. To assist the efficient utilization of land resources to promote sustainability of natural resource bases, the study made use of Geographic Information System and Remote Sensing models to integrate spatially complex and different land attributes for performing land suitability analyses and allocations. The current analysis revealed that, 9.7% and 2.6% of the study area were classified as moderately suitable for crop and livestock production, respectively. Noteworthy is also that nearly 65% of the study area is least suitable for the production of both crop and livestock. On the contrary, Satellite image classification of the study area has shown that 58.4 % and 21.7 % of the total area is currently used for the same land utilization types in the same order. The current study also showed that, 59.2% and 8.6%, of the total study area is allocated for livestock-crop specialization and crop-livestock specialization, respectively. It should be noted that, for sustainability of land productivity, integration of pastures in to crop farming system is inevitable measure. Hence, in the crop-livestock specialization zone, the study proposed crop as main

production and livestock as supplementary production. Furthermore, 6.23% and 7.38% of the total study area was allocated for productive and protective forestry respectively; the remaining portion of the area (18.57%) was carefully considered and retained its present spatial location. Finally the study serves as a stepping stone for understanding the potentials and limitations of the land in the study area.

KEYWORDS: Land use allocation, land use, Geographic Information Systems, Land suitability analysis

1. Introduction

Land resource is limited in nature and its use is not only determined by the user but also by land capability to sustain production (FAO 1993). Land capability is governed by the different land attributes such as the types of soil, underlying geology, topography, hydrology, and etc. These attributes limit the extents of land available for various purposes and the optimum and proper utilization of its resources is inevitable. Land cover is the product of human activities changing terrestrial ecosystem and is an element of complex ecological and economic system that needs periodic evaluation. Knowledge about the optimal allocation of land is important for understanding the magnitude of maximum return. Land use suitability was variously studied (e.g. Allen *et al*., 1995) taking environmental variables such as topography, soil, vegetation and landforms into consideration. However, the integration of various variables for a single assessment cannot result in accurate and efficient results unless Geographic information system (GIS) is used. GIS has found several applications in land suitability studies (Pereira and Duckstein 1993, Steiner *et al*., 2000, Zhang *et al*., 2011 and Joerin *et al*., 2001).

2. Literature Review

The ultimate aim of GIS is to provide support for making spatial decisions (Malczewski, 1999). The GIS system contains a set of procedures that facilitate the data input, storage, manipulation and analysis, and data output to support decision-making activities (Grimshaw, 1994). The GIS capabilities for supporting spatial decisions can be analyzed in the context of the decision making process.

Malczewiski (1999) stressed the complexity in attempting to acquire data and to process the data to obtain information for making decisions. This problem may require processing at a level that exceeds a decision maker's cognitive ability. To this end, the role of GIS and multi-criteria decision making (MCDM) techniques support the decision maker in achieving greater effectiveness and efficiency of decision making while solving spatial decision problems. Furthermore, the combination of GIS capabilities with MCDM techniques provides the decision maker with support in all stages of decision making, that is, in the intelligence, design and choice phases of the decision-

making process (Laaribi *et al*., 1996; Malczewski, 1999; Thill, 1999 and Chakhar and Martel, 2003; Chakhar and Mousseau, 2008; Malczewski, 1996; Pereira *et al*.,1993; Carver, 1991; Malczewski, 2006).

Land Suitability Analysis (LSA) implies the assignment of values to alternatives that are evaluated along multiple decisions or criteria (Pereira *et al*., 1993). These criteria are detrimental to land suitability analyses for different land use types. Generally, land suitability analysis evaluates many alternative land use types under various criteria from various disciplines. Decisions have to be taken at various levels starting from the selection process, for instance, different livestock species or crop until their allocation to an area that suits best. Analyzing suitability is mainly based on the land qualities such as erosion resistance, water and nutrient availability, rooting condition, drainage and flood hazard.

The value of land quality is the function of the assessment and grouping of land types into orders and classes in the framework of their fitness. Generally, land suitability is categorized as suitable (S) and not suitable (N). Whereas S features lands suitable for use with good benefits, N denotes land qualities which do not allow considered type of use, or are not enough for suitable outcomes (FAO, 1993, 1985). Suitability orders could be further sub-divided. Accordingly, three classes (S1, S2 and S3) are often used to distinguish land that is highly suitable, moderately suitable and marginally suitable for a particular use. Two classes of 'not suitable' can usefully distinguish land that is unsuitable for a particular use at present, but which might be useable in future (N1), from land that offers no prospect of being so used (N2).

The procedure for optimizing land use allocation will depend on whether the land uses are compatible or conflicting Mendoza (1997). When the land uses are compatible technically there is no pressure to allocate the land for alternative land uses. Hence, the allocation is simply based on a descending measure of overall or cumulative suitability for the compatible land uses. However, the optimal land use allocation procedure is a bit more complicated when the objectives are conflicting. In this case, land use allocations are exclusionary; that is, land units can be allocated to only one land use. Mendoza (1997) advised a 'prioritized allocation' to solve the problem. That is, the land uses are compared in terms of priority. Allocation is done first to the land use rated as the highest priority. Then, allocation of remaining land units is done for lower priority land uses.

3. Materials and Methods

The study area is the special zone (SZ) of Oromia Regional State surrounding Addis Ababa city which comprises 499,209ha (Figure 1). The SZ is located between latitude $8^0$34'25" and $9^0$32'41"N and longitude $38^0$25'50" and $39^0$07'53"E. The altitude of this area ranges from 1500 to 3443 meters above sea level (masl). The major towns of this SZ are Burayu, Sabata, Galan, Dukem, Holeta, LagaTafo, Sandafa, Bake, Sululta and Chancho (Figure 1).

The major soil types of the study area are Cambisols, Leptosol, Luvisols, Nithosols and Vertisols. Whereas vertisol covers the highest portion of the SZ (36.06%), combisol amounts to 25.8%. The remaining three major soil types collectively account for 29.03% of the total area of the SZ. Noteworthy is also that about 8.12% of the study area is covered by rock surface.

Agro-climatically, the study area is partitioned into seven parts. These are cool humid, tepid humid, cool moist, tepid moist, warm moist, cool sub-humid and tepid sub-humid. Whereas the largest proportion (37.62%) of the study area is classified as cool sub-humid, tepid and tepid moist accounted for 22.45% and 20.37 %, respectively. The remaining agro-climatic zones collectively fall below 20%.

The study area has mean temperature that ranges from 10°C to 26°C. Moist to humid moisture characterizes areas with the LGP greater than 120 days and annual rainfall of 1043.87mm to 1316.6mm. Eighty percent of the annual precipitation of the study area occurs from June to September, with a peaking from July-August. Furthermore, small and unreliable rainfall occurs in the month of April followed by dry spell in the month of May in some parts of the study area. The area is characterized by low annual rainfall variability (<30%) indicating the stable nature of the rainfall and no risk of drought hazard from low rainfall.

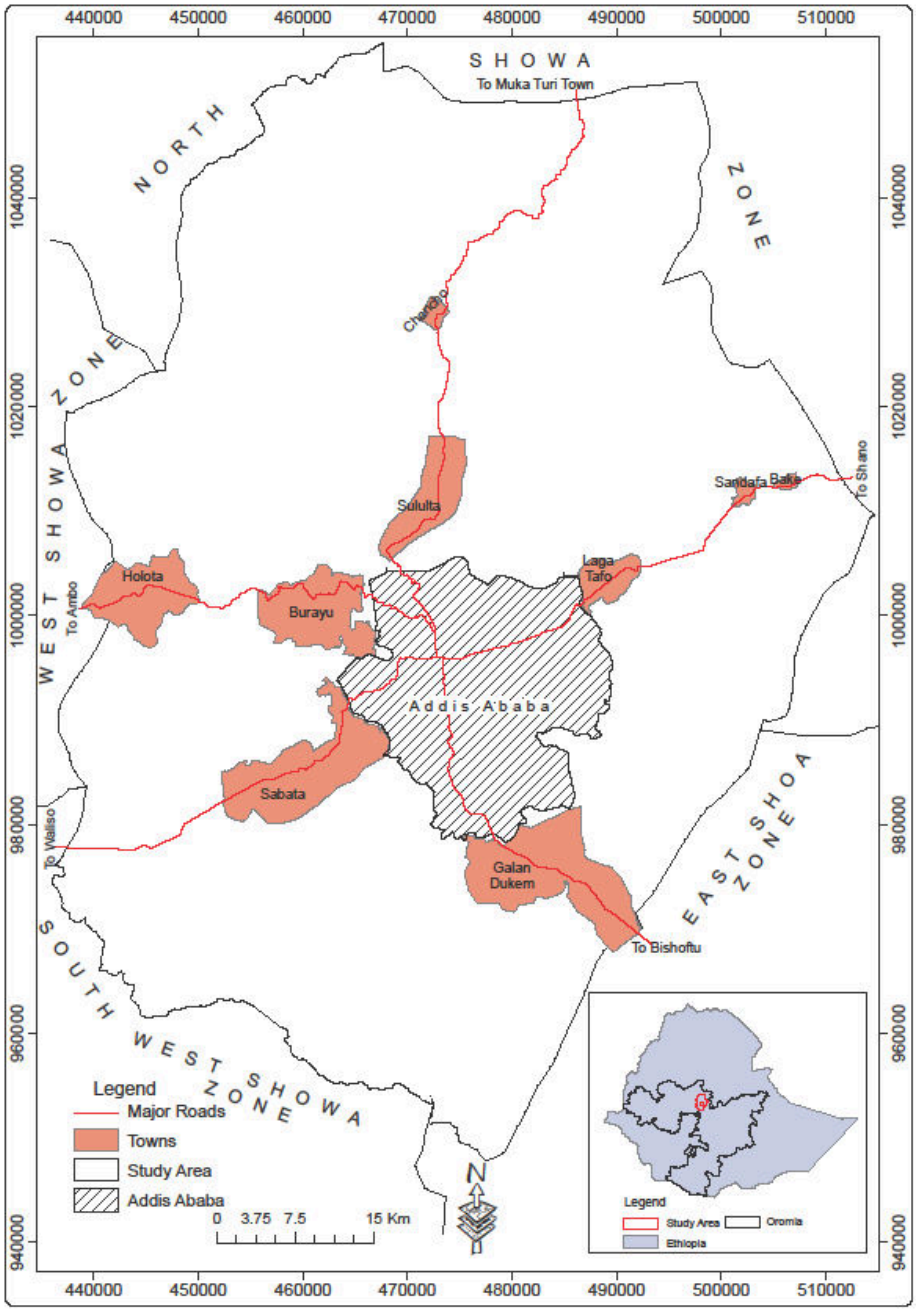

Figure 1. Location map of the study area

3.1 Mapping of Land Types

The robustness of the GIS-based land suitability studies depends on the quality of geographic data (Yeung, 2002). This is also cumbersome and expensive and yet constitutes the critical stage of these types of studies.

In the present study, both primary and secondary data were used. Mosaic of Landsat Enhanced Thematic Mapper plus (ETM+) imagery, path/row 168/53, 168/54 and 169/54, with 30-meter resolution taken in January 2005 by ETM+ sensor assumes a recent ETM+ data in the absence of up-to-date cloud free and useable imagery. The ETM+ image data was acquired from online archive of the USGS GLOVIS. In addition, QuickBird panchromatic image with 61-centimeter resolution acquired in 2007 and SPOT-5 panchromatic image with 5-meter resolution taken in 2006 were used for land use/land cover classification. These two datasets were purchased from Ethiopian Mapping Agency (EMA). Besides, topographic maps of scale 1:50,000 were also used to assist land use types classification of the study area. Land use/land cover types of the study area were analyzed using Erdas Imagine 9.2 software. Ground truthing was conducted to verify the reliability of the preliminary output of GIS and Remote Sensing. Information from local community elders were also gathered to further refine these data.

3.2 Assessment of Variables and Criteria

Topographic maps, aerial photographs, satellite images, digital elevation model, meteorological and soil survey data were used to extract environmental variables pertinent for the current studies. These are among others climate, hydrology, geology, landforms and soils. Climate is important because it affects the growth of vegetation and crop while hydrology determines the total availability of water. Furthermore, whereas terrain is important for maintaining slope stability and governs local scale microclimate, soil determines the type of vegetation. Table 1 shows land qualities and characteristics and environmental suitability rating for Teff.

In establishing the environmental suitability rating table (Table 1) reference were made to: national manuals, guidelines, research station publications, and relevant literature. In addition, community elders and local agriculturalists experience and opinion were considered in arriving at a factor rating of a given land use that can be used in the matching processes.

Above all, the following three parameter maps were used for land suitability analysis and allocation of the study area:

a) Agro-Climate Zone (ACZ) Map

This was produced by superimposing thermal zone and growing period zone maps. Hence, seven agro-climate zones were identified in the study area. These are cool

humid, cool sub-humid, cool moist, tepid humid, tepid sub-humid, tepid moist, and warm moist. The purpose of agro-climate zoning is to provide a condensed inventory of the agricultural potential and constraints as they are determined by the temperature and moisture in a given area.

b) Soil Map

Assessment of soil for land evaluation and crop suitability requires a detail evaluation and characterization. Thus, every activity was carefully conducted to examine soil distribution, types and all related factors. In the study area, there are basically six major types, namely; Vertisols, Luvisols, Fluvisols, Nitisols, Cambisols, and Leptosols. There are 33 soil-mapping units (SMU).

c) Land Use/Land Cover Map

This is useful for resource assessment, land use planning, land evaluation, and land use/land cover change detection. For the present study land use/land cover map was generated from satellite imagery (cf. section 3.1). It was created in order to inform development-planning process.

No	Land Quality	Land Characteristics	Unit	Suitability Ranges				
				S1	S2	S3	N1	N2
1	Temp. Regime	Alt.	m	1800-2600	2600-2800	2900-2800	2900-3200	>3200
					1400-1800	1400-1200	1200-900	<900
		Temp.	^{0}C	15-21	14-15	12-14	11-12	<11
					21-22	22-23	23-25	>25
2	GP	LGP	days	150-180	120-150	90-120	45-90	<45
3	Drainage	Soil drainage	class	I,MW, W,SE,E	P	NE	NE	VP
4	Nutrient availability	Total Nitrogen	%	≥0.25	0.15-0.25	0.1-0.15	<0.1	-
		Available phosphorus	ppm	>8	4-8	<4	-	-
		Soil reaction	P^H	5.5-7.5	5.2-5.5	5-5.2	4.5-5.0	<4.5
					7.5-7.8	7.8-8	8.0-8.5	>8.5
		OM	%	>3	2.5-3	2-2.5	1-2	<1
5	Rooting condition	Soil depth	cm	>50	30-50	20-30	10-20	<10
		soil texture	class	SiCL-C	L-SI	SL	LS	LS-S
		Stone & rock out crops	%	<0.1	0.1-15	15-40	40-50	>50
6	Erosion hazard	Slope	%	0-13	13-25	25-40	40-55	>55

Table 1. Land qualities/characteristics and environmental requirement suitability rating for Teff

3.3 Delineation of Land Unit

Land units are areas that are relatively homogeneous with respect to climate, landforms, soils and vegetation (FAO, 1993). Each land unit presents similar problems and opportunities and will respond in similar ways to management.

Land unit delineation involves the representation of land in layers of spatial information and combination of layers of spatial information using Geographic Information System (GIS). Therefore, to analyze the present situation in the study area it will be necessary to break the area down into land units. Accordingly, in this study, land units were delineated by overlaying Soil Mapping units, Lengths of Growing Periods (LGP) and thermal zones maps of the study area. As a result, 382 land units (LU) were identified in the study area.

3.4 Matching Land Use with Land Quality

The land qualities (i.e. temperature, rainfall, slope, altitude, soil and etc.) of each land unit were matched with the corresponding land use types. Furthermore, tables of specifications relating measurable land characteristics to the requirements of the land utilization types were formulated. We assigned each land unit to its land suitability class according to the most severe limitation.

3.5 Land Suitability Classification

Following the comparison of the requirements of land use types with properties of land units, a provisional land suitability classification was performed. Suitability was assessed separately for each land use type, i.e. whether the land is suitable or not suitable. This could be accomplished in a GIS environment utilizing a generic model of land suitability assessment $S = f(X1, X2\ldots Xn)$ Where S = Suitability measure; X1, X2… Xn = the factors affecting the suitability of the land.

The above model could be implemented to generate a suitability map for a particular land use. Hence, suitability maps reflecting the suitability values of each land unit relative to a particular land use could be generated. The following classes of suitability were used: highly suitable (S1), moderately suitable (S2), marginally suitable (S3), currently not suitable (N1) and permanently not suitable (N2).

3.6 Land Use Allocation

Land allocation involves the process of designing an optimal mix of land uses based on their estimated suitability and perceived management objective (Mendoza, 1997). Accordingly, the present study used different measures of land use suitability as guides to optimally allocate lands to their most suitable uses. We used the multi-criteria

decision-making (MCDM) framework to allocate land units by maximizing the overall suitability of a land area.

Following the creation of suitability maps for each land use, allocation of land to alternative uses was addressed using a general optimization model F(X) = max (S1, S2… Sn); where F (x) = overall cumulative suitability; S1, S2… Sn = measures of suitability for each land use.

Finally, land units were allocated to specific uses. A series of options for the allocation of land use types to land units were set out. These options are different types of crop, livestock species and forestry. Figure 2 depicts the flowchart of the methodological approaches followed for the current land suitability studies and allocation.

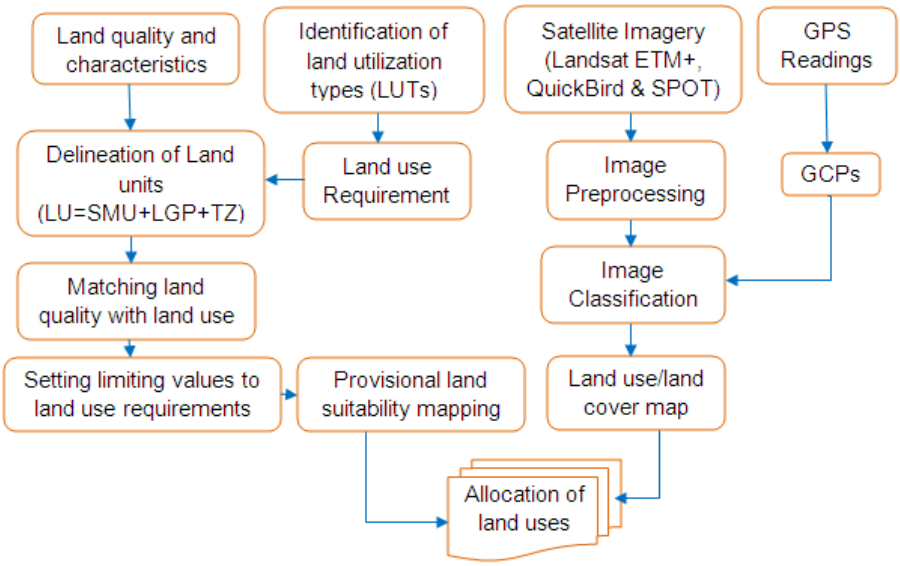

Figure 2. Methodology flowchart for suitability based land use allocation

4. Results and Discussion

4.1 Suitability Analysis Results

The overall suitability analysis result for crop and livestock showed five classes of land. These are: highly suitable (S1), moderately suitable (S2), marginally suitable (S3), currently not suitable (N1) and permanently not suitable (N2) (Table 2).

The area coverage of each suitability class for both crop and livestock under the present study was calculated in a GIS platform. Table 2 shows these findings. Table 2 also reveals that 9.7% and 2.6% of the study area are classified as moderately suitable for crop and livestock production, respectively. In addition, 11.07% and 21.36% were

found to be currently not suitable while 13.40% and 10.99% is permanently not suitable land for crop and livestock production, respectively. Noteworthy is that nearly 65% of the study area is least suitable for the production of both crop and livestock. The result of the study also showed that no part of the study area was recorded to be most suitable for livestock production and only 0.6% of the study area is highly suitable for crop production (Table 2).

Suitability Classes	Crop		Livestock	
	Area		Area	
	ha	%	ha	%
Highly Suitable (S1)	3192.8	0.64		
Moderately Suitable (S2)	48597.9	9.73	13283.8	2.66
Marginally Suitable (S3)	325270.6	65.16	324431.4	64.99
Currently Not Suitable (N1)	55264.9	11.07	106653.9	21.36
Permanently Not Suitable (N2)	66882.8	13.40	54840.2	10.99
Total	499209	100.00	499209	100.00

Table 2. Spatial coverage of suitability classes for both crop and livestock production

On the contrary, Satellite image classification of the study area has shown that 58.4% (291,756 ha) and 21.7% (108,263.3 ha) of the total area is currently used for crop cultivation and livestock grazing, respectively (Table 3). The remaining 19.9% is collectively covered by natural and plantation forest, flower farm, water body, woodlot, inundated land and settlement area. This discrepancy in the figures for current (existing) use of land and suitability analysis result is attributable to the major limitations of the land resources in the study area for crop and animal production which are high slope gradient, heavy soil texture, poor drainage conditions, acidic and alkaline soil reaction, low available phosphorous and total nitrogen.

Use type	Area	
	ha	%
Cultivated land	291756.0	58.4
open grassland	36327.7	7.3
Bushed shrub land	32305.0	6.5
Bare land	39630.7	7.9
Woodlot	1952.5	0.4
Inundated land	4538.6	0.9
Plantation forest	21697.0	4.3
Natural forest	3648.6	0.7
Flower farm	312.5	0.1
Water Body (Dam)	937.1	0.2
Settlement (Urban and Rural)	66103.4	13.2
Total	499209.0	100

Table 3. Present land use/land cover types of study area as extracted from satellite imagery

For instance, if we consider slope gradient of the study area independently, approximately 10% of the total study area is classed as high slope gradient (> 30%). This attribute of land alone made the area permanently unsuitable for crop production at intermediate level of management. Refer Table 4 for details of slope classes of the study area.

Slope Class	Area	
	ha	%
0-2	92473.9	18.52
16-30	72827.8	14.59
2-5	130379.9	26.12
5-8	53567.8	10.73
8-16	100560.8	20.14
>30	49398.7	9.90
Total	499209	100.00

Table 4. Spatial coverage of slope classes

4.2 Proposed Land Use Allocation

The present study proposed multi-objective land use allocation. Accordingly, based on their estimated land suitability, both land utilization types (livestock and crop) were allocated in combination instead of exclusive land use allocation. This type of land use allocation is purposeful especially for sustainability of land productivity. To this end, the present study followed an optimal mix of land uses based on their estimated suitability to allocate land uses. (cf. section 3.6 – land use allocation – for the explanation of how the allocation was accomplished). Hence, the study area was allocated for the following uses: Livestock-Crop Specialization, Crop-Livestock Specialization, Plantation Forest, Protective Forestry, Productive Forestry, Natural Forest, Flower Farm, Water Body, and Built up areas (towns and rural settlements) (Table 5 and Figure 3).

a) Crop-Livestock Specialization

The study has shown that 8.57% (42,788.5 ha) of the total study area is a combination of suitability classes one and two (S1 and S2) and have relatively better soil fertility status. Therefore, crop were allocated as main production and livestock supplementary production as the name crop-livestock specialization implies. However, currently 58.4% (29,1756 ha) of the total area is being used for crop cultivation (Table 3). This is an indication for lack of proper land use planning in the study area. Therefore, to get the optimum benefit out of the land, proper use of it for specific purposes is inevitable.

b) Livestock-Crop Specialization

Table 5 elucidates that 59.24% of the study area is allocated for livestock as main production and crop as supplementary production. This area is marginally suitable (S3) for crop production as estimated suitability analysis result showed (Table 2). It is because of this fact that crop were allocated as supplementary production under this spatial extent. To get better production while improving the soil productivity at the same time, the present study has allocated both livestock and crop for aforementioned spatial extent.

c) Productive Forestry

In the present study, the area which is very marginal for crop and livestock production (suitability classes N1) is allocated for Productive Forestry**.** Table 5 shows that 6.23% of the total study area is allocated for this particular land utilization type.

Land use Allocation	Area	
	Area	%
Crop-Livestock Specialization	42,788.5	8.57
Livestock-crop Specialization	295,752.5	59.24
Plantation Forest	21,697.0	4.35
Major Towns	40,588.3	8.13
Protective Forestry	36,846.2	7.38
Productive Forestry	31,123.1	6.23
Natural forest	3,648.6	0.73
Flower farm	312.5	0.06
Water body (Dam)	937.1	0.19
Rural Settlement	25,515.1	5.11
	499,209.0	100.00

Table 5. Spatial extent of suitability based land use allocation

d) Protective Forestry

The area which does not fit for productive forestry and more fragile environment and which is permanently not suitable for crop and livestock production was allocated for Protective forestry. This portion of land utilization type accounts for 7.38% (36,846.2 ha) of the total study area (Table 5).

e) Others

Noteworthy is that the present study made use of the land use/land cover classification output of the study area. As a result, rural and urban setups, plantation and natural forest, flower farm and water body were carefully considered and retained their present spatial location during the allocation process. As Table 5 elucidates major towns and rural settlement together accounts for 13.24% of the study area. Whereas natural and plantation forest, flower farm and water body collectively covers 5.33% of the special zone.

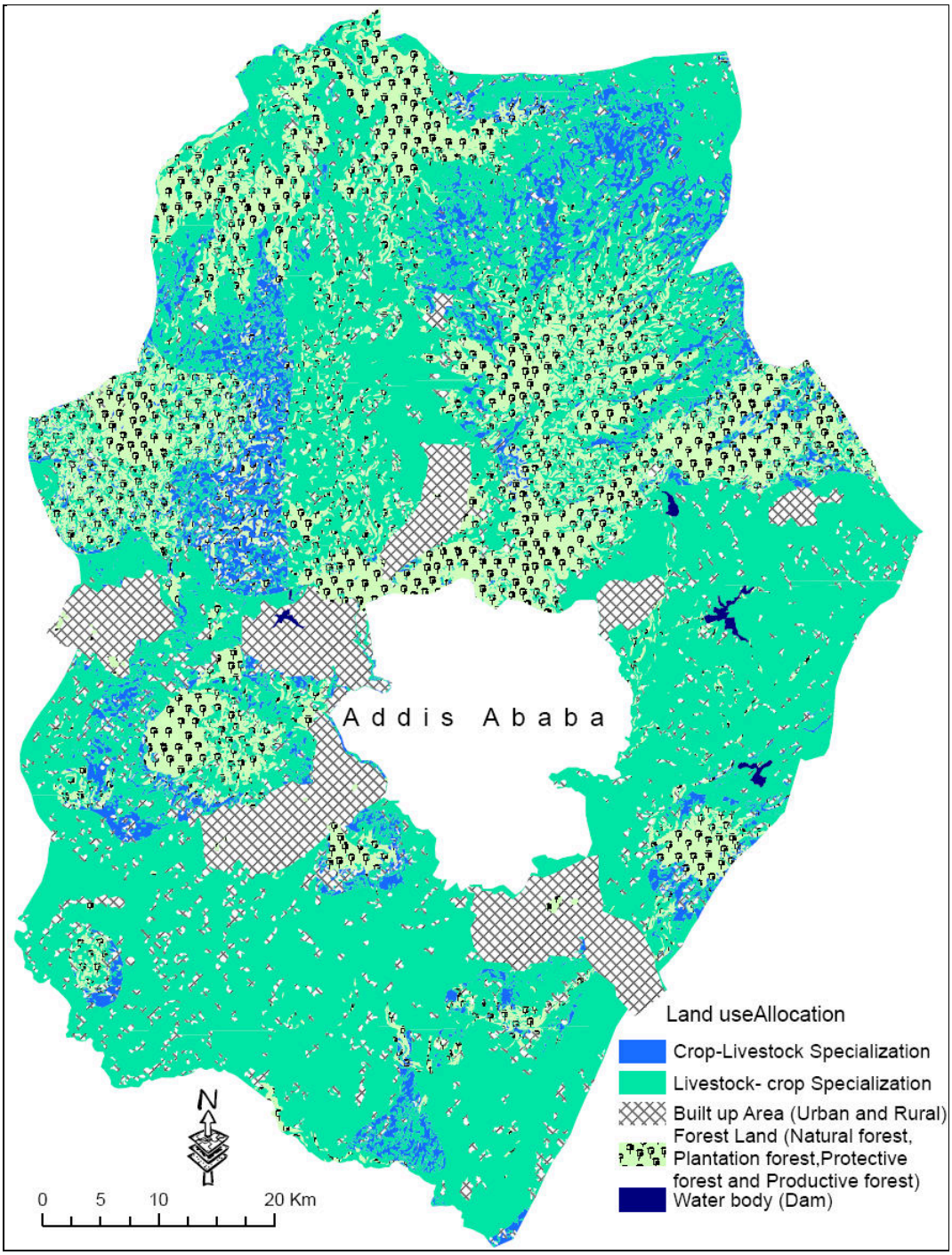

Figure 3. Land use allocation map of the study area

5. Conclusion and Recommendation

The main objective of this study was to contribute the efficient and effective utilization of land resources to promote sustainability of natural resources. In order to substantiate the suitability-based land use allocation in the special zone, necessary and reliable data were collected, processed, analyzed and interpreted to come up with a sound conclusion, feasible recommendations and practical land use allocation.

The study used multi-objective land use allocation techniques in a GIS platform to arrive at the final land use allocation for different purposes. As per the suitability analysis result, nearly 65% of the total study area is marginally suitable for both crop and livestock production. In addition, 9.7% and 2.6% of the study area are classified as moderately suitable for crop and livestock production, respectively. However, satellite image classification of the study area has shown that 58.4 % (291,756 ha) and 21.7% (108,263.3 ha) of the total area is currently used for crop cultivation and livestock grazing, respectively (Table 3). This discrepancy in the figures for current (existing) use of land and suitability analysis results is attributable to high slope gradient, heavy soil texture, poor drainage conditions, acidic and alkaline soil reaction, low available phosphorous and total nitrogen. Furthermore, with regard to the output of the suitability analysis, it should be noted that 24.47% and 32.35% of the total study area are not suitable for crop and livestock production respectively.

The present study followed an optimal mix of land uses based on their estimated suitability to allocate them. Accordingly, the study area was allocated for the following uses: Livestock-Crop Specialization, Crop-Livestock Specialization, Plantation Forest, Protective Forestry, Productive Forestry, Natural Forest, Flower Farm, Water Body, and Built up areas (towns and rural settlements). The study has shown that the larger proportion of the special zone (67.81%) was allocated for crop and livestock production collectively.

This study may not provide the ultimate explanation for all problems related to land and cannot be an end in itself. Nevertheless, it serves as a stepping stone for understanding the potential and limitations of the land in the study area. Hence, the designed activities have to be worked out in detail scale whenever needed for easy implementation.

Above all, major areas of land use conflicts between the study area and Addis Ababa city should be resolved between the Regional National State of Oromiya and Addis Ababa city Council.

References

Bizuwerk, A., Peden, D., Taddese, G. and Getahun, Y., (2005), GIS Application for analysis of Land Suitability and Determination of Grazing Pressure in Upland of the Awash River Basin, Ethiopia. International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia.

Carver, S.J., (1991), Integrating Multi-criteria Evaluation with Geographical Information Systems. *International Journal of Geographical Information Science*, 5: 321-339.

Chakhar, S. and Martel, J-M., (2003), Enhancing Geographical Information Systems Capabilities with Multi-criteria Evaluation Functions. *Journal of Geographic Information and Decision Analysis*, 7: 47-71.

Chakhar, S. and Mousseau, V., (2008), GIS-based Multi-criteria Spatial Modeling Generic Framework. *International Journal of Geographical Information Science,* 22: 1159-1196.

Cowen, D., (1988), GIS versus CAD versus DBMS: What are the Differences. *Photogrammetric Engineering and Remote Sensing*, 54(2): 1551-1555.

Food and Agricultural Organization (FAO) (1976), A Framework for Land Evaluation. FAO Soils Bulletin No. 32, International Institute for Land Reclamation and Improvement, Wageningen, The Netherlands.

Food and Agricultural Organization (FAO) (1985), Guidelines: Land Evaluation for Irrigated agriculture. FAO Soils Bulletin 55, Rome, Italy.

Food and Agricultural Organization (FAO) (1993), Guidelines for Land-use Planning. Rome, Italy.

Grimshaw, D.J., (1994), Bringing Geographical Information Systems in to Business. Harlow, Essex, England: Longman Scientific and Technical.

Joerin, F., Theriault, M. and Musy, A., (2001), Using GIS and Outranking Multi-criteria Analysis for Land-use Suitability Assessment. *International Journal of Geographical Information Science*, 15: 153-174.

Laaribi, A., Chevallier, J.J. and Martel, J.M., (1996), A Spatial Decision Aid: A Multi-criterion Evaluation Approach. *Computers, Environment and Urban Systems*, 20: 351-366.

Lo, C.P. and Yeung, A.K.W., (2002), Concepts and Techniques of Geographic Information Systems. New Delhi, India.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., (2005), Geographic Information Systems and Science. Second edition, John Wiley & Sons, Ltd, England.

Malczewski, J., (1996), A GIS-based Approach to Multiple Criteria Group Decision-making. *International Journal of Geographical Information Science,* 10(8): 955-971.

Malczewski, J., (1999), GIS and Multicriteria Decision Analysis. Wiley, New York.

Malczewski, J., (2000), Review Article on the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. *Transactions in GIS,* 4: 5-22.

Malczewski, J., (2003), GIS-based Land-use Suitability Analysis: A Critical Overview. *Progress in Planning*, 62: 3-65.

Malczewski, J., (2006), GIS-based Multi-criteria Decision Analysis: A Survey of the Literature. *International Journal of Geographical Information Science,* 20: 703-726.

Malczewski, J., (2006), Integrating Multi-criteria Analysis and Geographic Information Systems: The Ordered Weighted Averaging (OWA) Approach. *Int. J. Environmental Technology and Management*, 6(2): 7-19.

Mendoza, G.A., (1997), Introduction to Analytic Hierarchy Process: Theory and Applications to Natural Resources Management. In Proceedings of 1997 ACSM/ASPRS Annual Convention. Vol 4. Resource Technology. April 7–10. Seattle, WA. pp.130-39.

Pereira, J.M.C. and Duckstein, L., (1993), A Multiple Criteria Decision-making Approach to GIS-based Land Suitability Evaluation. *International Journal of Geographical Information Science,* 7: 407-424.

Simon, H.A. 1960. The New Science of Management Decision. New York: Harper & Row.

Steiner, F., McSherry, L. and Cohen, J., (2000), Land Suitability Analysis for the Upper Gila River Watershed. *Landscape and Urban Planning*, 50: 199-214.

Thill, J.C. (ed.) (1999), Multi-criteria Decision Making and Analysis: A Geographic Information Sciences Approach. Brookfield, VT: Ashgate publishing.

Zhang, Y.J., Li, A.J. and Fung, T., (2011), Using GIS and Multi-criteria Decision Analysis for Conflict Resolution in Land Use Planning. *Procedia Environmental Sciences*, 13: 2264-2273.

CHAPTER 12

Context-Aware Recommender System Based on Ontologies

Guillermo González Suárez[1], *Tatiana Delgado Fernández*[2], *José Luis Capote Fernández*[1] *and Rafael Cruz Iglesias*[1]

[1] GEOCUBA Enterprise Group, Cuba
[2] High Polytechnic Institute Jose Antonio Echeverria, ISPJAE, Cuba

guille@geomix.geocuba.cu; tatiana.tsp@gmail.com; capote@geomix.geocuba.cu; rcruz@geomix.geocuba.cu

Abstract

This chapter focuses on developing a method of hybrid context-aware recommender systems, based on ontologies in the environment of a Spatial Data Infrastructure, which will positively impact the effectiveness of decision-making and raise the capacity for analysis in applications for mobile users. The method developed has three components: data and semantic framework, others that implement the recommender system, and another which is responsible for displaying the recommendations. The recommender-system component implements spatial, semantic and collaborative filters. The recommendations generated show the elements of potential interest. These can be displayed in tabular form or combined with spatial information. It was through experiments and differently applied metrics, that the increase in its effectiveness for mobile users according to their preferences as to the recommendations given in a spatial environment, was discovered.

KEYWORDS: Recommender System, Context, Ontology, Spatial Data Infrastructure

1. Introduction

The domain of geographic information has grown rapidly due to computational development, resulting in a large amount of spatial data being available on the Web. However, there is an increasing need to share this information among different stakeholders and information systems to facilitate consistent and contextual use of it.

This need forms the basis for the emergence of a number of infrastructure at different levels, international, national and regional, for the collection and dissemination of geographic data such as Spatial Data Infrastructures (Vaccari *et al.*, 2009).

A Spatial Data Infrastructure (SDI) encompasses the policies, technologies, standards and human resources necessary for the effective collection, management, access, delivery, and use of spatial data at different levels, depending on economic decision making, political, social and sustainable development (Delgado Fernández and Capote Fernández, 2009).

In modern society, there is a growing amount of information that is considered critical to daily decision-making. As more information becomes available online, including geographic context, the ability to discover and access geographic data resources for visualization, planning and decision making is becoming an indispensable requirement to support society (Nebert, 2004).

Nowadays, much electronic content is created and delivered to users such that they are overloaded with information. Tools addressing information overload have thus become necessary. Such tools provide recommendation; help users to understand better the information needed so as to make more effective use of it (Belkin, 2000).

Given the large amount of information available on the Web, industries such as marketing and sales have in recent years developed and implemented various tools to provide users quick access to the appropriate information needed (Espinilla *et al*., 2009). Recommender Systems have emerged strongly in this area. Recommender Systems aims to personalize the information that users receive according to their needs, preferences and/or tastes. Because of its success, recommender systems can be applied in a wide range of uses (Schafer *et al*., 2001), especially for e-commerce and entertainment.

Recommender systems take into account user preferences to suggest information of potential interest to them, streamlining searches by providing a number of elements that is relevant. These provide information alerts and support navigation when you have large volumes of information (Espinilla *et al*., 2009).

In literature and in the market, there are different types of recommender systems that differ in the method or process of obtaining the recommendations and/or information sources used, especially: collaborative recommender systems (Adomavicius and Tuzhilin, 2005), content-based systems (Martínez *et al*., 2007), knowledge (Burke, 2000), utility or hybridizing some of these techniques (Burke, 2002).

The proliferation of mobile phones in society has begun to change the dynamics of communication, described by many as the way information is accessed and used in a world that is 'mobile centered' (Sacher and Loudon, 2002). In this new mobile paradigm, the focus shifts from the stationary user to the mobile user, one who has

different information needs. This information is closely linked to its geographical location (Bernardos, 2008).

Many of these devices are Java-enabled and support Wireless Application Protocols (WAP) via Global System for Mobile Communications (GSM) or General Packet Radio Services (GPRS). Integrations with Personal Digital Assistants (PDAs) and have incorporated technologies like Global Positioning System (GPS) and Bluetooth forming a device capable of storing and organizing their tasks and communicating with other users devices (Koh and Kim, 2000).

Users of these mobile systems are increasingly demanding more relevant and pertinent information to support their work, and without the use of the various device applications would mean additional effort and inconvenience. These users expect these technologies to be incorporated easily into their lives (Weiser, 1991).

Mobile devices are presented as autonomous objects that require minimal user intervention. They are distributed and interconnected physically and develop collaborative behavior, are able to form coalitions to offer more functionality to the user (Urbieta and Barrutieta, 2007), and are sensitive to changes in environment information, such as the location and status of users and devices (Saha and Mukherjee, 2003).

Reviewing the work in this area outlines several initiatives in the definition and design of context models. Some of these papers collect interesting theoretical proposals, although not always has its implementation been carried out. In this context, (Henricksen and Indulska, 2006) are prominent modeling trends, whereas (Lee and Meier, 2007), (Park *et al*., 2007) have proposed various hybrid models of context. However, in these proposed model approaches are limited context models that combine an ontology-based approach with the spatial (McGuinness and van Harmelen, 2004).

The aim of this chapter is to describe a method of context-aware recommender systems based on ontologies in the SDI environment. For this, the design of the method and its validation during the experimental stage are key elements.

2. Materials and Methods

The information that feeds the system is represented by ontology of destinations and a database that stores user profile and preferences.

We present the result of the analysis and the design of the proposed method, starting from the general to the specific, first showing a global scheme as part of the theoretical description of the method, and later detailing the components involved.

2.1 Technologies Used

For the development of the proposed system, PostgreSQL 9.2 was chosen for its features as an open-source database manager.

The implementation of the functionalities of the application has used a number of technologies, always prioritizing the free software listed below. The selected programming language was Java, using NetBeans Development Interface.

Google Web Toolkit (GWT) was used as a development framework, created by Google, which allows hiding the complexity of various aspects of AJAX technology.

Visual environments were achieved using Ext, a class library OpenSource JavaScript language that can be included in Web applications which comes with a well-designed visual interface and many components. This library of components is used in several application modules and its inclusion in the panels that display maps facilitates easy integration with other elements.

For visualization of geographic data, OpenLayers is used – a completely free open-source library developed in JavaScript and licensed from Berkeley Software Distribution – its function is to facilitate the location of dynamic maps in a webpage.

Points of Interest was designed using Protegé, and integrated with Jena – an open-source semantic Web framework for Java. It provides an API which extracts data from, and writes to RDF graphs. The graphs represent an abstract 'model'. A model can be sourced with data from files, databases, URLs or a combination of these. A model can also be queried through SPARQL.

The quality of the recommendations of the proposed method depends on the component that performs and logically influences the efficiency of implementation. Within existing engines, Mahout was chosen, an open-source library of machine learning of Apache project.

2.2 Description of the Solution

The proposed method is divided into three components, semantic, data framework and context-aware recommender system based on ontologies and viewing of the recommendations. Information flows through it in that order. Figure 1 shows the diagram that serves the proposed method description:

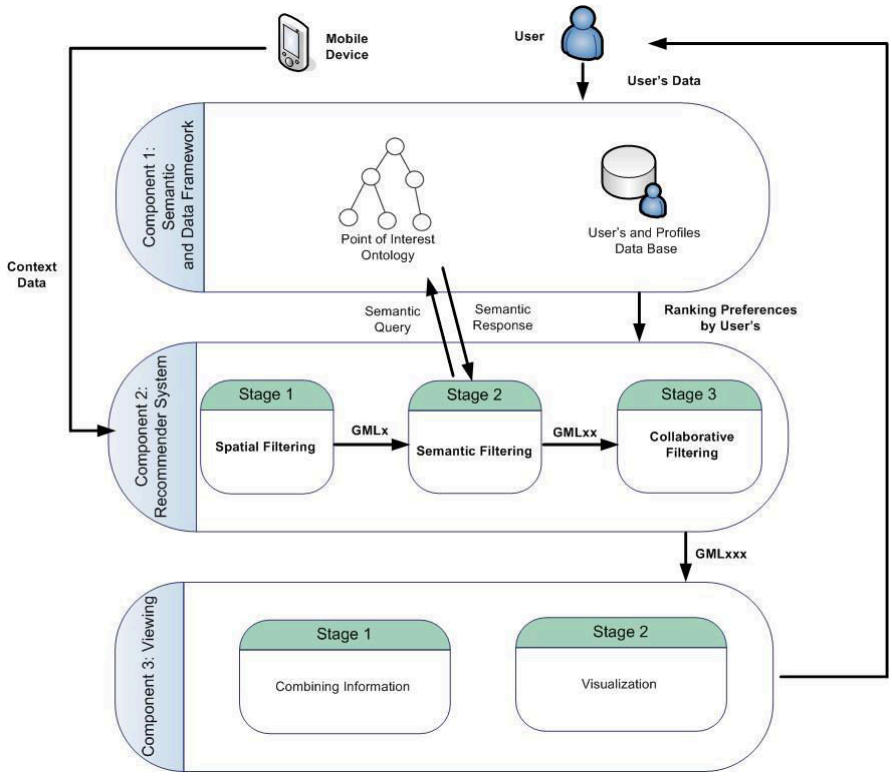

Figure 1. Diagram of a context-aware recommender system based on ontology method

Its performance begins when a user logs into the system, then from access data manages their preferences in the database of user profiles. These are sorted together with context data. The geographical position data and time forms the basic information of the recommender system component, a key component within the method, responsible for filtering data of the point of interest for the user using three criteria, Spatial, Semantics and Collaborative Filtering. Finally, it executes the display component of the method, which has the function of combining the data obtained in the previous component with the mapping services and obtains a ready representation for display. Geographic information is handled within the method in GML format (Cox *et al*., 2003). The following, describes in detail each of the component with its most important elements.

2.2.1 Component 1: Semantic and Data Framework

This component begins when the user logs into the system or when there is a change of context, such as a change of location. At the end of its execution, preferences obtained for the user are ordered by levels of preference queries from the database: building information for this component of the method recommender system.

In addition to the information in the proposed method, is an ontology of Points of Interest and a database of user profiles. This combination forms the semantic and data framework of the method.

The framework that supports the method consists of an ontology for Points of Interest – a resource of specific locations that have attractive user profiles. This has been designed using Protegé tool (Mayrhofer *et al*., 2003), an open-source environment for the development of knowledge-based systems, one of the editors of OWL (Nivala and Sarjakoski, 2003) ontologies software.

The purpose of the ontology for points of interest is providing information on potential places to visit by users for later reference semantics. For the design, we started from a super class called Place, which includes Name, Address, Phone, Latitude and Longitude and others. This than branches out to places you want to store, such as restaurants, hotels, cafes, etc. A fragment is shown in Figure 2.

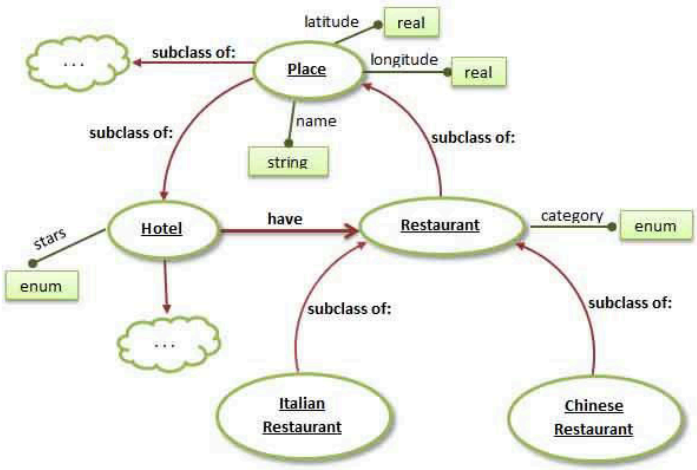

Figure 2. Fragment of ontology for points of interest

Recommender systems work based on the information known about the user, usually stored in their profile (Hong, 2008). Part of the method includes an initial process where you ask the user a series of data (name, age, sex, general tastes, etc.). This assists in initializing their profiles. Besides this, other data are recorded, including preferences regarding visited places. This information can be supplied by the customer or is recorded in the system which is used. The user profile is stored in a database, which completes the initial component of the method.

2.2.2 Component 2: Context-aware Recommender System

Component 2 is based on the information provided by the previous component. It is essential that the user is registered and that their preferences are recorded and analyzed. Another important requirement is that the mobile device provides location information.

During its execution, which is divided into three stages, it refines the set of possible destinations to be visited by the user based on his preferences.

In stage one, the spatial filter lists these destinations based on the criteria of spatial distance. In the second stage, the filter runs a semantic search and considers user preferences, while in the third and final stage, a recommendation process sorts destinations based on the preferences and tastes of the user. At all stages, data is in Geography Markup Language (GML).

For the implementation of the spatial filtering stage, it is essential to have the position data of the user which can be provided by the mobile device. This step reduces the search universe of the method, since it eliminates those destinations that are not within the buffer. The returned results are stored in a file in GML format, which serves as a source for the next stage. For the selection of the destinations included in the catchment area are made GeoSPARQL semantic query format, as shown below:

PREFIX geo:http://www.opengis.net/ont/OGC-GeoSPARQL/1
PREFIX geof:http//www.opengis.net/def/queryLanguage/OGC-GeoSPARQL
SELECT ?place WHERE {?place geo:hasGeometry ?pgeo.
FILTER (geof:distance (?pgeo, "PONT((-80.089005 23.913574))"^^geo-sf:WKTLiteral), units:m) < 2000)}

The above query is made to apply the places belonging to the ontology of destinations that are within 2000 meters (units: m <2000) from the site where the user corresponding to the coordinates: (-80.089005 23.913574).

In order to select an item p, within the extracted points of the previous stage, in the stage of filtering using the Semantic Filter deployed in the IDERC (Capote, 2011), which allows a selection of points of interest to represent the classification of all activities related to the context 'P' for data users 'U'.

This request is translated and converted to semantic filter format, the component related to the requested criteria administrator user context and as a result of running, you get a list of sites that match the context. Below is an example of this query:

PREFIX geo: http://www.opengis.net/ont/OGC-GeoSPARQL/1
PREFIX geof:http//www.opengis.net/def/queryLanguage/OGC-GeoSPARQL/
SELECT ?place **WHERE** {?place a place:Restaurant}

Selecting from the ontology of Points of Interest, places that match the description 'Restaurant' are expressed in a simple and abbreviated form for the reader.

As the semantics of the Points of Interest is described by ontology, the semantic search engine is aware of the class hierarchy of each Point of Interest. This means that the engine can select the appropriate forecasting strategy for each target class.

In the last stage of this component, called Collaborative Filtering, the recommendation engine uses multiple strategies to predict how each destination can respond to user preferences. A strategy selects and/or combines multiple prediction techniques to decide which is most appropriate to provide a recommendation based on the latest information provided by the user.

The recommendation process typically starts with an initial set of preferences, or provided by users 'U' explicitly and/or implicitly inferred by the system for a Point of Interest 'P'. Once specified, the recommendation system estimates the valuation function 'R' for a new pair (u, p); R: U x P → R.

As a result of the previous stage is a list of possible targets arranged in the form (U x P x R). The recommendation engine uses a database of items and users to generate predictions. Figure 3 shows the diagram of the recommendation engine components.

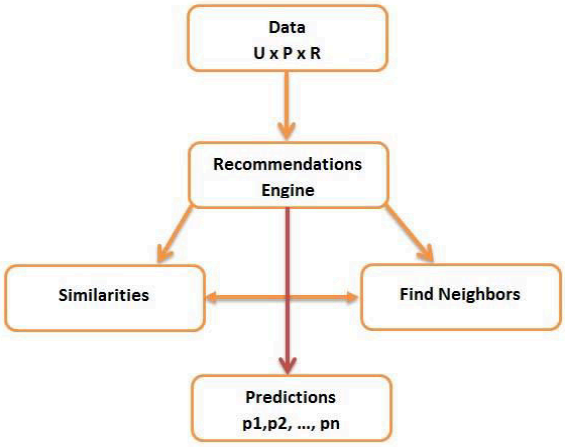

Figure 3. Diagram of the recommendation engine components

The recommendation engine uses a database of items and users to generate predictions. Firstly, statistical techniques are used to find neighbors, i.e. users with a history of reviews on items similar to current user. Once you have built a list of neighbors, combine their preferences to generate a list of N elements recommended for the current user. This recommendation technique is called 'nearest neighbors'.

First, it is necessary to compare the similarities among all users with the current user, for this will execute functions that allow you to calculate the degree of similarity within the 'neighborhood'. Several methods can be used to calculate the similarities:

One is the Pearson correlation function defined as follows:

$$w_{a,u} = \frac{\sum_{i=1}^{m}[(r_{a,i} - \overline{r_a})(r_{u,i} - \overline{r_u})]}{\sqrt{\sum_{i=1}^{m}(r_{a,i} - \overline{r_a})^2 \sum_{i=1}^{m}(r_{u,i} - \overline{r_u})^2}}$$

where:

$w_{a,u}$ is the value of the similarity between the active user 'a' and its neighbor 'u'.
m is the number of elements.
$r_{u,i}$ is the preference value assigned by the user or the element 'i'.
r_u is the average of all the values assigned by the user 'u'.

After compiling all the similarities among users, the 'neighborhood' is made up and the engine is able to make recommendations. For this, the method uses the following prediction function:

$$P_{a,i} = \overline{r_a} + \frac{\sum_{u=1}^{n}[(r_{u,i} - \overline{r_u}) * w_{a,u}]}{\sum_{u=1}^{n} w_{a,u}}$$

where:

$w_{a,u}$ is the value of the similarity between the active user 'a' and its neighbor 'u'.
$P_{a,i}$ represents the prediction for the active user 'a' for item 'i'.
n is the number of neighbors.
$r_{u,i}$ is the preference value assigned by the user or the element 'i'.
r_u is the average of all the values assigned by the user 'u'.

Once the recommendation engine terminates it execution and achieves an ordered list of Points of Interest where the first few hits are of most interest to the user, these points are stored in a file in GML format. They are then ready to send to the third and final component.

2.2.3 Component 3: Viewing Recommendations

This component is executed when it has made the selection of possible destinations to be recommended to the user. As a starting point we have a GML file obtained from the Collaborative Filtering stage of the previous component. This third component has two stages.

The information resulting from the component Recommender System is coded in GML format file that contains a list of Points of Interest relevantly recommended to the user. In the stage of combining the information, mix data with geographic information

from WMS or WFS, which are standardized specifications from OGC consortium and obtain the representation of these points on a reference map.

Through these services, one can request specific data from geographic information services located anywhere in the world. At this stage of the method, it is possible to combine obtained information with services like Google Earth™, Yahoo Maps and/or services published in a Spatial Data Infrastructures, as IDERC (www.iderc.co.cu).

At this stage, we propose results displayed in two forms. The first is to display the data in a thematic window, or to display in a form where you can see several properties.

The other is to transfer the file with the information to a spatial visualization component (OpenLayers) so that it can be combined with other information infrastructure or satellite images to form a basis for data visualization.

3. Results

The described method defines a workflow for the implementation of a context-aware recommender system based on ontologies for mobile device users SDI environments.

This method is like an orchestra performing different operations such as running spatial filters and collaborative semantics in a logical manner. It includes a contextual pre-filter linking the user environment and the use of any collaborative filtering implementations in 2D.

The data obtained as a result of Component Context-aware Recommender System, mixed easily with the IDERC map service, ensuring its standardization. Other scenarios may be used to link other map services which are OGC standards-compliant.

To ensure that the method complies with the purposes for which it was designed, the described component was implemented and a client application was developed as a case study which validated its effectiveness.

The validation of the proposed method focused on component context-aware recommender system based on ontology (Component 2). An essential element in the validation was to have the necessary data perform.

The main problem encountered during testing was that the system had no real user profiles (personal data, general preferences, etc.). To obtain user profiles, we simulated the system by conducting real user surveys.

Specifically, surveys were conducted on 60 different users, 36 men and 24 women. Respondents are divided into several age ranges and educational level, as summarized in Table 1, volunteers were selected so that there were no significant differences in demographics and were representative of as many individuals as possible.

Demographic Dimension	Quantity	%
Gender		
Female	24	40
Male	36	60
Age Range		
Between 13 & 18 years old	7	11.66
Between 19 & 24 years old	21	35
Between 25 & 50 years old	24	40
More than 50 years old	8	13.33
Educational Level		
Medium	2	3.33
High Medium	23	38.33
High	20	33.33
Other	15	25
Occupations		
Student	15	25
Worker	13	21.66
Professional	20	33.33
Housewife	7	11.66
Pensioner	5	8.33

Table 1. Information gleaned from the survey population

It also has an ontology for Points of Interest, in which there is information of 2,946 destinations in Cuba, distributed into several classifications, including restaurants, cafes, hospitals and more categories. Each point also has location, category, type of service, among other data, that is linked to user preferences.

The information gathered can then be used to form a database of user preferences forming the elements of the ontology for Points of Interest. This data universe would consist of 176,760 places ranked according to preference.

The validation process of this stage is to know if the recommendations meet users' tastes. For evaluation data to be reliable, it should be carried out in multiple runs and the average of the evaluation metric calculated. This decreases the likelihood of inaccurate algorithm results as the data has been tested.

For this type of validation, test protocols are applied. They determine what data can be used in an algorithm for model building (training phase) and how data should be calculated (evaluation phase). With these considerations taking into the account, the accuracy of the algorithm can be increased.

A configuration widely used in the evaluation of recommender-system algorithms is 80% training, 20% evaluation. This means that 80% of the initial data will be used to build the model and the remaining 20% will be used to verify the effectiveness.

To measure the quality of the generated recommendations, evaluation metrics are used to analyze the strengths and weaknesses, displaying the variables involved and seeing their variations how they affect the different parameters.

To demonstrate the dependence of the training set and the accuracy of the recommendations, we divided this into 8 sub-sets randomly taken from 30% of the data from the total training set. With each of them the method and data was trained, the recommendations were implemented and the mean absolute error rate calculated, using the following expression:

$$MAE = \frac{\sum_{i=1}^{N} |p_i - r_i|}{N}$$

where:

p_i: is the value calculated by the method recommendation.
r_i: is the value that the user has expressed a preference for the element 'i'.
N: is the number of elements in the set.

As seen in Figure 4, more data training of the method reduces the Absolute Mean Error, so we can say that the proposed recommender system improves the quality of their predictions. Even when the amount of data training is very high (more than 80%), there is a tendency to linearity. This may be due to the difficulty of finding Points of Interest that have not yet been visited by the user that could meet their expectations.

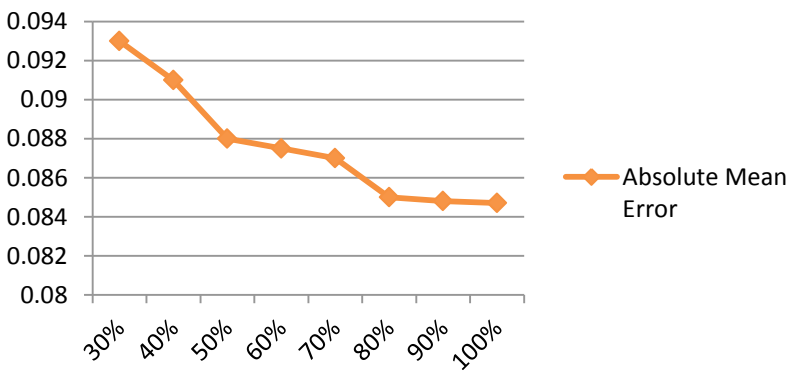

Figure 4. Absolute mean error behavior

To make precision tests, an average of 10 recommendations per user was generated. The metrics used were precision and recall, for its operation needs to define a criterion

of relevance, which is used to decide if the product is relevant or not. In this case, the users set their preferences on a scale of 1 to 10 and products with a score greater than or equal to 7 are relevant, and less than or equal to 6, are not.

The process is as follows: It simulates the recording of each user associated with the test set in the system, entering minimal information known to the user (personal data). Then, the system generates a recommendation, and user selection 'S'. With the intersection of 'S' and the places that the user actually visited, called 'R' (obtained from the questionnaire), we obtain the set of relevant and selected places.

Precision tests were performed using the following expression:

$$Precision = \frac{N_{rs}}{N_s}$$

where:

N_{rs} is the number of relevant items selected by the system.
N_s is the total number of items selected by the system.

Figure 5 shows the behavior of the Precision while performing the tests. We can see that the designed algorithm gained in precision for 4 or less recommendations, indicating its greater ability to recommend products correctly. Therefore, one can conclude that the algorithm is useful in fields which the user receives a list of recommendations, for example, when the system offers recommendations to the user without these previously being requested.

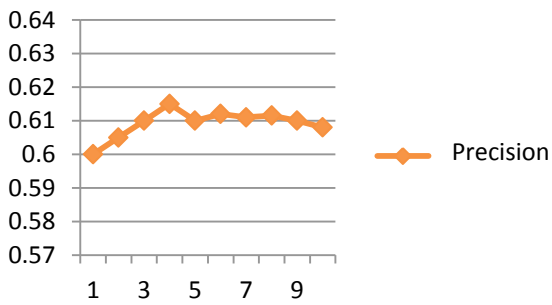

Figure 5. Precision of the algorithms in the experiment

Recall assessment was conducted from the expression:

$$Recall = \frac{N_{rs}}{N_r}$$

where:

N_r is the number of items that the user has classified as relevant.

N_{rs} is the number of relevant items selected by the system.

Figure 6 shows how the proposed algorithm improved in Recall, increasing the number of recommendations. This indicates that the weight is an improvement, which confirms that the users, by performing preference ratings, take into account the characteristics of the Points of Interest.

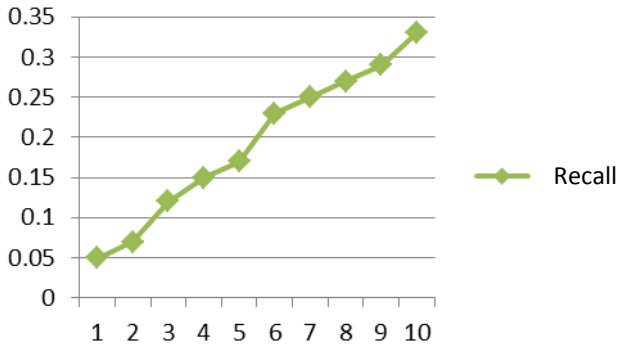

Figure 6. Recall of the algorithms in the experiment

Evaluating the results of the Collaborative Filtering algorithm using metrics Mean Error Absolute, Precision and Recall showed that the predictions calculated are in the range of actual scores given by users.

The case study demonstrated the validity of the proposed model, revealing that, for a number of recommendations less than or equal to 4, better results were obtained. This condition makes the optimal algorithm suitable for use in mobile applications, where the user does not request for recommendations, but are given them automatically.

4. Conclusion

This research method yielded a hybrid recommender system which filters the contextual, semantic and collaborative, thus increasing the analytical capabilities of the SDI. Its application increases the effectiveness of decision-making in SDI environments and analytical skills in the application of mobile systems users.

Mobile devices have become an everyday technology and have evolved into a new paradigm of context awareness. The tools that filter information for users of these devices have emerged as a very useful element, and within them, the recommender systems, play a leading role.

The spatial filters, semantic and collaborative, are key in the method of recommending, based on the preferences of the user and their geographical position at the time of the recommendation.

The feasibility of the proposed method is demonstrated with the implementation of a functional prototype that integrates all the components described in it. Additionally, we implemented a Web application, based on simulations, which allowed its efficiency to be determined.

It was confirmed by an experiment and the application of various metrics, that the algorithm is optimal for use in mobile applications. One where the user does not request recommendations, but recommendations are performed for them automatically.

Acknowledgements

This work it would not be possible without the contribution of many people who collaborated with us on the research, those who kindly reviewed this chapter. In particular, we thank Cristobal Pascual, David Farre, Maykel Wong, Mirella Perez, Andrea Samper, Ulises Pacheco and also our colleagues at GeoSÍ, or GeoMIX agency, for their unconditional support.

References

Adomavicius, G. and Tuzhilin, A., (2005), Toward The Next Generation Of Recommender Systems: A Survey Of The State-Of-The-Art And Possible Extensions. IEEE Transaction On Knowledge And Data Engineering.

Belkin, N.J., (2000), Helping People Find What They Don't Know. *Communications of the ACM*, 43(8): 58-61.

Bernardos, A., (2008), Modelo de Integracion de Tecnologias para la provisión de Servicios Móviles basados en Localización y Contexto, Tesis Doctoral: Universidad Politécnica de Madrid.

Burke, R., (2002), Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12(4): 331-370.

Burke, R., (2000), Knowledge-Based Recommender Systems. In A. Kent (ed.), Encyclopedia of Library and Information Systems. Volume 69, Supplement 32. Marcel Dekker.

Capote, J. L., (2011), Modelo de Servicios Semánticos en la IDERC. La Habana.

Cox S., Daisey P. and Lake, R., (2003), OpenGIS Geography Markup Language Implementation Specification. OpenGIS project document reference.

Delgado Fernández, T. and Capote Fernández J.L., (2009), Marco teórico de Infraestructuras de Datos Espaciales en el Proyecto CYTED IDEDES. Semántica Espacial y Descubrimiento de Conocimiento para Desarrollo Sostenible.

Espinilla, M., Rodríguez, R.M., Martínez, L., Pérez, L.G. and Barranco M., (2009), SR-REJA. Sistema de Recomendación Híbrido.Georreferenciado. Jaén, España.

Henricksen, K. and Indulska, J., (2006), Developing Context-aware Pervasive Computing Applications: Models and Approach, *Pervasive and Mobile Computing*, 2(1): 37-64.

Hong, S.K., (2008), Ubiquitous Geographic Information (UBGI) and Address Standards. ISO Workshop on Address Atandards: Considering the Issues Related to an International Address Standard: 28-37.

Koh, J. and Kim, Y.M., (2000), Mobile Internet Applications Primer, UBS Warburg, Global Equity Research Report.

Lee, D. and Meier, R., (2007), Primary Context Model and Ontology: A Combined Approach for Pervasive Transportation Services, Workshop on Pervasive Transportation System, White Plains (NY), USA.

Martínez, L., Pérez, L.G. and Baranco, M., (2007), A Multigranular Linguistic Content-Based Recommendation Model. *International Journal Of Intelligent Systems*, 22(5): 419-434.

Mayrhofer, R., Radi, H. and Ferscha. A., (2003), Recognizing and predicting context by learning from user behavior. In W. Schreiner G. Kotsis, A. Ferscha and K. Ibrahim, editors, The International Conference On Advances in Mobile Multimedia (MoMM2003), volume 171, pages 25-35. Austrian Computer Society (OCG).

McGuinness, D.L. and van Harmelen, F., (2004), OWL Web Ontology Language, Available from: http://www.w3.org/TR/owl-features/

Nebert, D., (2004), Spatial Data Infrastructure Cookbook v2.0. Technical report, Global Spatial Data Infrastructure.

Nivala, A.M. and Sarjakoski, L.T., (2003), Need for Context-aware Topographic Maps in Mobile Devices, In ScanGIS'2003 – Proceedings of the 9th Scandinavian Research Conference on Geographical Information Science, Finland.

Park, M., Sug Gu, M. and Ryu, K.H., (2007), Context Information Model using Ontologies and Rules Based on Spatial Object. *Communications in Computer and Information Science*, 2: 107-114.

Perez Cordon, L.G., (2008), Modelos de recomendación con falta de información. Aplicaciones al sector turístico. Universidad de Jaen.

Sacher, H., and Loudon, G., (2002), Uncovering the New Wireless Interaction Paradigm. *Interactions*, 9(1): 17-23.

Saha, D. and Mukherjee, A., (2003), Pervasive Computing: A Paradigm for the 21st Century. *IEEE Computer*, 36(3), 25-31.

Schafer, J. B., Konstan, J.A. and Riedl, J., (2001). E-commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, 5(1/2):115-153.

Urbieta, A. and Barrutieta, G., (2007), Estado del arte de composición dinámica de servicios en entornos de computación ubicua II Simposio Internacional sobre Computación Ubicua e Inteligencia Ambiental.

Vaccari, L., Shvaiko, P. and Marchese, M., (2009), A Geo-service Semantic Integration in Spatial Data Infrastructures. *International Journal of Spatial Data Infrastructures Research*, 4: 24-51.

Weiser, M., (1991), The Computer for the 21st century, Scientific American.

CHAPTER 13

A Discovery Geospatial Portal for Promoting Geo-ICT Use in Rwanda

Felicia O. Akinyemi[1], Bernard Hakizimana[2] and Jean Damascene Mazimpaka[3]

[1] Kigali Institute of Science and Technology, B.P. 3900 Avenue de l'Armee, Kigali, RWANDA
[2] Rural settlement task force, Ministry of Local Government, Kigali, Rwanda.
[3] Centre for GIS and Remote Sensing, National University of Rwanda, Box 212 Huye, Rwanda.

felicia.akinyemi@gmail.com; hakibry@gmail.com; mazimpaka@cgisnur.org

Abstract

The level of awareness with regards to the use of spatial data for national development is high in Rwanda. Within the context of Geo-information and Communication Technologies (Geo-ICT), various institutions, especially government ministries and agencies are aspiring to use spatial data in their day-to-day activities in order to fulfill their various mandates. Efforts in this direction include the production of different types of spatial data sets. Alongside, they are organizing customized short-term training and refresher courses for their staff in different aspects of spatial data production and management, as well as geographic information technologies such as GIS, Cartography, Surveying and Remote Sensing. To implement different Geo-ICT based applications and projects, there is increasing demand for the use of geographic information (GI). Considering that spatial datasets are produced by different ministries and the inability of a single department or organization to meet its spatial data needs, it is necessary to share data and information with different organizations. As an option for facilitating this process of data access and sharing, we consider the development of a geoportal to organize Web-based content and services. This chapter describes the content and features of the Rwanda geospatial portal as a discovery portal to promote GI use, access and dissemination. It also demonstrates the technological feasibility of implementing web-based mapping services in Rwanda at minimal cost using both commercial and free open-source software (FOSS) to achieve the goal of performance. The application of FOSS is particularly noted as financial considerations are a major deterrent to developing geoportals and other Geo-ICT applications in developing countries.

KEYWORDS: Geographic information, Spatial data, Geo-ICT, Geoportal, WMS, Rwanda

1. Introduction

Geographic data/information (GI) are increasingly used in society to make evidence-based, intelligent decisions while leveraging on the power of location. Geographic data is as diverse as the phenomenon being mapped and includes non-spatial data for which a geographic reference is needed. In a broader sense, the term GI also includes geospatial data and the services used in providing it, which is referred to as Geo-ICT (Poplin, 2010). Geo-ICT is the addition of the geographic dimension to Information and Communication Technologies. It encompasses fields such as GIS, GPS, Spatial Decision Support Systems, LIS, SDI, Spatial Information Infrastructure, Internet GIS.

GI is becoming more important everyday at all levels of society as it has a central role in supporting economies, improving business effectiveness in the private sector, enabling more efficient governments and increasing citizens' participation in the decision-making process, thus enhancing their quality of life. As virtually everyone is a user of GI, the same information can be used by all segments of society (Genovese *et al*., 2009; Akinyemi 2011). The GI domain is experiencing rapid growth of both computational power and quantity of information, making large geospatial data archives available on the Internet with the availability of platforms such as Web 2.0 technologies (Green 2002; Jackson *et al*., 2009; McDougall 2009).

The emergence of geoportals is a consequence of the need to access and share GI across different platforms. As World Wide Web (WWW) gateways, geoportals organize content and services such as directories, search tools, community information, support resources, data and applications. They provide capabilities to query metadata records for relevant data and services, and then link directly to the online content services themselves. They can also control commercial usage of services by facilitating the sale/purchase of data and services (Maguire and Longley, 2005). Tait (2005) defines a geoportal as a website considered to be an entry point to geographic content on the web or, more simply, a website where geographic content can be discovered. It enables geo-processing interoperability that makes it possible to exchange heterogeneous geographic information content and share a wide variety of geospatial services over the WWW (Yu and Yu, 2009).

Existing geoportals are too numerous and diverse to enumerate, especially in developed countries. However, the emergence of geoportals in developing world contexts, particularly in Africa is slow; Some examples are the Volta Basin Authority Geoportal (http://131.220.109.2/geonetwork/srv/en/main.home) which is a data sharing platform for spatial and non-spatial data on the Volta Basin's water resources management and its related issues (Shumilov *et al*., 2008). UNEP Africa GRID GeoPortal provides national, sub-regional and regional environmental statistics and data in support of environmental assessment and early warning activities

(http://gridnairobi.unep.org/Portal). AISA geo-portal (Africa Institute of South Africa) provides means to search for spatial data sets and spatial data services on socio-economic, demographic and geographical data for the African continent http://www.ai-geoportal.org.za/. Others are SERVIR's East-Africa Geospatial Catalog (http://www.servir.net, Gitau and Grant 2010), AGEOS (Gabonese Agency for Space Studies and Observations) http://www.ageos.ga/en/web/guest/carte-catalogue, Observatoire du Massif du Fouta Djallon GÉOportail http://www2.fouta-djallon-programme.org/FDHWebGis/. Regional geoportals are a good gateway to regional geospatial data but they often lack the local data and local context of countries. There is a need to really catch up on its use and benefits to access GI. On a positive note, institutions in Rwanda already have realized the need for such portals and are developing them as gateway to making data available on different themes. Examples of portals worth mentioning in Rwanda are the AMIS Rwanda (Agricultural Information Gateway) under the Ministry of Agriculture and Livestock (MINAGRI) which is multilingual (http://amis.minagri.gov.rw/map), the Integrated Multi-Sectoral Information System (IMIS http://www.imis.statistics.gov.rw/) and Rwanda Statistical Data Portal (RSDP http://prognoz.statistics.gov.rw/Map.aspx) of the National Institute of Statistics Rwanda (NISR).

Designing and developing a geoportal is no longer a matter of convenience or want; it has become a necessity for disseminating GI and related documentations. GIS communities are recognizing that providing access to geographic content is an important GIS activity that requires a long-term vision in order to realize the possible impacts to society that GIS offers (Cutter *et al*., 2003). This chapter highlights the design of a geoportal tagged, Rwanda Geospatial Portal (RGP) with the aim of facilitating the discovery of geographic and non-geographic information about Rwanda. Also, it is to serve as a discussion platform for GI users locally. The chapter first gives a background to the Rwandan Information and Communication Technology (ICT) context and describes the design components of the geoportal and its salient features. This chapter concludes with the challenges in developing and sustaining the RGP.

2. Rwandan ICT Context

According to the International Telecommunication Union (ITU), Rwanda improved its ICT development index (IDI) ranking at 133 in 2011, from its ranking of 136 in 2010. The IDI is divided into the following three sub-indices, namely: Access sub-index which captures ICT readiness, and includes five infrastructure and access indicators (fixed-telephone subscriptions, mobile cellular telephone subscriptions, international Internet bandwidth per Internet user, percentage of households with a computer, and percentage of households with Internet access); the ICT Use (intensity) sub-index which includes three ICT intensity and usage indicators (percentage of Internet users, fixed (wired)-broadband subscriptions, and active mobile broadband subscriptions); and, ICT Capability (Skills) sub-index which includes three proxy indicators (adult literacy, gross secondary enrolment and gross tertiary enrolment), and therefore is

given less weight in the computation of the IDI compared with the other two sub-indices. For detailed definition of indicators, see the report (ITU, 2012).

Looking at some of the indicators in details, there are 818,048 Internet users as of Dec. 31, 2011, that is, 7.2% of the population. The percentage of households with a computer and with Internet access both almost doubled between 2010 and 2011 to 2% and 5% respectively. Mobile-broadband penetration rose from 1% in 2010 to 6% in 2011. International Internet bandwidth per Internet user doubled, from around 2,000 bit/s in 2010 to over 4,000 bit/s in 2011. This jump is explained by the completion of a 2,300km fiber-optic backbone roll-out in December 2010, linking landlocked Rwanda with neighboring Tanzania and Uganda (ITU 2012, http://www.telegeography.com/products/commsupdate/articles/2011/01/07/rollout-of-national-fibre-optic-backbone-complete/).

Government statistics show that as of 2010, only 159,516 subscribers have access to electricity. Ownership of mobile phones used by both rural and urban dwellers rose to 5,155,697 million subscribers with a mobile penetration rate of 48.1% as of September 2012 from 4,619,429 million (43%) as of May 2012. (Rwanda Utilities Regulatory Agency - RURA 2010, 2012, http://www.rdb.rw/departments/information-communication-technology/overview.html). With ownership of mobile phones and Internet subscribers increasing, the Government of Rwanda (GoR) is seeking to transform the country into a knowledge-based economy and ensure it becomes a regional ICT hub. As part of the GoR's ongoing efforts, it is investing in numerous initiatives to take advantage of ICT's to foster Rwanda's economic development. In recent years, Rwanda has funded computers in schools, built tele-centers (cyber cafes) in every district, ICT buses (mobile connectivity installed in buses) going to remote rural areas to assist local people to access facilities and services online. These ongoing initiatives are aimed at bringing ICT applications closer to the people of Rwanda (Akinyemi and Uwayezu, 2011).

The role of spatial data in national development, social and economic planning has been recognized in Rwanda as far back as 2000. A national SDI (NSDI) was envisaged in the major ICT policies as shown by the NICI-Plan which describes strategies for setting up the national GIS centre and its role (Government of Rwanda, 2006). The major focus in the field of geo-information has been on data and information production. Spatial data is seen as an essential input to implementing various development strategies and activities in Rwanda (Schilling *et al*., 2004). However, the work on spatial data and information discovery and sharing is lagging. Spatial data and information are produced by various government agencies such as the NISR and the Rwanda Natural Resources Authority (RNRA) and research institutions such as the Centre for GIS of the National University of Rwanda (CGIS-NUR). Geo-data are generally available on administrative sub-divisions, physical aspects such as elevation, and data on socio-economic aspects such as infrastructure, population distribution and economic activities. Despite the availability of spatial data and information, more effort is needed to improve its accessibility for use. Currently, geo-data and information are mainly held by their producing institutions. Hence, there is the need to set up clear

policies for data sharing and develop tools for data discovery. For details of the status of NSDI in Rwanda and availability of spatial data, see Akinyemi (2012a; 2012b).

Examples of some Geo-ICT applications are the use of GIS for the determination of Rwanda coffee appellation regions and the eSoko Rwanda initiative (http://www.esoko.gov.rw/, http://repository.uneca.org/tiga/?q=node/50) which is an agriculture pricing information system which enables users to access prices of agricultural produce through the use of mobile phones and the Internet (Schilling *et al*., 2004; 2008; Akinyemi, 2013).

3. Key Considerations and Architecture

Although the awareness of GI value is high, missing are data-access policies, online access points to existing GI and a comprehensive stakeholders' framework to facilitate data sharing. To foster discussion of common issues and data sharing, GI users' discussion platform (formal or informal) was proposed at various times and events in Rwanda. It is with this need in mind that the Rwanda geospatial portal is being designed and developed to spatially enable society.

The acquisition, storage, management and proper dissemination of GI to a wide variety of users are major issues. For the task of disseminating GI, geoportals take up an important role (Put, 2010). The main objective of this current endeavor is to develop the geospatial portal as a discovery portal, that is, a gateway for accessing available GI and featuring Web-mapping services (WMS). It will aid the discovery and use of GI not only for professional GI users but non-professionals alike. It is also to serve as a discussion forum where GI related issues are shared and comments/solutions sought from other stakeholders.

3.1 Key Considerations

Two main aspects that were considered crucial to the effective use of the geoportal are design and implementation.

3.1.1 Design Considerations

The main requirement considered is that the user needs a user-friendly interface that allows for easy navigation and the visualization of maps as well as the ability to interact with the maps (Hennig and Belgiu, 2011). Tait (2005) highlighted a number of design issues: portal sites are usually accessed by users with a wide range of education and technology skills and the site must be simple in design and perform quickly. These two attributes of a portal are key to user acceptance. Based on these requirements, the WMS are designed to provide the following functionalities: 1) Selecting spatial data layers to visualize and make on-demand maps; 2) Interacting with the map through functions such as zooming in and out and panning to aid the user in thoroughly

examining the map content; 3) Reading the geographic coordinates at any point on the map; 4) Selecting a spatial data layer and displaying the attribute values for this layer; and 5) Searching and discovering metadata from linking to the Rwanda Metadata Portal (RMP), 6) Registration on the portal as users. All these functionalities are provided for in the current design. Overall, the design is made to be scalable, which enables more functions to be added as requested by users. For example, in the future, we would extend the functions to allow users to publish and expose their own GI for others to discover.

3.1.2 Implementation Considerations

The two main considerations under the implementation category are interoperability and performance. To ensure interoperability, the technologies and tools being used are based on well-known industry standards. As an example, the work presented in this chapter adopts the Simple Feature Specification (SFS) for spatial data storage, the Catalog Service for the Web (CSW) for spatial data description and discovery, and the Web Map Service (WMS) standard for spatial data visualization. The major tools used in this work implements these Open Geospatial Consortium (OGC) standards with PostgreSQL/PostGIS based on SFS, MapServer based on WMS and GeoNetwork based on CSW. The Web's heterogeneous environment requires that for example, query mechanisms be both software- and hardware-neutral. This creates a need to use mechanisms that are distributed and open (Coetzee and Bishop 1998; Wang *et al*., 2004; Sikder *et al*., 2013). Performance is enhanced by building lightweight websites that are easy to download, even in low bandwidth situations. A portal must minimize the number of user 'clicks' to get to content and, at the same time, maximize the functionality available to the user (Tait, 2005).

3.2 Architecture of the Web Mapping Services

The portal's WMS' section comprises two parts: the frontend and the backend. The frontend provides an interface to use the services while the backend provides spatial data management functionality. A link on the portal's homepage provides access to the frontend part. The architecture of the WMS' section of the geoportal is shown in Figure 3.

The frontend section of the WMS has been implemented using MapServer, OpenLayers library and Apache. Apache is used as the Web server that receives client's request and relay them to MapServer as the map-serving tool. OpenLayers library is used to enhance the performance of the services and to ease the development of the services. OpenLayers library is based on AJAX (Asynchronous JavaScript And XML) principles which enhance the performance of applications in which it is used. For example, the following features of MapServer have been the reason for its choice in this implementation:

- MapServer complies with the OGC's (Open Geospatial Consortium) Web Map Service (WMS) standard.
- MapServer is compatible with the OpenLayers library which also implements the OGC's WMS standard
- MapServer supports additional OGC's WMS requests such as GetLegendGraphic which are needed in this implementation.
- MapServer is widely used by the open source software community.

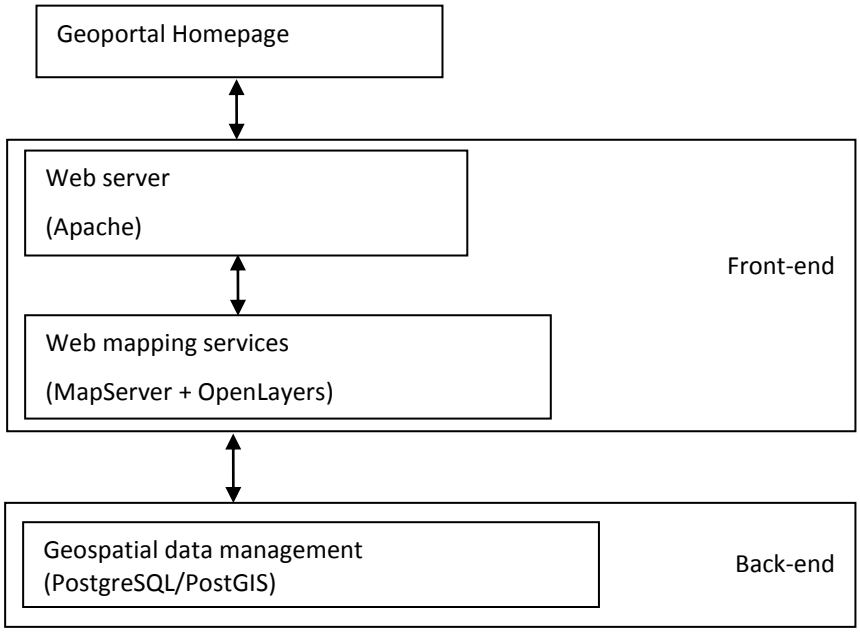

Figure 3. Architecture of the web mapping services' section

The backend is implemented using PostgreSQL/PostGIS is chosen because it is the mostly used open-source Spatial Database Management System for which we can easily get resources and help. This architecture demonstrates the technological feasibility of implementing Web-based mapping services in Rwanda at minimal cost using both commercial and free and open-source software (FOSS). We opted for a hybrid use of software (commercial and FOSS) in order to achieve our goal of performance, which is a key factor in our choice of each technology. The application of FOSS is particularly encouraged as their use significantly reduces software costs and can help build local level Geo-ICT knowledge/technical skills that are required for successful implementation (Herold and Sawada, 2013). Financial considerations are a major deterrent to developing Geo-ICT applications in developing countries.

4. Overview of the Geospatial Portal

The portal is conceptualized as a discovery gateway to facilitate the communication and sharing of geographic data and information about Rwanda. It has functions to search and access a wide variety of information that is potentially of interest to the geospatial community as well as posting news events, particularly latest geospatial related events both locally and internationally (see Figure 2). Figure 2 shows the home page of the portal.

The RGP is hosted at www.geo.cgisnur.org and made temporarily publicly accessible at various times, but it is internally accessible for the purpose of testing. We would need to seek for a domain name which is appropriate for the website to publish it soon. Features on the portal are categorized into the following: 1) A web catalog service linking to the Rwanda Metadata Portal (RMP) which was developed in 2009; 2) Discussion forum; 3) GIS applications drawn from ongoing and past projects in Rwanda; 4) Existing maps that are non-copyrighted or made available free by the producers; 5) Opportunities – this is mainly postings about vacancies in related GI fields; 6) Publications provides links to relevant documentations such as GIS reading materials, scientific papers, reports from projects, etc. that are available online; 6) Rwanda geography mainly focuses on the tourist industry by providing information about Rwanda and links to other sites with guides; 7) Training announces available geographic information science courses that are being organized locally, lastly 8) it has a geospatial business directory with specific focus on Rwanda. Here information about geospatial service and data providers (both public and private sector) can be found. Additionally, there are links to websites of the different organizations. Also, one image is displayed on the homepage on a weekly basis tagged, *the image of the week*. The sample image shown in Figure 2 is the 3D bathymetric map of a section of Lake Kivu, Rwanda. Some of these features are further elaborated on in the next sub-sections.

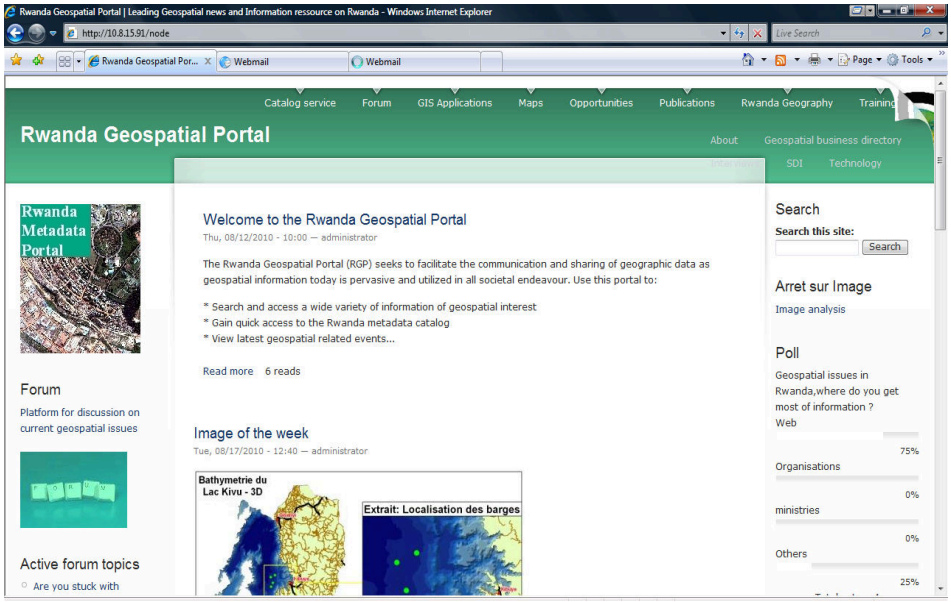

Figure 2. Home page of the Rwanda geospatial portal

4.1 Rwanda Metadata Portal (RMP)

The RMP is a web catalog service accessible at this link http://www.cgis.nur.ac.rw/geonetwork/srv/en/main.home (Figure 3). This portal was developed by CGIS-NUR in 2009 (Akinyemi and Kagoyire, 2010).

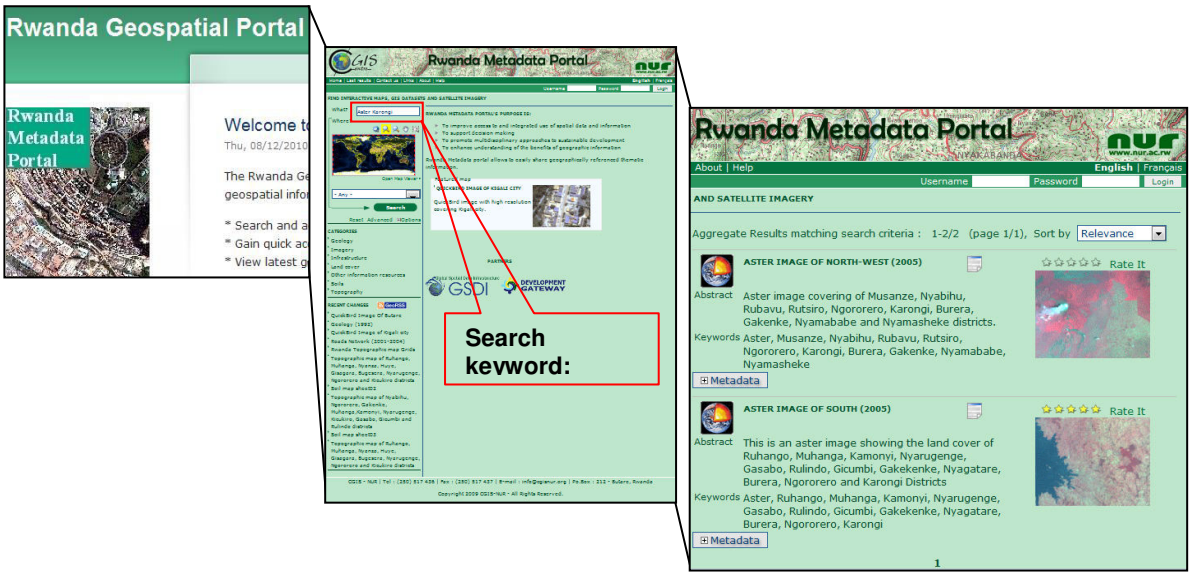

Figure 3. Rwanda metadata portal interface with search results

The main goals in creating the portal were to improve the metadata status of existing geospatial datasets on Rwanda, raise awareness about the benefits of web-based metadata catalogs and lay the foundation for the establishment of a local, sub-national SDI node at NUR. It enables users and producers of GI to locate and evaluate existing geospatial datasets on Rwanda that meet their needs by accessing metadata. This improved the status of metadata, eased the discovery of geospatial data on Rwanda and raised awareness about the numerous benefits of a web-based metadata catalog. The RMP has demonstrated the technological feasibility of implementing open web catalog services in East Africa. The next move in the right direction in further developing the RMP is to enable user's access and/or download actual data. Although GeoNetwork has data access function, the RMP still has this function deactivated (Akinyemi and Kagoyire, 2010).

4.2 Forum for Geospatial Discussion and Polls

This is the platform for discussing burning issues in the geospatial industry (Figure 4). Figure 4 shows some discussion topics such as *are you stuck with projection issues*? D*ata sharing policies, how should they be implemented in Rwanda*? There is possibility to take polls, especially where opinions of users are sought on a particular issue. An example is shown in answer to the question *Geospatial issues in Rwanda, where do you get most information?* (the lower right hand side of Figure 4). The answers to the different options given are displayed such as Web (75%) and others such as verbal communication and referrals from peers (25%). Additional features on the portal consist of posting news and advertisement for marketing purposes (see the folded top right hand side of Figure 4).

This forum is meant to features issues that are particularly of relevance to the local community of GI users (GI professionals and non-professionals). The need for a GI users' discussion platform was identified at several events in Rwanda such as the 2009 GIS day event organized by CGIS-NUR in Kigali and the workshop on Geo-information, ICT, and the Private sector in Rwanda, that was jointly organized by the Lands and Mapping Department of the RNRA (LMD-RNRA), ITC (faculty of the University of Twente) and ESRI Rwanda Ltd. in 2011.

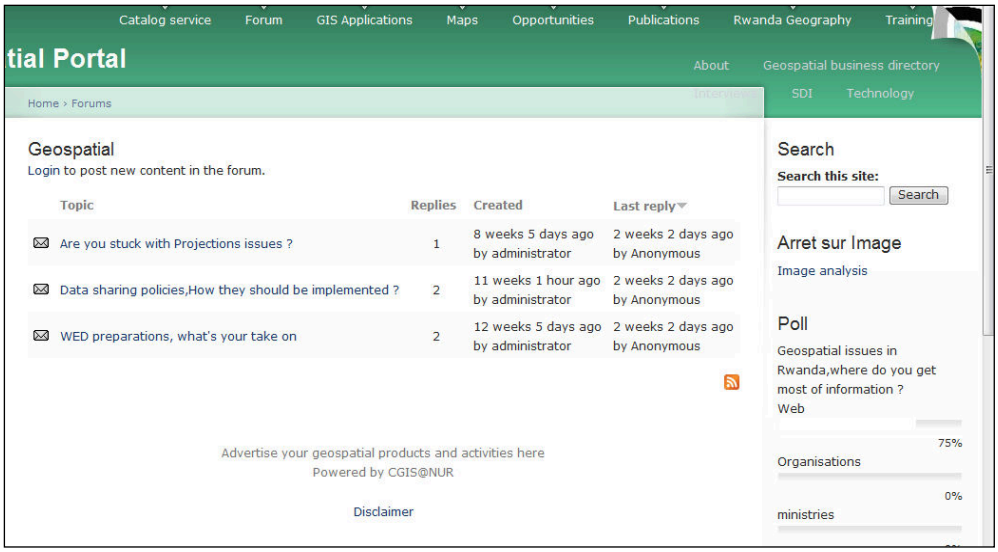

Figure 4. The portal's discussion forum page

4.3 The Geospatial Business Directory

This is to serve as a directory to geospatial businesses in Rwanda. It is realized that although several initiatives are being undertaken in the country, these efforts are often not properly documented. This directory will make information available to people and the industry at one place which could be used as a reference or rallying point for planners, decision makers, academia or every citizen including the common man on the street. This reference could be used to identify organizations (both government and private sector) for procurement of the geospatial information and services. Information includes: 1) Mapping organizations and geospatial data producing agencies; 2) Geospatial data user organizations; 3) Educational institutions providing geospatial education and consultancy services; 4) Geospatial associations or related professional bodies; and 5) Private geospatial companies (both domestic and international which are active in the country). A brief profile of each company/institution could be presented with links to their websites (if it exists) with full contact details.

An output from the WMS tools is shown in Figures 5. Figure 5 is a sample webpage returned by the WMS. This example relates the location of health facilities with access to roads in Rwanda.

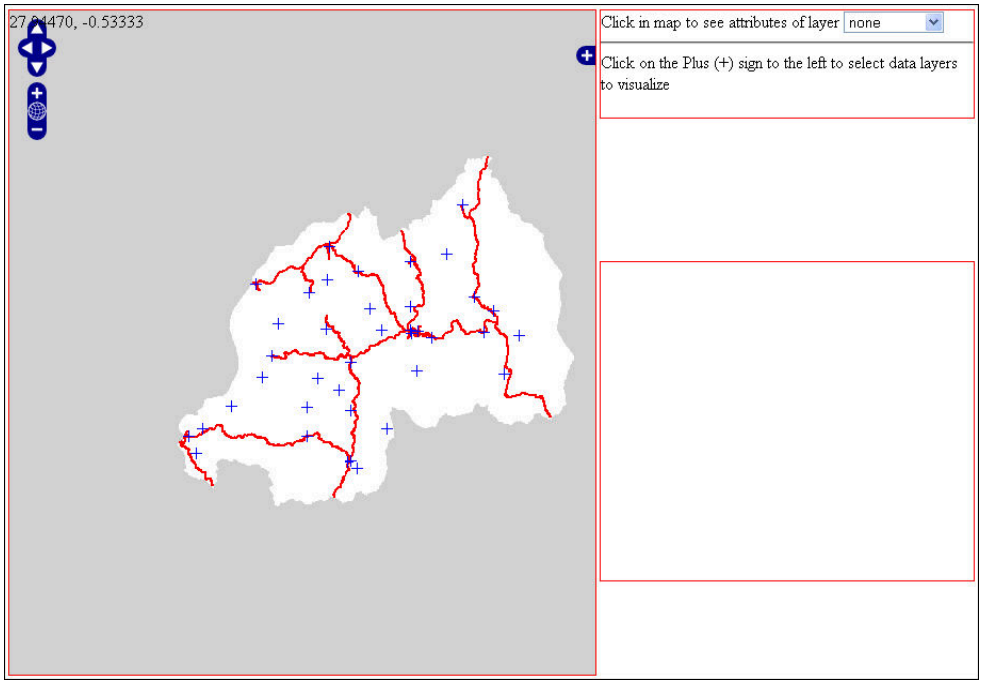

Figure 5. A sample webpage returned by the Web mapping services

5. Discussion and Conclusion

Despite the possibilities with web-map services for example, this option is largely unexplored in Rwanda. This finding was reported by Akinyemi and Uwayezu (2011) based on the survey of 35 organizations using geospatial technologies in Rwanda (government institutions: 66%, consulting firms: 20%, academia: 11%, NGOs: 3%). The problem is not really that organizations are not aware that geoportals exist, it is more of a lack of capably skilled people in GIS and web mapping. Furthermore, they found that as regards the format in which GI is shared, most spatial datasets are shared as map printouts/hardcopies (47%), 32% are shared as shapefiles (.shp), 3% in portable document format (.pdf), and 8% are downloadable from the internet as non-dynamic maps. With 47% of data shared as hardcopy and zero web-map services, the main technical barriers experienced are limited GIS web mapping skills and poor Internet access.

The foregoing reveals the need to show the feasibility and demonstrate the possibility of developing geoportals, especially in Rwanda as an example of a developing country context where the use of Geo-ICT is being promoted. Tait (2005) noted that Web services, service-oriented architectures and distributed GIS are the foundation technologies through which society will realize the benefits of GIS, and geoportals play a key role, guiding the way to the emergence of societal GIS.

5.1 Challenges

Some of the challenges faced in the process of developing the RGP are four fold, namely: 1) Diversity of GI users forced us to ask the question, "who really are the GI users we are designing the portal for"? 2) Up-to-date datasets; 3) Need for partnership to cooperate in GI access and sharing; and 4) Sustainability of the portal.

Who really are the GI users for whom the geoportal is being designed? The fact that nowadays GI users are a diverse lot must be acknowledged. Spatial data users are individuals or institutions that use spatial data. Sometimes spatial data producers are also users when they require datasets produced by other institutions for their activities. With Geo-ICT applied in different fields today, we are dealing with a broad spectrum of users with or without geospatial skills (Oana *et al*., 2010). Since it is very difficult to design to meet the need of everyone, the design of the portal allows for scalability which will enable the portal to evolve over time based on user input and feedback. Geoportals as components of SDI are to be seen as dynamic systems, rather than static systems, with conceptual models used to create SDI frameworks accommodating user requirements which are changing as new environmental, societal or economic conditions and technological improvements appear (Maguire and Longley 2005; Hennig and Belgiu, 2011).

There is a need for updated datasets to serve in the Web-mapping application development and distributed services because decisions made are only as valid as the currency of the data input used. There are issues of copyright, access to data and financing involved here which can be handled with some forms of partnership. Many available spatial datasets which are in dire need of updating are currently being updated or were recently updated. Examples are the topographic maps which were prepared and published in 1988 and the forest cover map produced in 2007. Some datasets are actually available for free and can be downloaded from the respective websites of their copyright owners. A very good example that is worth emulating by other institutions, especially government-owned institutions is the Carte pédologique du Rwanda – soil map at scale 1:50.000 that was produced between 1981 to 2000. The boundaries of these maps have been revised in 2012 and published by Agricultural Information and Communication Center in the Ministry of Agriculture and Livestock Resources (CICA/MINAGRI). These maps are copyrighted, but made available under the Creative Commons Attribution license. Users are free to distribute and modify content as long as the original author(s) or licensor(s) is attributed.

Another challenge is to have a rallying point for the many GI stakeholders in order to maintain their high level of commitment and to get them to contribute to the development of the Rwanda NSDI. The lack of standard collaboration practices within and between government agencies and other stakeholders have been cited as a major bottleneck to ICT adoption, implementation and use (Lance, 2005). This is not a trivial matter, as it requires a lot of commitment in terms of finance, time and expertise. An appropriate mechanism for coordinating this effort to foster data access and sharing in Rwanda is needed.

The long-term sustainability of the geoportal is an issue that begs to be addressed. Financial considerations are essential to be considered, especially to sustain the effort of the geoportal design and development team. Particularly vital is the need for a Geo-ICT technician who will be saddled with the daily upkeep of portal content. This person reviews and approves content submitted for publishing on the portal website as well as edit and validate published content. System usability remains as a challenge because any geoportal is only as good as the content it exposes (Tait, 2005). Many organizations with the mandate to do this publishing do not have the necessary skills or resources in terms of staff and monies needed for the job. These are crucial considerations to address if the goal of widespread GI dissemination is to be achieved.

5.2 Conclusion

The Rwanda geospatial portal is conceptualized to serve as a gateway to GI services and products in Rwanda. It will also serve as a platform for discussion by an audience interested in GIS and related applications. It is designed to offer series of services such as news and information relevant to the geospatial field. Although the primary focus is on the local community of GI users, nevertheless, it will prove equally useful to those online users interested in finding GI and other related information about Rwanda. Issues related to design and implementation such as user-friendliness, performance and interoperability were the main considerations while developing the geoportal.

Role of Funding Source

The financial support from the GSDI (2009/2010 small grant program) and the Rwanda Development Gateway Group grant at the CGIS-NUR are gratefully acknowledged.

References

Akinyemi, F.O. and Kagoyire, C., (2010), The Rwanda Metadata Portal: A Web Catalogue Service. *International Journal of SDI Research*, 5: 1-20.

Akinyemi, F.O., (2011), Evaluating Access to Spatial Data and Information in Rwanda. *Urban and Regional Information Systems Association (URISA) journal*, 23(2): 39-47.

Akinyemi, F.O. and Uwayezu, E., (2011), An Assessment of the Current State of Spatial Data Sharing in Rwanda. *International Journal of Spatial Data Infrastructures Research,* 6: 365-387.

Akinyemi, F.O., (2012a), Towards a Rwandan NSDI: An Update. Hershey, PA USA: IGI Global, *International Journal of Applied Geospatial Research* 3(1): 31-40.

doi:10.4018/jagr.2012010103, Available from http://www.igi-global.com/article/towards-rwandan-nsdi/62045.

Akinyemi, F.O., (2012b), Mainstreaming the Use of Geographic Information in a Developing Country Context: Examples from Rwanda. South Africa *Proceedings of the Geo-Information Society of South Africa conference*, 3-5 Oct.'12, Guateng, SA. Available from www.eepublishers.co.za/images/upload/GISSAUkubuzana2012/Felicia_Akinyemi.pdf.

Akinyemi, Dave T., (2013), The e-Soko Initiative and its Effect on Agricultural Traders in Rwanda – A Case Study of Markets in Kigali. Unpublished IB (International Baccalaureate) Diploma Extended Essay.

Cutter, S.L., Richardson, D.B., and Wilbanks, T.J., (2003), The Geographic Dimensions of Terrorism. New York: Routledge.

Coetzee, S. and Bishop, J., (1998), A New Way to Query GISs on the Web. *IEEE FOCUS,* pp. 31-40.

Genovese, E., Roche, S. and Caron, C., (2009), The Value Chain Approach to Evaluate the Economic Impact of Geographic Information: Towards a New Visual Tool. In B. van Loenen, J.W.J. Besemer, & J.A. Zevenbergen (Eds.), *SDI Convergence, Research, Emerging Trends, and Critical Assessment* (pp.175-188). Netherlands Geodetic Commission 48.

Gitau, J. and Grant, C., (2010), Integrating Regional Geoportals within a Web Content Management System Framework. Proceedings, ESRI User Conference 2010, San Diego (USA), July 12-16 2010. 1-11. Available from http://proceedings.esri.com/library/userconf/proc10/uc/papers/pap_1494.pdf.

Green, N., (2002), On the Move: Technology, Mobility, and the Mediation of Social Time and Space. *The Information Society,* 18: 281-292.

Hennig, S. and Belgiu, M., (2011), User-centric SDI: Addressing Users Requirements in Third-Generation SDI. The example of Nature-SDIplus, *Perspektiv*, 20: 30-42.

Herold, S. and Sawada, M.C., (2013), A Review of Geospatial Information Technology for Natural Disaster Management in Developing Countries. In Information Resources Management Association, *GISs: Concepts, Methodologies, Tools, and Applications*. Hershey, PA: IGI Global. pp. 175-215.

ITU (2012), Measuring the Information Society. pp. 230.

Jackson, M.J., Schell, D. and Taylor, D.R.F., (2009), NSDIs: Coordinating Framework or Battleground for the Management of Geospatial Data. *GIS Professional,* 28: 20-22.

Lance, K.T., (2005), Cross-agency Alignment of Geospatial Investments for Spatial Data Infrastructure Development. Proceedings, *ISPRS Workshop* October 14-16, 2005. Hangzhou, China.

Maguire, D.J. and Longley, P.A., (2005), The Emergence of Geoportals and their Role in Spatial Data Infrastructures. *Computers, Environment and Urban Systems,* 29: 3-14.

McDougall, K., (2009), Volunteered Geographic Information for Building SDI. In B. Ostendorf, P. Baldock, D. Bruce, M. Burdett & P. Corcoran (Eds.), Surveying and Spatial Sciences (pp. 645-653). Adelaide: Surveying and Spatial Sciences Institute.

Oana, C. Vasile, C. and Busuioc, M., (2010), Romanian Geoportal – A Catalyst for the National Spatial Data Infrastructure Activities. Proceedings, GSDI 12, 19-22 October 2010, Singapore. Available from mihai.busuioc@ancpihttp://www.gsdi.org/gsdiconf/gsdi12/papers/3.pdf [Accessed 22 November 2012].

Poplin, A.K., (2010), Methodology for Measuring the Demand Geoinformation Transaction Costs: Based on Experiments in Berlin, Vienna and Zurich. *International Journal of Spatial Data Infrastructures Research (IJSDIR),* 5: 168-193.

Put, S., (2010), Key Elements of the Dutch NSDI. *Sensors and Systems Magazine*. Available from http://www.vector1media.com/articles/features/12823-key-elements-of-the-dutch-nsdi [Accessed 30 August 2010].

RURA - Rwanda Utilities Regulatory Agency (2010), Statistics Information in Energy sector/electricity. Available from http://www.rura.gov.rw/docs/stastics/Energy_statistics_End_%20June%202010.pdf [Accessed 2 July 2012].

RURA (2012), Statistics: Active Mobile Subscribers (90-days revenue generating subscribers) as of May 2012. Available from http://www.rura.gov.rw/index.php?option=com_content&view=article&id=296 [Accessed 2 July 2012].

Schilling, M., Dushimumuremyi, J.P., Bizimana, J.P., Rutamu, D., Kajuga, J., Bizimana, J.D., Maniratora, Z., Schilling, T., Kayitsinga, J.C and Rusabasonga, D., (2004), Coffee of the Thousand Hills. The use of GIS for Determination of Rwanda

Coffee Appellation Regions. First Eastern Africa Fine Coffee Association Conference (EAFCA), Nairobi, Kenya 2004.

Schilling, M., Lyambabaje, A. Mukashema, A., Bizimana, J.P., Schilling, T. and Songer, P., (2008), Geo-IT as marketing Tools for Rwanda Coffee Sector: A contribution to Rural Livelihoods and Economic Development, in Agribusiness Forum 2008: The Agro Industry as Engine for Growth and Wealth in Africa, 18-20th June, FAO Headquarters, Rome, Italy. Available from http://www.emrc.be/en/projects/project-incubator-award/project-incubator-history.aspx

Shumilov, S., Laubach, J., Rogmann, A. and Wittkoetter, P., (2008), GLOWA Volta Geoportal – an active geodata repository and communication system. Available from http://www.glowa.org/eng/conference_eng/pdf_eng/volta/posters/GVP%2052.pdf

Sikder, I.U., Gangopadhyay, A. and Shampur, N.V., (2013), Web-based Geospatial Services: Implementing Interoperability Specifications. In Information Resources Management Association, *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications*. Hershey, PA: IGI Global. pp. 47-65.

Tait, M.G., (2005), Implementing Geoportals: Applications of Distributed GIS. *Computers, Environment and Urban Systems,* 29: 33-47.

Tu, S. and Abdelguerfi, M., (2006), Web Services for Geographic Information Systems. *IEEE Internet Computing,* 10(5): 13-15.

Wang, L.Y., Li, Q., and Yu, W.C. (2004). Web Service and Spatial Geographic Information interoperability [J]. *Science of Surveying and Mapping*.

Yu, Xiao-gao, Yu Xiao-peng, (2009), The Research on Service-oriented Geospatial Portal. *World Congress on Software Engineering*, 3: 286-290.

CHAPTER 14

Spatially Enabling Information to Support Liveability: A Case Study from the North Melbourne Metropolitan Region Australia

Serryn Eagleson and Abbas Rajabifard

The Centre for Spatial Data Infrastructures and Land Administration, Department of Infrastructure Engineering, The University of Melbourne

serrynle@unimelb.edu.au; abbas.r@unimelb.edu.au

Abstract

The importance of creating liveable and sustainable urban environments is widely recognized, and may be why international city rankings and benchmarking systems exist. However, at the neighborhood level, the data required for planners to enact local change and support decision making remain isolated within different local and state government departments.

This project has developed an open-source platform in the context of a Spatial Data Infrastructure (SDI) for accessing and distributing a series of integrated spatial datasets pertinent to the designing of liveable neighborhoods including: transport networks, land valuation, health services and locations of employment. The project has integrated over 100 datasets from disparate sources and now provides them to support researchers from across Australia.

To demonstrate the value of integrated data, four web-based tools have been developed. These tools include an agent-based 'PedCatch' modeling tool for assessing the walkability of neighborhoods, a land supply tool to assess the development potential or land within close proximity of existing infrastructure, an employment clustering tool to assess the agglomeration and spatial clustering of jobs and a tool for the exploration of risk factors associated with Type 2 diabetes and health care locations. Using these tools, the decision makers and users now have the potential to test different scenarios and ask questions to inform the livability of local areas.

The project has been supported by the Australian National Data Service (ANDS) and the Australian Urban Research Infrastructure (AURIN) and contributes to the national agenda for data sharing to improve research about urban environments.

KEYWORDS: Spatial Data Infrastructures, Data Integration, Metadata, Liveability, Walkability, Housing, Health, Employment Clusters

1. Introduction

1.1 Livability and Spatial Information

Defining liveability is a complex task as the concept is highly subjective. However, the common elements which liveable communities share have been broadly identified: that they are healthy, safe and walkable. They offer choices for timely transportation to schools, jobs, services and basic needs. They are cost effective for individuals and local governments. To plan for liveability requires wise decisions about land use and housing coupled with the ability to realize the potential impact of seemingly small decisions at multiple scales and across time (CMAP, 2013).

Several international indices and studies purport to reveal the livability of a City. Two of the most well-recognized systems are the Mercer Quality of Life Survey and the Economist Intelligence Unit liveability ranking system, the basis of these studies is primarily used to calculate the renumeration of expatriates, and the media then use these studies to report the attractiveness of a City. In reality, given the data used to develop these surveys, and the practical application which is for large companies to assign hardship allowances, the practical use of these surveys within cities is limited (Holloway and Wajzer, 2008; Urbecon, 2009; City of Melbourne, 2011).

To actually plan liveable neighborhoods requires small-scale data at the land parcel, building and street level. This type of data is often collected by government departments and stored in database systems. Such information includes: land use, value of property, provision of services and number of people enrolled in a school, the opening hours of a doctor's clinic and location and patronage of transport services. In addition, there are aspects of the socio-economic environment such as the age distribution of residents, gender of workers and origin of visitors. Such data cannot be observed physically, but are geographic in nature and are aggregated to geographic units (Martin, 1996). Just as events occurring at the location have space and time coordinates attached to them, they can also be integrated to understand the use of the place throughout different temporal scales for example: hour, day, week. Figure 1 provides a conceptual diagram of the different types of spatial information relating to a typical urban environment. It is important to recognize that the different activities have data recorded by different departments. For example, in Victoria Australia Public Transport Victoria records data on the patronage of transport, whilst the Valuer

General collects data on land value and use, a combined view is required to plan for a liveable neighborhood.

Figure 1. Spatial information within a neighborhood

2. The Current Problem and Approach

2.1 The Problem

There are increasing concerns about rising rates of serious physical and psychological conditions in the urban populations of developed nations. These conditions include obesity, heart disease, diabetes, asthma, depression and emotional stress. Research shows that urban planning and health patterns are closely related (Butterworth *et al*., 2013) and addressing these issues require integrated planning with a strong foundation of integrated data.

This research project is focused on the North and West Melbourne Region (NWMR) of Victoria. This region covers almost 3,000km^2, includes 14 local government areas and includes four of Melbourne's six Growth Areas. The region currently has a population of 1.68 million, which is expected to grow by over 20% to 2.04 million by 2020 (Department of Planning and Community Development, 2008).

Challenges for NWMR include high levels of disadvantage affecting particular groups from low socio-economic status, limited access to public transport, high rates of unemployment and low levels of access to affordable housing. To address these issues the region has developed a Regional Management Forum (RMF). The role of the RMF is to identify and address critical social, economic and environmental issues facing each region and also to consider the strategic priorities for the region. They also aim to

encourage cooperation between state government departments and councils, and work with statutory authorities, businesses and local communities to set and deliver key priorities. In 2011, the North and West Melbourne Region RMF established an Integrated Data Working Group. The role of this group was to explore the issues and solutions to accessing and sharing data. In doing so the group established a steering group for this research project.

The steering group recognized that the key data required for urban planning, monitoring livability, strategic planning and policy development in the region was fragmented between departments and tiers of government. As a consequence, it is often difficult to conduct research, monitor or forecast with any certainty the local indicators of liveability, for example, the number of jobs, and proximity to open space and facilities. In an absence of data strategic planning, policy development can often be ad hoc, based on partially substantiated assumptions, or delayed until circumstantial evidence demands a reactive response. This significantly impacts on the value of policy; the later an intervention takes place to correct an existing problem then, in general, the less efficacious the intervention is in managing the problem (Ley, *et al*., 2010).

At the same time as missing the data, there is often a mismatch between the research happening within academic institutions and the policy makers who need ready access to the knowledge and tools to effectively make decisions. In response to this problem two national initiatives, the Australian National Data Service (ANDS) and the Australian Urban Research Infrastructure Network (AURIN), both funded by the Australian Government's Super Science scheme, united together to support this project titled *'North and West Melbourne Data Integration and Interrogation and Demonstrator Projects'.* The project aimed to facilitate access to data sets for the North and West corridor of Melbourne and in turn provide researchers, planners, practitioners and policy makers with access to the data. By providing access to data, the project aimed to support focused research to address issues relating to liveability of the North and West Melbourne Region.

The first initiative ANDS, recognizes that research is producing larger and more complex data than ever before, and the imperative to manage and share this data. ANDS commenced in 2008 and has established protocols for enabling research data collections to be more valuable by connecting and supporting the reuse of data and research. Further information relating to ANDS can be access through the website: http://www.ands.org.au/

The second initiative AURIN was established shortly after ANDS in 2009. The objective of the AURIN project is to provide urban- and built-environment researchers with a portal providing seamless access to data and tools for interrogating a wide array of distributed data sets to support multiple research activities within Australian Cities.

2.2 The Approach

The approach taken in this project is to establish a spatial platform of fine-scale data and enable the collaboration between researchers and policy makers. The benefits of this approach are that researchers are able to use the latest data to develop rigorous techniques and scientific knowledge. For the policy makers, data, techniques and knowledge are available to improve decisions related to liveability. The work conducted provides the evidence of a new paradigm of data management with a focus on spatially enabled infrastructure, which integrates urban data from distributed resources. Figure 2 provides an illustration of the framework. Each of the datasets is located with the data custodian and is integrated into the AURIN portal along with metadata which provides information on the coverage, purpose and timeliness of the data along with individual records describing each of the attributes associated with the dataset.

The value of the data is then demonstrated in four policy-relevant demonstrator projects. The demonstrators on walkability, employment, housing affordability and health service, were defined by the North and West Melbourne Regional Management Forum, as tools required to assist with the planning issues of the region. The linkages between the tools then enable multi-disciplinary research teams to provide an evidence-based approach to decision-making. They are supported by an integrated Web-based framework giving access to the datasets in their custodial institutions, and sets of open-source analysis tools. Importantly, the approach is supported by a high-level governance structure, which facilitates access to resources and licensing agreements, as well as dissemination channels, directly to policy makers.

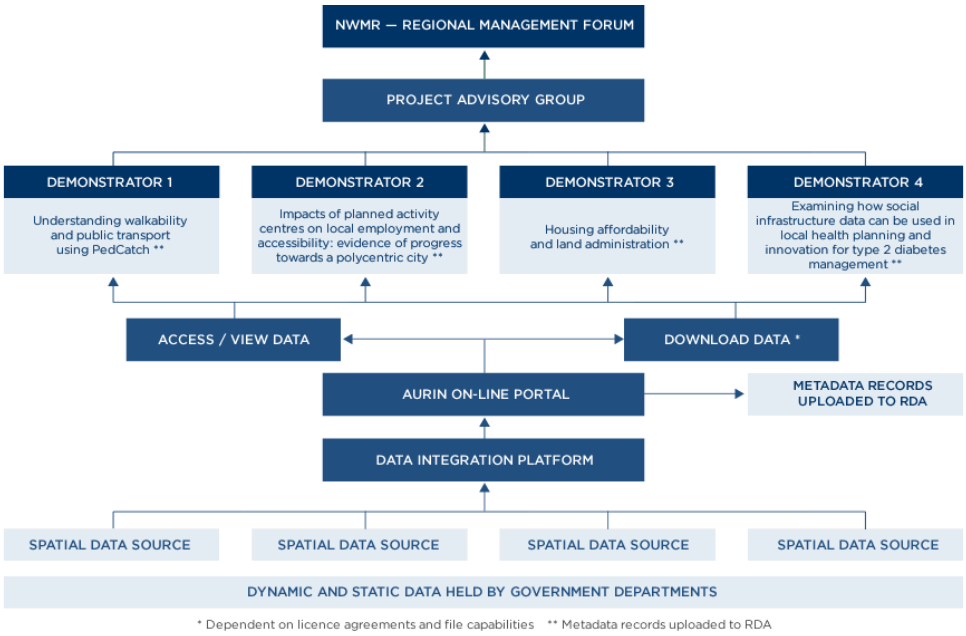

Figure 2. Structure of the data integration project (Source: Eagleson, 2011)

2.3 Project Governance

One of the key elements to the project is the governance framework with appropriate seniority within the academic sector to guide the research coupled with seniority across governments to allow access to data, input into the research design and implementation of the demonstrators within a policy setting. Within academia, the project has been guided by Professors from the following: Centre for Spatial Data Infrastructures and Land Administration (CSDILA), Architecture Building and Planning (AP), Department of General Practice and the McCaughey VicHealth Centre for Community Wellbeing. From the Victorian government's perspective, the project has been facilitated by the North West Melbourne RMF (NWM-RMF) established in 2007 as a resource to strengthen advocacy platforms. The RMF has a mandate to share data with the intention to guide policy decisions and collaborate in integrated planning activities.

The aims of a RMF include strategic priority setting, regional planning, implementation of strategic initiatives and projects, information sharing, networking consultation on major whole-of-government initiatives and projects, sharing lessons and building an evidence base of effective practice. The RMF program has established a collaborative relationship between state and local governments, and provides a mechanism for constructive, regular dialog.

Relevant to this project, the NWM-RMF identified four critical areas where better data

integration were needed – transport access, education opportunities across the life course, housing affordability and health and employment opportunities. In collaboration, researchers from the University of Melbourne, in collaboration with the NWM-RMF, AURIN and ANDS, devised a series of demonstrator projects to respond to these areas of concern and have applied these demonstrators to an integrated data platform.

Members of the NWM-RMF and the University of Melbourne formed working groups to design the data needs and development of the four demonstrator projects focused on walkability, housing, employment and health.

3. The Method

The method applied in this project can be divided into two components. The first was the ability to access and describe spatial data within an integrated platform known as the Data Hub. The second was the demonstrator projects each of which are outlined below:

3.1 Data Hub

To construct a Data Hub required a method and infrastructure for integrating data from multiple sources. It is important to note that the process of integrating data from multiple sources required considerable data cleaning, manipulation and validation to ensure they are 'fit for purpose'. The focus on the data collection has been on data held by Local and State Government agencies which have routinely collected databases. The data hub contained two main components: a GeoServer that had the ability to harvest data; and a metadata tool that provided information on varying levels of metadata from the title and abstract through to the individual attributes. The following paragraphs outline the technical components of each.

3.1.1 GeoServer

The core infrastructure for the project has been the data hub. This hub accesses data through a Web Features Service (WFS) GeoServer (an open-source GeoSpatial server), which supports a large spectrum of Open Geospatial Consortium (OGS) services. Through the GeoServer web-based dashboard, various datasets in various formats and standards could be registered as WFS.

In total, 102 datasets were available and harvested from the Department of Environment and Primary Industries (DEPI) WFS including: VicMap, Public Transport Victoria, and Department of Planning and Community Development (DPCD) datasets. Connectivity to DEPI's WFS server and the existence of all layers was tested by AURIN. In a number of cases, the directly harvested metadata was lacking enough detail and a process of manual enrichment needed to be employed.

For departments who did not have the capability to allow data to be directly harvested an interim solution was established where data was cleaned, formatted and uploaded into a virtual machine. Approximately 50 datasets were integrated using this approach.

3.1.2 Metadata

Metadata contains the information about data and is important for the discovery and sharing of spatial datasets. There are various formats that are designed to structure the way we define metadata. In Australia and New Zealand, the Australian and New Zealand Land Information Council (ANZLIC) Metadata Profile has been developed to enable the consistent collection of metadata across Australia and New Zealand. The profile defines a minimum set of elements that must be collected for spatial datasets and other resources.

In this project, the ANZLIC metadata profile was extended to incorporate the ability to add elements specific to layers' attributes, the new metadata profile (known as a new metadata schema) was created. This was by extending the ANZLIC ISO 19139 metadata profile, called the AURIN metadata profile. This new metadata profile included XSD (XML Schema Definition) files describing the elements of metadata records. Each metadata record needed to be validated against its metadata definition schema. To ease the creation of the metadata record, compliant with its definition schema, some commercial and open-source tools were developed. The most commonly known tool in the open-source world is GeoNetwork, which is supported by a strong community. Further information relating to the technical elements of the project is described in (Nasr and Keshtiarast, 2013).

3.2 The Demonstrator Projects

Utilizing the data from the data hub for the following four demonstrator projects have been conducted. These demonstrator projects have all been developed using open-source code and represent a significant contribution to the four dimensions of liveability: walkability, employment, housing and health.

3.2.1 Demonstrator 1: Walkability Demonstrator Outcomes

Recognizing the importance of walking for health and well-being, planners are in need of spatial tools to map walking paths and test scenarios so as to effect changes to street networks. This project has established an online pedestrian catchment modeler delivered via Web-based mapping tool. The tool includes a scenario testing functionality to enable planners to change walking speeds, assign wait times to road crossings and edit the walking paths to test scenarios. The results of the tool are available as a download for integration into traditional spatial analysis software.

As described by the demonstrator leaders *"…to our knowledge, this is the first open-source GIS tool that allows built environments to be manipulated and evaluated for*

walkability with an animated agent-based simulation within a web interface. Providing a tool such as this to translate research into practice is a substantial contribution to the health- and place-based research agenda". (Badland *et al*., 2013 p. 32*)*

Within the project context, this project also has close links to the other demonstrator projects, being employment, housing and health, as the ability to design environments that facilitate walking from home to local employment and services is critical to the overall liveability of the neighborhood. Further information on the tool is available from the project Blog: http://blogs.unimelb.edu.au/aurinands/demonstrator-1-walkability/

3.2.2 Demonstrator 2: Employment Demonstrator Outcomes

This tool provides a combination of gravity and clustering methodologies to understand the formation of overall sector-specific job clusters across space and time. The tool draws from the data hub an integrated set of pre-processed jobs data. This data has been processed using a combination of journey to work, planning scheme overlays and valuer general data on space use.
This project responds to a consensus among local policy makers, that Melbourne needs to adopt a multi-nodal metropolitan planning strategy in order to foster local economic development and reduce commuting. For decades, metropolitan planning strategies have sought to promote non-CBD centers in Melbourne. The tool further responds to a consensus among economic development planners that ABS data is insufficient to identify local urban clusters for analysis. The tool enables the users to understand whether spatial policies aimed at cluster development have actually resulted in employment clusters. This tool moves us toward examining those policies by providing a framework to identify whether and where local employment clusters have formed (Day *et al*., 2013).

Further information on the tool is available from the following link: http://blogs.unimelb.edu.au/aurinands/demonstrator-2-employment/
To use the tool requires the user to logon to the AURIN portal via: https://apps.aurin.org.au/gate/index.html

3.2.3 Demonstrator 3: Housing Demonstrator Outcomes

The focus of this project is to demonstrate the link between availability of developable land and space and location to infrastructure and planning restrictions. This requires an analysis of Residential Development Potential Index (RDPI) that is essentially land value divided by the capital improved value (Agunbiade *et al*., 2011). Providing the RDPI within an online interface provide ways of analyzing and communicating, the challenges and prospects of discovering developable land for housing.

This tool is supplemented with a set of exploratory analysis parameters relating to land supply and planning activity which have been integrated to enable policy makers to explore opportunities for urban intensification, housing development, change of use,

spatial analysis and analysis of development approvals. This tool provides the infrastructure to identify potential land for development within close proximity to existing infrastructure such as transport routes, education facilities and open space. A second component of the study provides an income-based assessment to regions where housing supply is affordable.

Further information relating to this tool is available from the project Blog site: http://blogs.unimelb.edu.au/aurinands/demonstrator-3-housing/

3.2.4 Demonstrator 4: Health Demonstrator Outcomes

The aim of this demonstrator tool is to improve access to an integrated set of health-related (prevalence and service use) and social and physical infrastructure data to aid policy makers and planners. This tool is based on collaborative work previously undertaken by The Department of General Practice at the University of Melbourne and the Department of Health North West Metropolitan Region (NWMR). This project uses findings from The Care and Systems Experience – Diabetes (CASE-D) project to specify key social and physical infrastructure, health and social services for this population in the NWMR. CASE-D established in-depth case studies about the pathways to, and experiences of, health care for people from disadvantaged backgrounds with Type 2 diabetes.

The result of this project is a dynamic Web-mapping portal which enables exploration of disadvantage indicators to visualize concentrations of vulnerability and physical infrastructure. For example, locations which combine high levels of obesity, socio-economic disadvantage, depression and smoking with an absence of a General Practice.

The results of the Demonstrator will contribute to ongoing partnership work between the NWMR-RMF and the University of Melbourne; the results are currently being presented back to meetings with stakeholders (including the RMF) as well as through conferences and peer review publication. Further information on all the demonstrator projects and information relating to the teams and level of stakeholder engagement is available from the project website: http://blogs.unimelb.edu.au/aurinands/

4. Challenges

The challenges of the project can be divided into the two sections of the data hub and the demonstrator projects. In terms of the data hub, the challenges were as follows:

- **Metadata:** Unfortunately a number of the basic metadata elements were not descriptive enough to meet the AURIN requirements. To overcome this problem, a solution for the manual enrichment of metadata was implemented.

- **Diversity of data:** The wide variety of data formats provided a number of challenges. Data was often presented in spreadsheets in MS Excel and MS Access formats. Spatially enabling these datasets required geocoding tools to be developed and the linking through identifiers.
- **Spatial-temporal data:** Where some of the most complex data are not able to be supported by the WFS currently available. To overcome this issue separate time-slices of the data have been extracted.
- **Lack of service availability:** In the situation where data was not able to be directly harvested, an interim solution has been established where data has been cleaned and formatted and uploaded into a virtual machine, hosted within the Centre for Spatial Data Infrastructures and Land Administration (CSDILA). These datasets have been gathered from various custodians including the Department of Planning and Community Development, Shire of Melton, Environmental Protection Agency, Public Transport Victoria and the Department of Human Services.

The challenges for the demonstrator projects have been the limited time and budget available for developing the tool. The stakeholder engagement process has identified many potential applications for their use. For example, an expansion of the walkability tool could be applied to the evacuation of buildings. Also by extending the data available, it would be possible to incorporate land use characteristics and enable agents to gravitate towards particular land uses, i.e. residential and commercial.

5. Conclusion

This project illustrates how data can be integrated from a series of data custodians, modeled and analyzed at a level of detail and extent not possible in the past. To do this, an integrated data platform had to be developed. The platform has sourced data from 15 different agencies in over eight different formats and provided them into one integrated platform connected with the AURIN portal.

The four demonstrator projects have been connected with the platform and each of them demonstrate the value of combing human scale data with experienced researchers and linking them with policy objectives in the North and West Melbourne Region.

Currently, this first phase of the project is being tested with users and stakeholder training and engagement is underway. A subsequent phase of the project has been funded by AURIN and will include the integration of a further 100 datasets, the benefits of the demonstrator projects will be published in a series of subsequent papers.

References

Agunbiade, M., Rajabifard, A. and Bennett, R., (2011), Land Administration for Housing Production In FIG Working Week 2011, Marrakech Morroco 18-22 May 2011. Available from http://csdila.unimelb.edu.au/publication/journals/ts05b_agunbiade_rajabifard_et_al_4809.pdf

Badland, H., White, M. and MacAulay, G., (2013), Using Agent-Based Modelling to Inform Neighbourhood Walkability. In Rajabifard, A. and Eagleson, S., (Eds) Spatial Data Access and Interrogation to Support Liveability: A Case Study in North and West Melbourne (Ch. 3). pp. 23-38.

Butterworth, I., Giles-Corti, B. and Whitzman, C., (2013), Chapter 2 Perspective 1: Setting the Scene for the North West Melbourne Data Integration and Demonstrator Projects. In Rajabifard, A. and Eagleson, S., (Eds) Spatial Data Access and Interrogation to Support Liveability: A Case Study in North and West Melbourne (Ch. 3). pp. 11-16.

City of Melbourne (2011), International City Comparisons. Available from http://www.melbourne.vic.gov.au/ABOUTMELBOURNE/STATISTICS/Pages/Citymonitoring.aspx

CMAP (2013), Chicago Metropolitan Agency for Planning – Livable Communities Available from http://www.cmap.illinois.gov/2040/livable-communities

Day, J., Sturup, S. and Chen, Y., (2013), *An Open-Source Tool for Identifying Industrial Clusters in a Data-Poor Environment*. AESOP / ACSP 5th Joint Congress 2013 Planning for Resilient Cities and Regions July 15-19, 2013 University College Dublin Ireland.

Eagleson, S., (2011), North West Melbourne Data Integration Project In A. Rajabifard *et al*. (eds.), A National Infrastructure for Managing Land Information – Research Snapshot (Ch. 13). pp. 116-121.

General Practice and Primary Health Care (2011), *CASE-D Research Project Care And System Experiences for people with Type 2 Diabetes in the North and West Metropolitan Region*. Information Brochure Available from http://www.gp.unimelb.edu.au/docs/communityreports/Iss2-Jun2011_CaseD.pdf

Gunn, J., Furler, J. and Palmer, V., (2012), *Care and Systems Experience – Diabetes CASE-D.* Paper presented at the 41nd Annual Scientific Meeting of the Society for Academic Primary Care. Glasgow.

Holloway, A. and Wajzer, V., (2008), Improving City Performance through Benchmarking. Paper presented at International Cities Town Centres & Communities Society Conference, Sydney.

Ley, A., Konstantinidis, G., Owen,C., Kilmartin, C., Allahmoradi, M. and Eagleson S., (2012), *Measuring Liveability: The Contribution of a Census of Land Use and Employment* Presented at the Local Government Researchers Forum, Sydney 20th January 2012. Available from http://www.csdila.unimelb.edu.au/publication/books/AurinAnds/NW_digital2.pdf

Martin, D., (1996), *Geographic Information Systems: Socioeconomic Applications,* London, Psychology Press.

Nasr, A., and Keshtiarast, A., (2013), *Datahub for AURIN and ANDS Project.* In Rajabifard., A. and Eagleson., S. Spatial Data Access and Interrogation to Support Liveability: A Case Study in North and West Melbourne (Ch. 7). pp. 75-82.

Stimson, R., (2011), *The Australian Urban Research Infrastructure Network (AURIN) Initiative: A Platform Offering Data and Tools for Urban and Built Environment Researchers across Australia*. State of Australian Cities Conference, The University of Melbourne, Victoria, Australia.

Urbecon (2009), Improving city performance through benchmarking, volume 1 2009. Available from: http://www.sgsep.com.au/system/files/SGS_Urbecon_1_%202009.pdf

CHAPTER 15

Cloud GIS in Geothermal Resource Data Management: A Case Study of the Kenya Electricity Generating Company

Daniel Waweru Mwaura[1] and Hunja Waithaka[2]

[1] Department of Geothemal Resouce Development, The Kenya Electricity Generating Company (KenGen)

[2] Department of Geomatic Engineering and GIS, Jomo Kenyatta University of Agriculture and Technology

dmwaura@kengen.co.ke; hunja@eng.jkuat.ac.ke

Abstract

GIS technology has revolutionalized data analysis and dissemination in a myriad of disciplines. Geothermal resource development has been a major beneficiary of the budding technology especially within Kenya. Most of the GIS applications for geothermal resource development have been desktop based, utilizing the technology's inherently powerful tools to prospect for geothermal potential.

This project evaluates the shortcomings at the Kenya Electricity Generating Company, where there is great need for harmonizing and sharing centrally stored geothermal data sharing a large number of departments. Currently disseminating this data is difficult due to lack of a well-structured and efficient system.

The project employs cloud computing as a solution for distributed parallel processing of a large set of data, storing and sharing the end results with users from KenGen. This research proposes a system for geothermal GIS data management using Cloud Computing technologies. The project is aimed at providing alternative options to serving GIS applications over the Web. In comparison to costly servers, interoperability, extensibility and performance, cloud technologies have been proven in recent years to meet, and in some cases, surpass the abilities of traditional servers to produce effective and robust Web-based spatial applications.

A prototype system hosted on Amazon cloud servers was created using GeoServer and its potential impact to KenGen was evaluated. The results show that cloud GIS is a

worthwhile venture especially for large organizations which spend a lot of money on regular GIS software and hardware updates. The study encourages adoption of the technology with due caution on common pitfalls.

KEYWORDS: GIS, Cloud GIS, Geothermal resource, Cloud Computing, Web-based.

1. Introduction

1.1 Background of the Problem

A significant proportion of geothermal information is geothermal well-related and location-oriented, thus geographic information systems (GIS) become central to any geothermal information system. In reality, the management of the flow of spatial and attribute data within the geothermal development departments and with other bodies is difficult and time consuming, because the departments and other agencies maintain separate digital records. At present in KenGen, the Internet and GIS are becoming useful means for involving all in having a holistic approach to geothermal development.

In KenGen collection of data for geothermal exploration involves large amount collected within a short period of time. Due to the large data formats involved in the data collection and the scope involved in most cases, the data becomes bulky and is difficult to manage on a single computer or perform different analyses from the same computer. This leads to data being stored in different computers according to the departments.

1.2 Geothermal Potential in Kenya

According to Ofwona (2006), Kenya is endowed with geothermal resources mainly located in the Rift Valley. Electricity demand in Kenya has continued to grow steadily over the years and has caused great pressure on the conventional sources of energy like hydropower, which is normally affected by weather changes. It is estimated conservatively that the Kenya Rift has a potential of more than 10,000 MWe of Geothermal Power.

Exploration first started by drilling two wells in 1956 in Olkaria I and was followed by increased interest in the 1970s. Initial production started in 1981 when the first plant of 15MW was commissioned in Olkaria I. Currently 45MWe is generated by Olkaria I geothermal power station; 110MWe is produced from Olkaria II (both operated by KenGen) and an IPP is producing 40 MWe at Olkaria III. KenGen and the IPP produce a total of 200 MWe of geothermal energy and this is expected to increase to 1,200 MWe within the next 20 years (Ofwona, 2006).

1.3 Geographic Information System (GIS)

Geographic Information System (GIS) applications often involve acquiring and processing data from multiple sources followed by intensive spatial computations provided by expensive computer systems. The exact same data are hosted in different locations and need to be processed the same way many times when used by different parties; and in many cases, in order to process or conduct spatial analysis over these data would require expensive investments in hardware, software and training of personnel.

Traditionally, GIS applications would need dedicated clusters and storage space to host large amount of data. With the help of Cloud Computing, this processing and storage responsibility can be offloaded to a Cloud service provider. The user can just use a Web interface to control the execution and flow of data. The end results could then be obtained with minimal user intervention (Rajkumar, 2009).

1.4 Cloud Computing

According to the National Institute of Standards and Technology (NIST): Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models.

2. Problem Statement

The Kenya Electricity generating company KenGen under its Geothermal Resource Development division has the mandate of provision of adequate clean reliable geothermal energy for Kenya. In an ambitious plan, the company seeks to develop about 1,000 MWe of geothermal energy by the year 2018. This is to be done in Olkaria geothermal field. In order for this goal to be realized, all departments within the division have to work together generating voluminous amounts of data that have to be analyzed and output of information provided to fast track geothermal development.

Currently there is no centralized data storage system provided and thus it becomes difficult to know where to locate data from the different departments. Procurement procedures can take years in order to upgrade hardware of computers or even the software contained within. This ultimately leads to problems when the current existing machines in the Geomatics section become overwhelmed with the many resource thirsty tasks involved in the processing and visualization.

On the Geomatics section specifically, problems are experienced in; data acquisitions, which are usually expensive and time-consuming; software and hardware needed to run GIS applications, which are usually expensive and require professional knowledge

to set up and use. The process of acquiring, storing and pre-processing data are time-consuming, expensive and often lead to unnecessary data replications. Spatial analyses often require expensive software and computer systems. Furthermore, for the general KenGen staff, it is not practical for a non-professional occasional user to acquire several gigabytes of data and to spend thousands of dollars on professional GIS software.

Moreover, when data needs to be disseminated to end users in order to make key decisions with regard to geothermal development, the process is slow, usually taking the form of exchange of hardcopy maps or soft copy maps emailed to interested clients.

3. Overall Objective

The overall objective of this study is to design and build a cloud GIS prototype that enables the effective management and utilization of geothermal resource data by KenGen employees and highlight the advantages it offers in comparison to the existing system.

3.1 Specific Aims

1. Analyze and understand cloud computing and its potential for GIS applications within KenGen.
2. To investigate the issues involved in development and implementation of a cloud GIS.
3. Explore a solution to host and serve large volumes of GIS data efficiently and speedily.
4. To compare the created Cloud GIS prototype with the current installed GIS infrastructure at KenGen.

3.2 Research Questions/Hypothesis

1. How could KenGen and other organizations use cloud GIS to help optimize, improve, or enable geospatial data dissemination?
2. What are the fundamental designing, implementation, and application issues for spatial cloud computing?
3. How is the cloud approach different from the traditional approach?
4. Is there any savings made in time and cost by use of cloud GIS infrastructures as opposed to the traditional GIS system especially as relating to geothermal resource development?

4. Study Area

Olkaria is home to Africa's largest geothermal project with an installed capacity of about 260 megawatts. Located in Naivasha, it is about 140km from Nairobi, the capital of Kenya, and is an area of recent volcanic activity surrounded by many volcanoes such as Suswa and Longonot. In theory Olkaria has a geothermal potential of about 7,000 megawatts and is central to Kenya's geothermal projects expansion with many future expansion development activities targeting the area.

The Olkaria geothermal field covers an area of approximately 70km^2. The main production zones are generally between 750-900m depth below the surface which are steam dominated and from 1,100m to 1,300m which is richer in water though steam can be intercepted down to the full depth of the well (Ofwona, 2006).

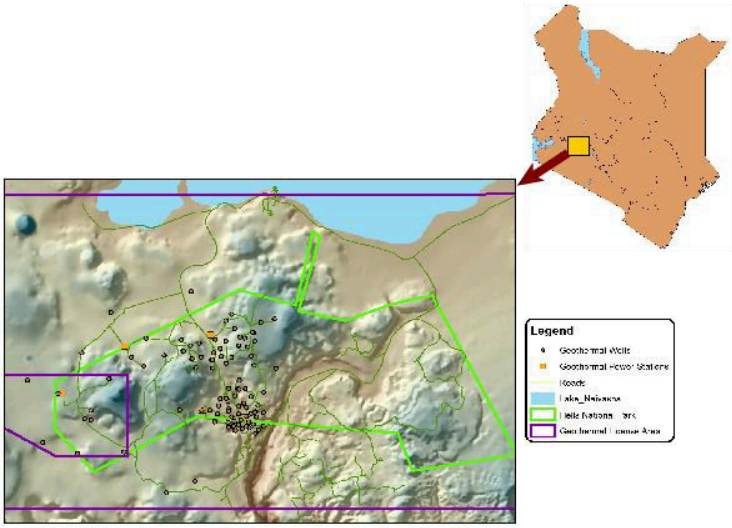

Figure 4. Study area

5. Methodology

5.1 System Design

5.1.1 Existing System

The existing system for management utilized ArcGIS version 10. KenGen staff collect data from the field and convert it to shapefiles which are uploaded to a folder. If a client wants to get a map one can either use the specific computer into which the data was copied to, or copy the shapefiles and use them to create another independent map for the client.

Visualization for the end users is in form of hard or soft copy maps which are emailed or delivered in person. There is a lot of replication observed in the system and this is justified by the need to have backed up copies of data. Most of the computers in the department are frequently affected by viruses which paralyzes operations in the department.

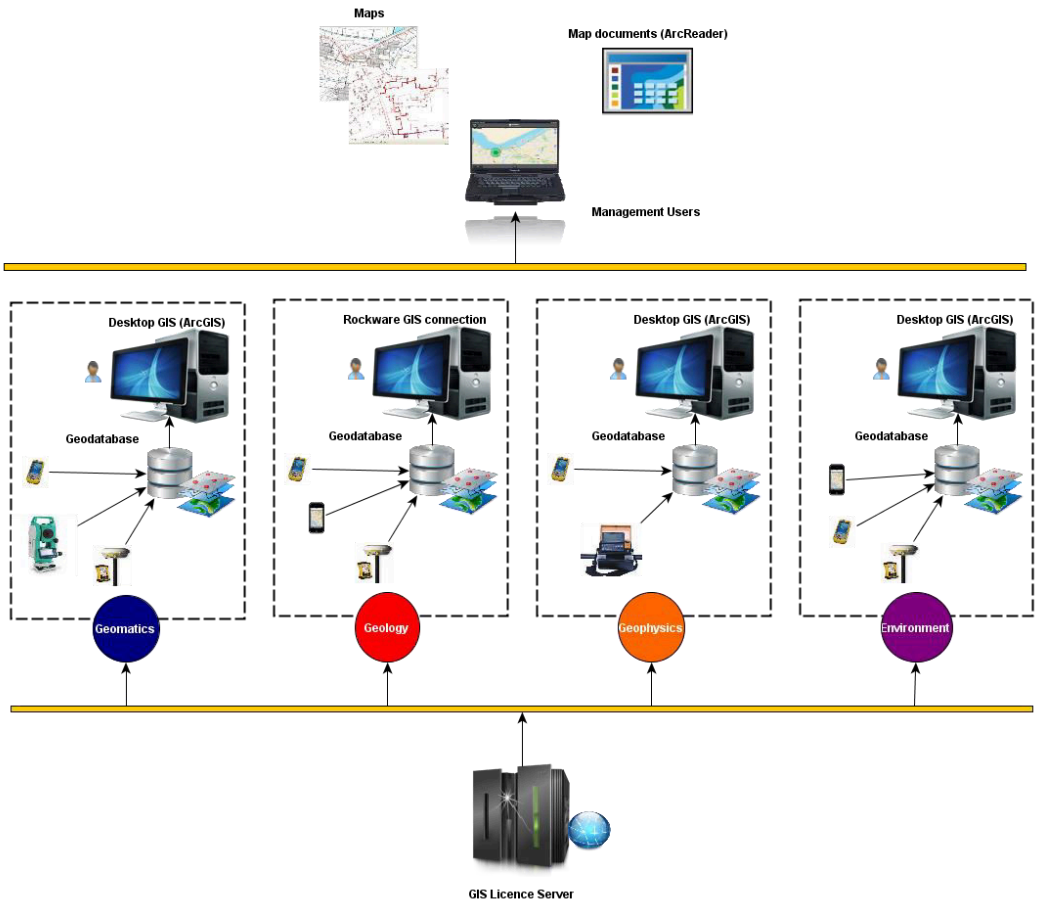

Figure 2. Current system

5.1.2 Users' Requirements

The general requirement is the information management system which could show the information of geothermal resources and infrastructure by Web map. There are two types of end users: the GIS department staff and the general KenGen staff. The GIS department staff are well conversant with most GIS interfaces and tools while the general staff do not have this privilege, so the user-friendly interface is very important. For the GIS staff, a centralized storage for all data is required and interfaces to

manipulate information contained therein. If so, it needs to be compatible with the original database as much as possible. The information can also be managed by the Web interface efficiently. For the general staff, the Web map showing the information should be easy to use and the interface should be user-friendly.

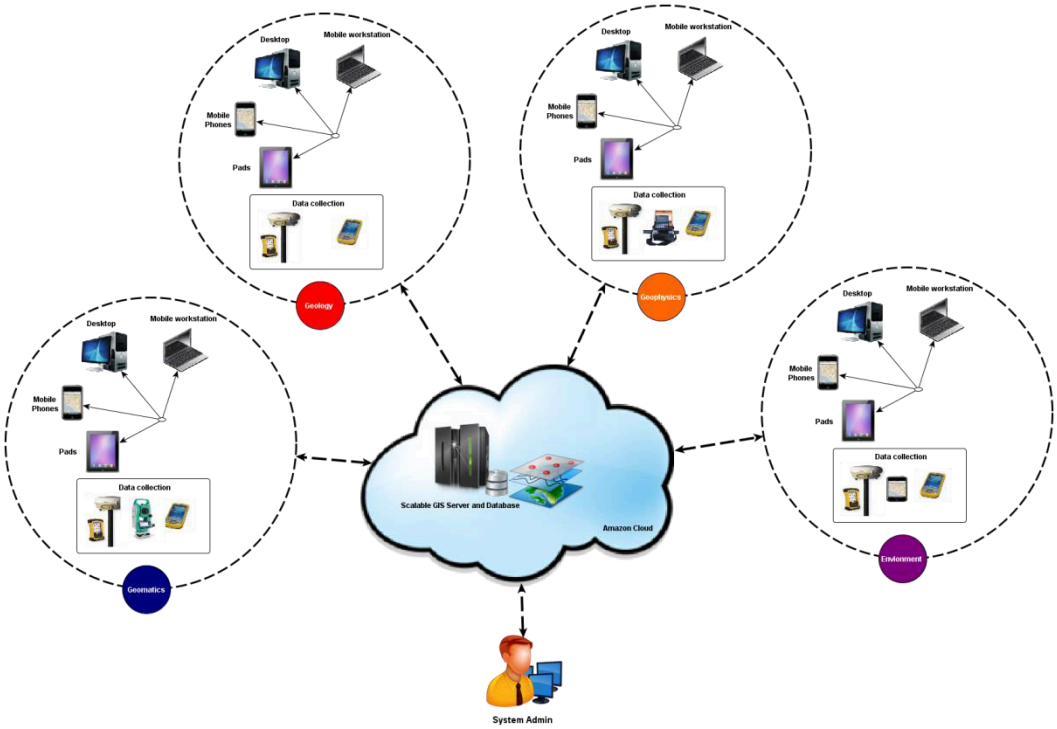

Figure 3. Proposed system bridged using cloud computing technology

5.1.2 Scope and Limitations

The design of the cloud system will utilize the Infrastructure as a model and specifically employ open-source GIS within the cloud environment. The main idea of the study will not be developing a high end and complex GIS on the cloud, but a basic prototype that can be used as a guide for a fully fledged system.

5.2 Data

5.2.1 Sources of Data

The data to be used for the research was entirely derived from the Kenya Electricity Generating Company's computers from the different departments that exist. This includes; Geochemistry, Geology, Geophysics, Geomatics, Environment and Drilling. The data will include geothermal wells, boundaries, power stations, power lines, steam

pipes, geothermal fields, fumaroles, roads, brine collection ponds, geologic maps and geophysical maps.

5.2.2 Collection of Data

Data collection techniques involved extracting data from existing databases from the Geomatics section. Data stored in files from the other departments were picked and assembled in one central computer.

As the main concept of the research is to develop a prototype cloud GIS, spatial data will only undergo minimal analysis. This will involve simple searches. The data collected from questionnaires will be compiled in a table and analyzed at the end of the development and a graph of similarities and differences plotted.

5.3 Architectural Design

The design of the whole system can be summarized into four basic components:

5.3.1 Cloud Service Provider

There exists a number of commercial cloud services providers. Out of this host of providers, the study was narrowed down to Amazon as the provider under which the architecture would be based. The Amazon Elastic Compute Cloud (EC2) is an Infrastructure as a Service (IaaS) cloud. This means that it provides computing power and resources that you can use for a fee. You take care of running the software; Amazon EC2 provides the hardware, (Amazon, 2012). Conceptually, the system has all components installed on a virtual computer provided by Amazon at a fee.

5.3.2 Database Server

A database was considered to be one of the most important components of the whole system. The design of the database took into consideration the ability of current database servers and their ability to handle spatial data. There exists powerful open-source solutions that were given first priority due to their effect in reducing overall cost of the project. PostgreSQL server was taken to be the systems database server due to strengths which go beyond the scope of this project. PostGIS component was incorporated into the database server to spatially enable it. All layers were to be loaded into a database created in this server.

5.3.3 Map Server

A map server is responsible for serving the map output from the Web server into the browser's user interface where the user can interact with it. For this case, the design utilized Geoserver due to its nature of being readily installed in a virtual computer as most of its components can be configured via a Web browser. All layers conceptually

were to be imported form PostgreSQL into the Geoserver layer stores ready for use by the user interface.

5.3.4 User interface

Using OpenGeo suite, the component geoexplorer was used as the basis for building the client user interface. This component employs GeoExt as the base for creating most tools. Due to the scope of the project it was not deemed necessary to create an interface from scratch and thus geoexplorer was an important starting point in the final user interface.

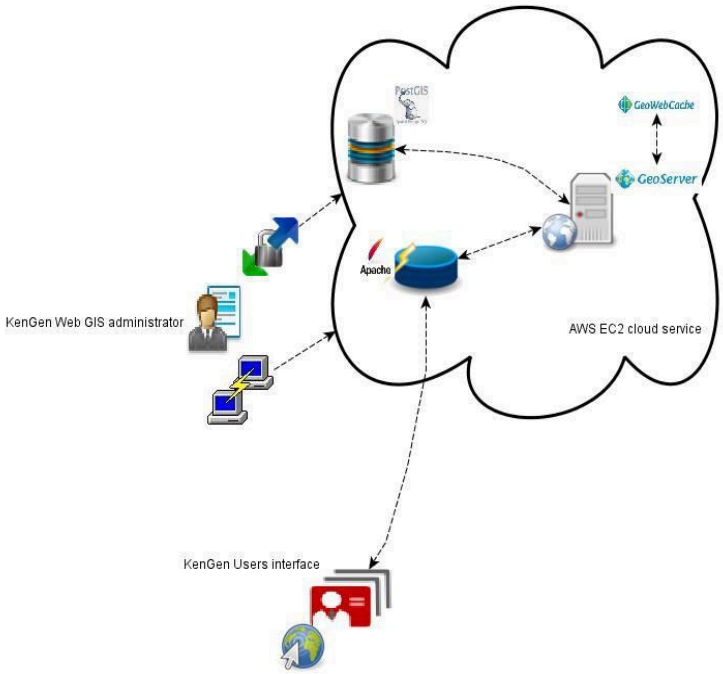

Figure 4. System architecture

5.4 GIS Cloud System Implementation

5.4.1 AMI Creation

The first step in the whole implementation process was to create an instance of a machine on the Amazon cloud infrastructure. This involved using an AMI provided by OpenGeo to the public. An Amazon Machine Image (AMI) is a template that contains a software configuration (for example, an operating system, an application server, and applications). From an AMI, you launch instances, which are copies of the AMI running as virtual servers in the cloud (Amazon, 2013).

You can launch multiple instances of an AMI, as shown in the following Figure.

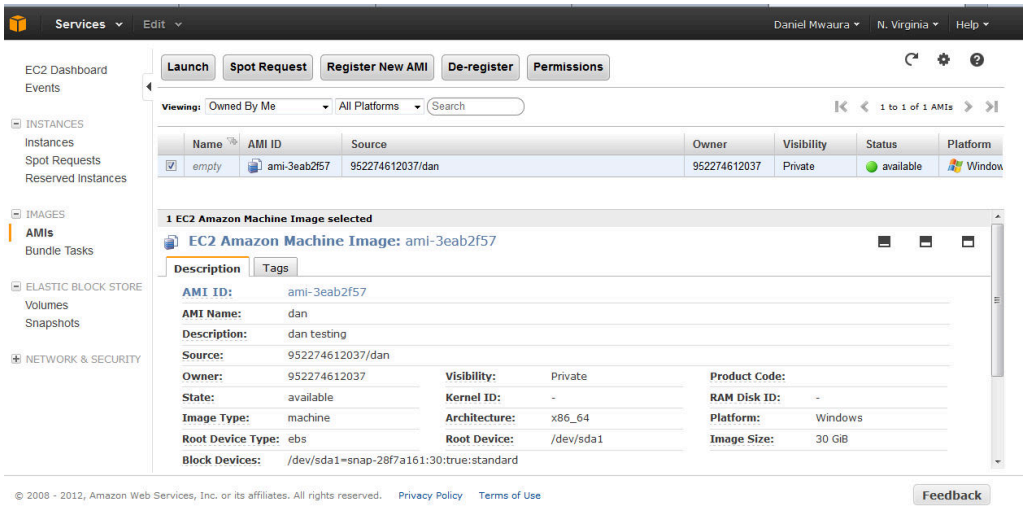

Figure 5. AMI description screenshot

Amazon publishes many AMIs that contain common software configurations for public use. In addition, members of the AWS developer community have published their own custom AMIs. You can also create your own custom AMI or AMIs; doing so enables you to quickly and easily start new instances that have everything you need. For this case, the application is a website and a geo-Web service, the AMI thus included a Web server, a GIS application server a database server, the associated static content, and the code for the dynamic pages (Amazon, 2013).

5.4.2 Launching the Instance

After creating the AMI, an instance is launched in the AWS management console. Once the instance was launched the Web server started, and the application was ready to accept requests. An instance type essentially determines the hardware of the host computer used for the instance. Each instance type offers different compute and memory capabilities. The instance type selected was the one with least cost offering minimal computing capabilities.

5.4.3 Regions and Availability Zones

Amazon has data centers in different areas of the world (for example, North America, Europe, and Asia). Correspondingly, Amazon EC2 is available to use in different regions. This specific instance was launched in the US-East region, the default adopted region. Each Availability Zone is engineered to be isolated from failures in other

Availability Zones, and to provide inexpensive, low-latency network connectivity to other zones in the same region (Amazon, 2013).

5.4.4 Instance Store

All instance types, with the exception of Micro instances, offer instance store, which provides your instances with temporary, block-level storage. This is storage that is physically attached to the host computer. The data on an instance store volume doesn't persist when the associated instance is stopped or terminated. This was the type of storage that the system was based on (Amazon, 2013).

5.4.5 Networking and Security

The instance was launched into the Amazon EC2 network space and assigned a public IP address. To control access to the instance, a security group was set up. This is analogous to an inbound network firewall that enables you to specify the protocols, ports, and source IP ranges that are allowed to reach your instances. The following is a screenshot of the network group settings.

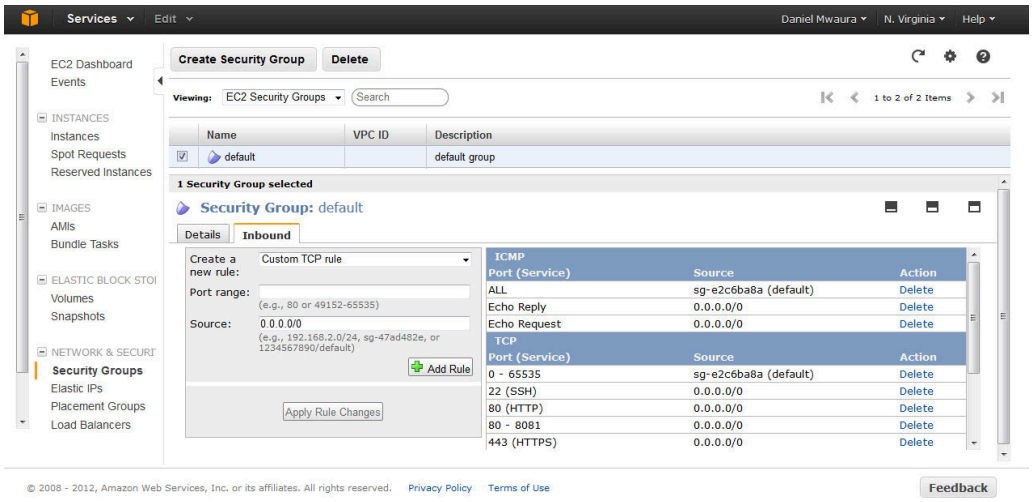

Figure 6. Security configuration on the administration panel screenshot

5.4.6 SSH Key Pair

One must create an RSA public/private key pair. This is used to ensure that only the administrator has access to instances which are launched. AWS doesn't store a copy of the private key. Amazon EC2 only stores the public key, and associates it with a friendly key pair name. Whenever an application is launched using an instance using the key pair name, the public key is copied to the instance metadata. This allows access to the instance securely using your private key (Amazon, 2013).

5.4.7 Transfer of Files

The main protocols that were employed for transfer of files were FTP and SFTP. FTP protocol provides a standard way to transfer files between computers, regardless of the operating system on each computer (Windows, DOS, Macintosh, UNIX, etc.). It is the most common protocol used for exchanging files between computers on the Internet. Secure FTP (SFTP) is similar to FTP, but with SFTP the entire session is encrypted, so that passwords are never sent in the clear, and are therefore much less vulnerable to interception.

5.4.8 Spatial Database Design and Creation

PostgreSQL database server was the main database server that was employed for the GIS backend. This is because of its inherent capability to easily support spatial datasets. A database scheme was created in PostgreSQL using PgAdmin client provided for in the software. This was logged into using SSH. All data that was acquired from existing databases was prepared to be in form of shapefiles and then imported into the database using PostGIS shapefile loader.

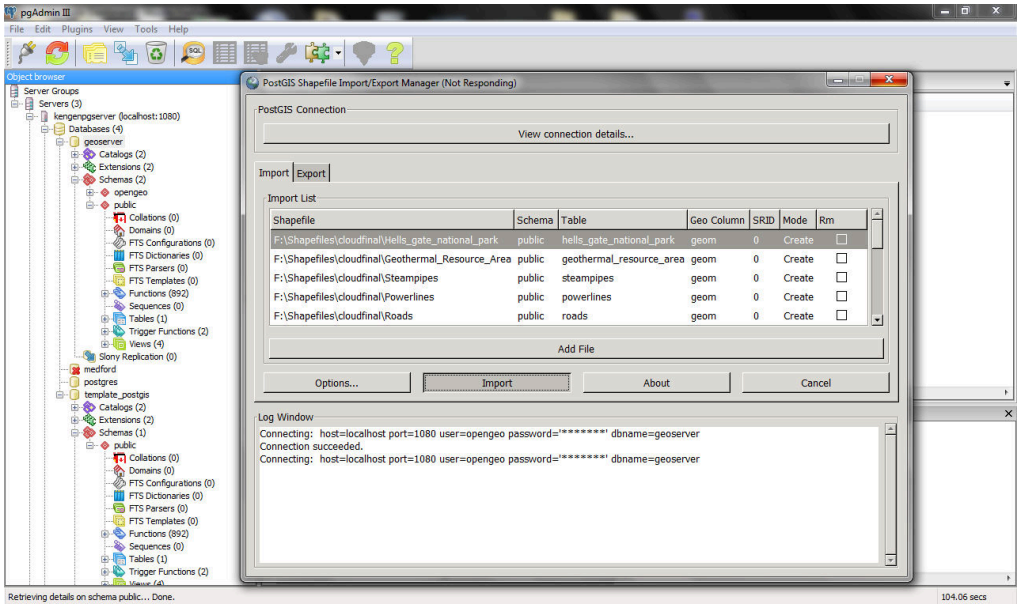

Figure 7. Database configuration on PgAdmin

5.4.9 Geoserver Import

After all the layers were brought into the database server they were then imported into Geoserver, the Mapserver employed for the application as earlier discussed. OpenGeo provides an interface which can be accessed for the default Apache Tomcat port using any Web browser on the public IP provided by Amazon. In this case, the port used by Tomcat was 8080. This is as illustrated below. From the Geoserver interface the spatial reference of the layers was set. In this case it was EPSG: 21037 the spatial reference code for UTM Zone 37 S using the Arc datum 1960.

5.5 Interface Results

After all the data was successfully configured onto Geoserver, OpenGeo's frontend client geoexplorer was employed as the main client to visualize the layers and add various functionalities for interacting with the data. This interface is created using Geoext. The main map panel is shown in the center of the webpage, with a toolbar panel attached under the title of the panel, and the control panels in the left of map. Next to the map panel, it is the panel including two sub-panels which are for layers and legends in the left side, while in the right side is the panel providing information based on querying. The components in the website are described in details as following:

- **Main map**: It is located in the center of the webpage, and is the main component to show the geographic information and link non-geographic information of Geothermal resources.
- **Zoom bar**: It provides the buttons of zoom in and out on the top and bottom of the bar respectively.
- **Navigation Panel**: It provides panning capabilities. The functionality is made simpler and quicker by rolling the mouse.
- **Toolbar**: In the toolbar, the first three buttons provide the functions related with zooming and moving. One advantage of zooming buttons is that it could zoom in or out either by one simple left click to zoom one level up or down, or by drawing the box to zoom in or out to the selected areas. The left and right arrow allows going back or forward in the the history of main map. It is designed to save time for user to get the map from history whenever they want. The left panel next to the main map includes layers and legends, and its width can be adjusted by the edge or hidden by the arrow button on the top.
- **Layers**: It is the place to list available layers in the map. Considering that the list of layers is long, the list is grouped by their characteristics. This includes other external layers from providers such as MapQuest and Google. Only one of base map can be selected (Adopted from geoexplorer, 2013).

5.5.1 Built-in Tools for Graphical Styling and Editing

The styling tool has a rule editor, where one can set options such as color, opacity, and shape. One can also set conditions for display, such as scale rules. The results are saved directly back to GeoServer and are displayed in real time.

5.5.2 Adding and Removing Layers from the Map

This allows one to actually upload shapefiles and PostGIS layers directly into GeoServer through the interface.

5.5.3 Makes Use of Server Caching

The interface uses GeoWebCache, the built-in caching server in GeoServer, to cache tiles on the fly. To avoid stale tiles, when a change happens to a layer in the interface (via styling or editing), a request is sent back to GeoWebCache to truncate the cache. This makes serving of data very fast (Geoexplorer, 2013).

5.5.4 Export Maps to PDF

With this interface, one can compose a map and click the Print button to export the map view as a fully vectorized PDF. While the tool is still a bit rudimentary for professional map-publishing standards, it is sufficient for basic uses.

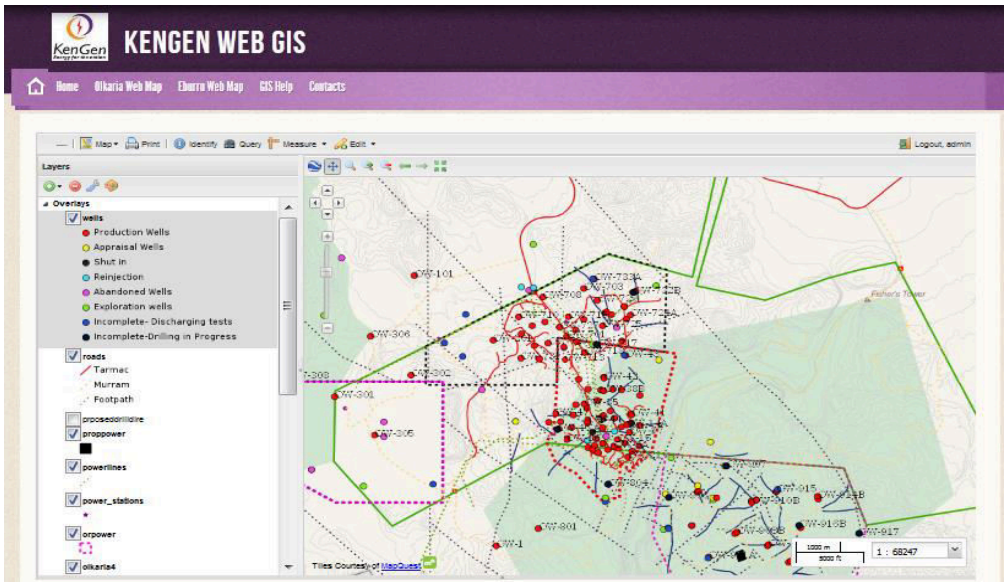

Figure 8. Web GIS interface

5.6 Estimation of Cloud GIS Application

The instance that the application is using has been running for 368 hours. The instance is charged on an hourly basis or the nearest partial hour. The charges are on an hourly basis while utilizing OpenGeo software. The Developer small instance size was used for the test application and thus incurred a cost of 0.13 USD per hour. The usage as of today adds up to around 48 USD which translates to around 4,000 Ksh. This instance has purposely been left on to incur the maximum charges in order to work with the upper limits when it comes to estimating the cost. When the currently existing system is in use, computers are normally replaced after every two years to keep up with current processor speeds.

Current system	Cost (USD)	Cloud system	Cost (USD)
20 Computers replaced every two years	150,000	20 computers served by one server in the cloud at 0.13 per hour	11,388
Software renewals	440,000	20 Licences at 0.13 per hour	227 ,760
Total costs	**590,000**		**239,148**

Table 5. Cost comparison of the existing and proposed system
Assumptions: Software runs 24 hour per day for the ten-year period

5.7 Performance Comparisons

The proposed system is significantly different when compared to the existing system. This meant that it was difficult to find a standardized test for performance comparison.

Current system	Cloud system
Decentralized with a hub in each department	Centralized with one hub in the GIS department
Desktop based	Web based
Uses commercial software	Uses open source software
Uses many local servers	Uses one cloud based server

Table 2. Components comparison of the existing and proposed system

The following performance graphs shown gives the hourly network request and response rates of the cloud-based system.

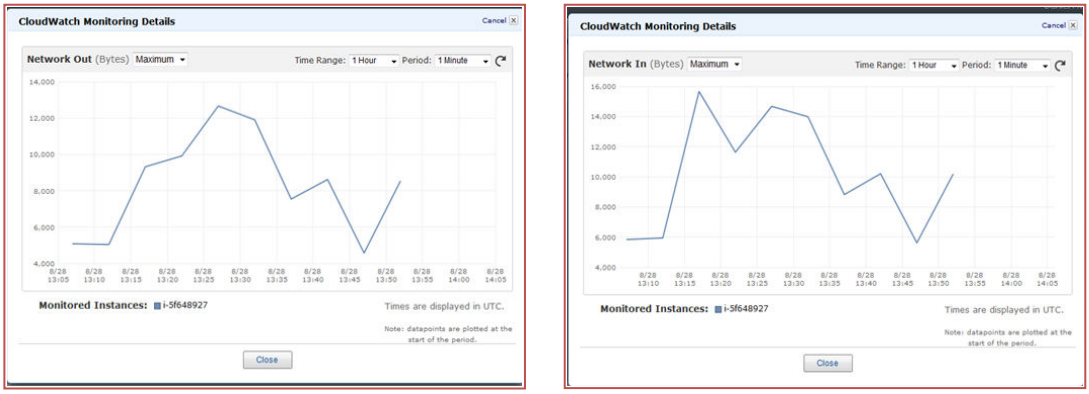

Figure 9. Graphs of cloud-system performance in requests and responses respectively

5.8 User Survey

5.8.1 Survey Design

Due to time limitation factor, a limited group of testers were pre-selected to conduct the beta-test of the Web application: Projects planning staff; Geomatics staff; Research and consultancy staff; Drilling staff and Infrastructure staff. All the potential testers were requested to access the website, explore the tool and then later respond to the survey feedback questions. The purpose of the survey was to evaluate the effectiveness of the website application and the potential for future use of such a concept as a tool and to facilitate data dissemination within KenGen. Interview questions were developed from the literature review, in general and specifically, to gather data on the users' usability of the interface and their thoughts on the online application as an effective dissemination tool.

Out of the people who were invited to participate, there were twenty respondents to the survey that included: seven from Geomatics, four from Infrastructure, three from Drilling, three from Research and Consultancy and three from Projects Planning.

The quantitative methods were analyzed in a descriptive fashion as the sample size was too small to come to any conclusions. The qualitative data was analyzed by reviewing the participant responses to the questions that requested comments from the users.

5.8.2 Quantitative Survey Results

There was overwhelming confidence in the application's value in assisting in accessing maps with 100% feeling the tool is either excellent or good for this purpose. 90% of the respondents felt that the tool both enables time saving in access of information

and also adds value in decision making with the Geothermal Resource Development department.

6. Conclusion and Recommendation

This study is based on a case-study implementation of an application prototype developed to analyze the performance of a Web-based GIS in a virtual server environment. The application was built with open-source tools and implemented on an Amazon virtual cloud server environment.

Cloud GIS is envisaged to provide opportunities for KenGen to become more cost-effective, productive, easier and flexible in order to rapidly deliver new capabilities especially in geothermal resource exploitation. Because Cloud computing offers a scalable virtual infrastructure to users and developers, it gives the illusion of an unlimited resource for computation and data storage. It will allow KenGen users to start small and increase computation or storage resources only when they need, it will also provide users with access to large amounts of "hardware resources" in a short time interval without requiring the users' fixed investments or maintenance costs for expensive hardware.

KenGen's GIS software developers do not need to worry about the limitations or the hardware specifications of the 'computers' they are working on because these 'computers' have a one interface and specification provided by Amazon.

It is envisaged that the technology will save KenGen money because the application uses hardware resources more efficiently and effectively: The cloud GIS allows different users to use resources of the cloud infrastructure at different times, which leads to less system idle time, i.e. less waste of resources; cloud computing providers will be able to charge users for the resources they used.

Users in the different departments can query, edit and manipulate large volumes of data stored in servers hosted by different parties without buying expensive professional GIS software or facing a steep learning curve to understand how to use it. By avoiding the downloading, storing and pre-processing duplicated data, efficiencies can be obtained in the sharing of data over the cloud.

6.1 Implementation of the System and Government Legal Position

There is no explicit existing legislation pertaining to cloud computing in Kenya. This could be attributed to the fact that cloud computing is a relatively new technology in Kenya and in Africa as a whole. This has led to difficulty in openly embracing technology for organizations due to lack of a legal framework to protect them against potential conflicts over data and information shared on the cloud. Computer hardware tasked with storing personal data on customers in a cloud environment can be

distributed across many countries. Cloud computing by nature is nebulous. When data is stored in a public cloud it is difficult to establish where that data is stored, it is impossible to know who has access to the data, from a security perspective and finally it is difficult to ascertain what security mechanisms are in place to keep the data safe (Harris, 2012).

The current Amazon licensing includes a clause that absolves them from blame in case of any data loss or leaks. "*The service offerings are provided 'as is.' We make no representations or warranties of any kind... regarding the service offerings or the third-party content, including any warranty that the service offerings or third-party content will be uninterrupted, error free or free of harmful components, or that any content, including your content or the third party content, will be secure or not otherwise lost or damaged.*" This has not helped in the implementation process.

Due to sensitivity of information in KenGen's domain coupled with the uncertainties described above, it had been difficult implementing the proposed system. Before approval can be granted for implementation, it is a company requirement that the measures taken to secure data are clearly outlined and guaranteed to be breach proof. This in a public-based cloud, is impossible at the moment, and thus the proposal still remains at the proposal stage. It is hoped that an evaluation of improved security using a private cloud can be evaluated in the future, but this is anticipated to drastically change costs.

6.2 Implications of Findings

This step-by-step process prepared for tool development could be used by anyone planning to develop an online GIS application based on the cloud for participation from different tools and technologies and for the specific requirements dictated by the context. This study also gives a descriptive overview on how to set up an interactive online GIS to facilitate data management using open source products and specifically with GeoServer.

6.2.1 Online Application as a Data Management Tool

With the application, it is easier to manage the data in the geothermal resource development section since all of it is contained in a central location. The login requirements ensure that only authorized access and manipulation is permitted ensuring the safety of the data.

6.2.2 Online Application as Data Dissemination Tool

The application ensures that an interactive map is always available for users from all departments within the organization. This makes it easy for them to access updated information when they need it. The users can even access the information at the comfort of their homes if they need to work at home.

6.2.3 Online Application as a Decision Making Tool

Due to the importance of the tool in disseminating the information quickly and efficiently, it ensures that decisions within the organization are made in a timely manner. Wells can be sited and marked on the interactive map with terrain being evaluated much more conveniently.

6.2.4 Online Application as Cost Saving Tool

Though the savings envisaged by the application could not be fully realized because of the small size of the virtual server. Currently about seven licenses of ArcGIS are maintained in the GIS department within the organization. This amounts to about $5,000 every year. With this kind of application, the licenses can be much reduced thus saving the company a lot of money. There is also enormous savings made in terms of time taken in accessing information and also in paper resources reduced by using digital maps as opposed to current paper-based maps.

6.3 Recommendations for Future Work

6.3.1 Real-world Setting

The next step would be to test the tool on a current, real-life issue, where the data and content are. The criteria, their sources, and their consequences need to be explained in much more in depth, and hopefully in a more engaging manner. Making the website user friendly would have to be a priority, in order to attract and maintain the attention of stakeholders.

6.3.2 Increased User Interactivity

As suggested by some of the members, another key feature that could be very helpful is: An option to switch between maps while preserving a field of view for easy comparison of the different geothermal suitability sites. This will allow the user to easily transfer from one scenario to the other like it is possible on desktop GIS software applications.

References

Amazon Web Services. Available from http://www.amazon.com/. [Accessed 5 July 2013].

Armbrust, M., Fox, A., *et al*., (2009), Above the Clouds: A Berkeley View of Cloud Computing. Technical Report No. UCB/EECS-2009-28, University of California at Berkley, USA.

Armstrong, M.P., Cowles, M., and Wang, S., (2005), Using a Computational Grid for Geographic Information Analysis. *Professional Geographer*, 57 (3): 365-375.

Bernhardsen, T., (1999), Geographic Information Systems, an Introduction (2nd edition) Wiley, New York, pp. 550-572.

Bhat, M.A., Shah, R.M., and Ahmad, B., (2011), Cloud Computing: A Solution to Geographical Information Systems (GIS). *International Journal on Computer Science and Engineering*, 3(2): 594.

Bunze K., Ager A. and Schrader-Patton C., (2010), Up in the Air: Adventures in Serving Geospatial Data Using Open Source Software and the Cloud, ACM International Conference Proceeding Series, 2010, COM.Geo 2010 – 1st International Conference and Exhibition on Computing for Geospatial Research and Application, Article number: 35, 21-23 June 2010.

Buyya, R., Yeo, C.S. and Venugopal, S., (2008), Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities, Keynote Paper, Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC 2008, IEEE CS Press, Los Alamitos, CA, USA), Sept 25-27, Dalian, China.

Buyya, R., Pandey, S., and Vecchiola, C., (2009), Cloudbus Toolkit for Market-Oriented Cloud Computing, Proceeding of the 1st International Conference on Cloud Computing (CloudCom 2009, Springer, Germany), Beijing, China.

Chappell, D., (2010), GIS in the Cloud, the ESRI example. ESRI Technical report.

Cornel Ofwona (2006), The Status of Geothermal Energy Exploitation in Kenya, United Nations University Journal Volume 4, pp. 56-63.

Cui, D., Wu, Y. and Zhang, Q., (2010), Massive Spatial Data Processing Model based on Cloud Computing Model. In: Proceedings of the Third International Joint Conference on Computational Sciences and Optimization, IEEE Computer Society, Los Alamitos, CA, 28-31 May 2010, Huangshan, Anhui, China, IEEE, 347-350.

Estimating the Cost of a GIS in the Amazon Cloud An Esri White Paper, August 2012.

Geoexplorer (2013), Available from http://blog.opengeo.org/tag/geoexplorer/ [Accessed 20 August 2013].

Gonzalez, H., *et al*., (2010), Google Fusion Tables: Data Management, Integration and Collaboration in the Cloud. In: Proceedings of the 1st ACM Symposium on Cloud Computing, SoCC'10, 6_11 June 2010, UC Berkeley, Berkeley, CA, USA. New York: ACM, 175-180.

Goodchild, M., Huang, Q., Nebert, D., Raskin, R., Xu, Y., Bambacus, M. and Fay, D., (2011), Spatial Cloud Computing: How Geospatial Sciences Could Use and Help to Shape Cloud Computing. *International Journal on Digital Earth,* 4(4). In Press.

Harris, A., (2012), A Masters in Computing Thesis Submitted to the Dublin Institute of Technology.

Li, Z., *et al*., (2010a), An Optimized Framework for OGC Web Service Seamlessly Integration to Support Geospatial Sciences. *International Journal of Geographic Information Science*.

Longley, P., Goodchild, M., Maguire, D., and Rhind, D., (2001), Geographic Information Systems and Science: John Wiley & Sons, Ltd.

Monmonier, M., (1990), Strategies for the Visualization of Geographic Time-series Data. *Cartographica*: *The International Journal for Geographic Information and, Geovisualization*, 27(1): 30-45.

Montgomery, K. and Mundt, C., (2010), A New Paradigm for Integrated Environmental Monitoring. In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application (COM.Geo'10), 21-23 June 2010, New York, NY, ACM, USA.

Nah, F., (2004), A study on Tolerable Waiting Time: How Long are Web Users Willing to Wait?. *Behaviour & Information Technology*, 23(3): 153-163.

National Institute of Standards and Technology, The NIST Definition of Cloud Computing, Special Publication No. 800-145, September 2011.

Nicolae, B., *et al*., (2011), BlobSeer: Next-generation Data Management for Large-Scale Infrastructures*. Journal of Parallel Distributed Computing*, 71: 169-184.

NRC, 2010. Living on an Active Earth: Perspectives on Earthquake Science. Washington DC: The National Academies Press.

Oliveira, D., Baião, F., and Mattoso, M., (2010), Towards Taxonomy for Cloud Computing from an e-Science Perspective, Cloud Computing: Principles, Systems and Applications (to be published), Heidelberg: Springer-Verlag.

Spatial Cloud Computing (SC2) White Paper 2009: A New Paradigm for Geographic Information Services, Presented by: Hugh Williams, SKE Inc. August 2009.

Srinivas, J. and Kanti Kumar, K., (2011), Geoportal-A Spatial Cloud Information Service. *International Journal of Engineering Science and Technology (IJEST)*, 3(11).

Taylor, D.R.F., (1991), GIS in Developing Nations – Special issue, *International Journal of Geographical Information Systems.*

Wong, D.W.S. and Victor Wu, C., (1996), Spatial Metadata and GIS for Decision Support, Proceedings of the 29th Annual Hawaii International Conference on System Sciences - I996 IEEE.

Yang, C. and Raskin, R., (2009), Introduction to Distributed Geographic Information Processing Research. *International Journal of Geographic Information Science*, 23(5): 553-560.

Yang, C., *et al*., (2008), Distributed Geospatial Information Processing: Sharing Earth Science Information to Support Digital Earth. *International Journal of Digital Earth*, 1(3): 259-278.

Zhang, L. and Zhou Q., (2009), CCOA: Cloud Computing Open Architecture, IEEE International Conference on Web Services icws, pp. 607-616.

Zhelu, Y., (2009), A Web-based Geographical Information System Prototype on Portuguese Traditional Food Products – Master's Thesis in MSc. Geospatial Technologies.

Publication Sponsors

Global Spatial Data Infrastructure (GSDI) Association

The GSDI Association is an inclusive organization of organizations, agencies, firms, and individuals from around the world. The purpose of the organization is to promote international cooperation and collaboration in support of local, national and international spatial data infrastructure developments that will allow nations to better address social, economic, and environmental issues of pressing importance.

Centre for SDIs and Land Administration (CSDILA), the University of Melbourne

CSDILA Provides a focus for research in Spatial Data Infrastructures and Land Administration by building on ongoing research relationships and creating new links through extended collaboration both nationally and internationally. The centre undertakes original worldwide research in land policy, land management, land information, spatial data, technical applications and other areas. It also supports spatial data developments and land administration particularly in the Asia-Pacific region.