METADATA CHALLENGES FACED BY PRODUCERS AND USERS OF SPATIAL DATA IN SOUTH AFRICA

by

Judith Alford

Submitted in partial fulfilment of the academic
requirements for the Degree of Master of Environmental and Development,
(Land Information Management),
School of Environmental Sciences, University of KwaZulu-Natal.

Pietermaritzburg, 2009

Disclaimer

The work described in this dissertation was carried out at the Centre of Environment, Agriculture and Development, University of KwaZulu-Natal, Pietermaritzburg, from September 2008 to July 2009, under the supervision of Dr Denis Rugege.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma at any other University. Where use has been made of the work of others, it is duly acknowledged in the text.

Signed:…………………………………..
Judith Alford (Candidate)

Signed:…………………………………..
Dr Denis Rugege (Supervisor)

Signed:…………………………………..
Mr Dorman Chimhamhiwa (Co-Supervisor)

Abstract

A large number of spatial datasets have inconsistent and/or outdated metadata. In certain cases, metadata is entirely absent. Some spatial data producers suggest that metadata creation and maintenance is a time consuming and labour-intensive process. Conversely, users experience difficulties in understanding and accessing spatial datasets, if associated metadata is insufficient or non-existent. Eventually, deficient metadata use may lead to loss of spatial data meaning and cause its very existence to be forgotten. The purpose of the study was to assess the main challenges hindering metadata creation and maintenance on the part of producers and its usage on the part of users in South Africa. The main findings showed that: data was accessed at expected levels via the internet; most data users accepted alternative spatial data media including compact disks and hardcopy; the spatial data industry is generally under financial budget constraints; particularly in the public sector, lack of skilled personnel in spatial metadata management resulted in staff turnover problems; the framework datasets indicated outdated metadata; and different producers used inconsistent metadata standards and a number of organizations were at rudimentary stage of spatial metadata development. In conclusion, spatial data producers should be encouraged to maintain data with complete documentation in a standardized spatial metadata to assure information consistency for users. Raising awareness about spatial metadata benefits may encourage data managers and top leaders to build on metadata priorities. Moreover, strong compliance with the SDI policy necessitates solid cooperation amongst the spatial data community.

TABLE OF CONTENTS

Acknowledgements

I am heartily thankful to my supervisor, Dr Denis Rugege, whose knowledge and constructive feedback enabled me to improve the outcomes of my dissertation. Most of all, I appreciate his invaluable guidance, support and financial assistance throughout my master's degree. Also, special thanks go to Dorman Chimhamhiwa for his time and advice. I acknowledge the CEAD department (particularly LIM programme) for providing me with the equipment and space to complete my thesis. I convey much appreciation to those who took interest in being part of my sample survey, thank you.

I owe my deepest gratitude to my mother, Moitobo Alford, on whose constant encouragement and unconditional love I relied on from a distance. I cannot end without being grateful to God, for granting me the strength to continue when I felt lost and discouraged.

List of Acronyms

ANZLIC	Australian and New Zealand Land Information Council
CEAD	Centre for Environment, Agriculture and Development
CEN	Comité Européen de Normalisation
CSDGM	Content Standard for Digital Geographic Metadata
CSI	Committee for Spatial Information
DLA	Department of Land Affairs
ESRI	Environmental System research Institution
FGDC	Federal Geographic Data Committee
GDI	Geospatial Data Infrastructure
Geo-ICT	Geospatial-Information and Communication Technologies
GIS	Geographic Information System
GISSA	Geo-Information Society for South Africa
GPS	Global Position System
GSDI	Global Spatial Data Infrastructure
ICT	Information and Communication technology
ITC	Information Institute for Geo-Information Science & Earth Observation
ISO	International Organization for Standardization
NGO	Non Governmental Organization
NSDI	National Spatial Data Infrastructure
NSIF	National Spatial Information Framework
PCGIAP	Permanent Committee on Geographic Information for Asia and the Pacific
SABS	South African Bureau of Standards
SASDI	South African Spatial Data Infrastructure
SDDF	Spatial Data Discovery Facility
SDF	Spatial Development Frameworks
SDI	Spatial Data Infrastructure
SMDF	Spatial Metadata Discovery Facility
SPI	Spatial Planning and Information

WYGISC	Wyoming Geographic Information Science Center
UNECA	United Nations Economic Commission for Africa
URL	Uniform Resource Locator

List of Tables and Figures

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

1.1 Background

Over the years, nations have continued to recognize the importance of investing in spatial data. The management and efficient use of spatial information may support economic, social and environmental development improvements for a country (Warnest *et al.,* 2005, Deng and Di, 2008). However, spatial data quality and quantity alone is insufficient to facilitate sound decision-making for development purposes or to assess spatial damages such as natural disasters. Evolution Spatial Data Infrastructures (SDIs) has motivated the ability to discover, organize, analyze, visualize, access and share digital spatial data resources (Harris, 2008). Therefore, enhancing support to governments and private organizations in decision-making. Furthermore, Nebert (2004) adds that, SDIs encompass extensions of services and information and communication technologies (ICT) to facilitate data accessibility. Also, functional SDIs consist of institutional frameworks coordinated and administered from local to national scales through appropriate policies and standards connected to spatial data applications (Nebert, 2004). Moreover, several authors have pointed out that, documentation of spatial data such as metadata are amongst the most important SDI elements (Odongo and Rodrigues, 2007, Pierkot *et al.,* 2006).

The term metadata is most commonly defined as data about data or more simply, information about data. In essence metadata may be understood as a source of description of the content, quality, condition, authorship, and any other characteristics of spatial data (Nogueras-Iso *et al.,* 2004). Waugh (1998) adds to that, metadata describes: what the data is (title, subject, keywords); how to use the data (where to retrieve the data from); and how the data should be managed (relationship with other spatial data). (See section 2.4 for further metadata description).

To improve spatial data sharing and accessibility amongst the data community requires the existence of a metadata culture within an SDI. However, it is imperative to build effective and appropriate set of spatial metadata elements that describe the internal structure and content of spatial databases (Moellering *et al.,* 2006). Examples of metadata description may include information about: data collection methods, accuracy of the sources, projections and scales (Thompson *et al.,* 2003). Limbach *et al.* (2004) note that, metadata documentation enable users with a better understanding of data, thus leading to efficiency in spatial data utilization. Additionally, Duval (2001) suggests that, metadata encourages data preservation as users are able to reuse digital content through easy location of descriptive data. Furthermore, metadata may stimulate cooperation to spatially related organizations (Campbell, 2008) because metadata provides information to other sources associated with data sharing such as clearinghouses (Pierkot *et al.,* 2006).

To foster effective data documentation, organizations should use similar metadata standards (Pierkot *et al.,* 2006). Such standards facilitate the sharing of spatial data (FGDC, 2008), thus allowing for a greater data acceptance by the spatial data community. Similarly, Thompson *et al.* (2003) maintain that, participation in metadata standards may increase interoperability of datasets for ease of exchange. Also, metadata standards lead to consistent data structure (Thompson *et al.,* 2003) by providing common terminology and definitions for data element documentations. In other words, the standards provide adequate information for commonly designed metadata. To encourage spatial data sharing with different user communities across global network catalogue servers, Nogueras-Iso *et al.* (2004) insist data contents should be adjusted to international metadata standards. Some internationally known metadata standards extend to: ISO standard TC/211 and Federal Geographic Data Committee's content standard for Digital Geospatial Metadata (FGDC, CSDGM), which have become widely adopted by some developing countries like South Africa (Nogueras-Iso *et al.,* 2004).

Ezigkbalike *et al.* (2000) reported from their observations that developing countries, particularly in Africa, possess an extent of spatial information in hard copy formats. The authors predicted that, with the use of computers in spatial data management, efforts of hard copy map conversion into digital format will expand gradually. However, Ezigkbalike *et al.* (2000) claimed slow growth of computer technology partially impeded metadata development in most African countries, thus limiting efficient spatial data sharing and accessibility.

In contrast, some cases in developed countries are ahead in metadata developments because of technological advancements. In such efficient data sharing environments, more concern is based on meeting spatial information demands because spatial data users continue to increase. Therefore, Batcheller (2008) designed an automated metadata generation system through ESRI ArcGIS 9.1 customization with considerable data preparation. The aim was to reduce efforts of manual metadata production. Batcheller (2008) acknowledged that, existing voluminous data documentation tends to be tedious and error prone when manually produced, thus the need for automated metadata generation. Similarly, Parekh *et al.* (2004) observed that, much of spatial data is widely accessible to researchers and other end-users. Numerous spatial data providers produce and distribute spatial data in different formats, which result in interoperability and data discovery problems (Parekh *et al.,* 2004). Therefore, Parekh *et al.* (2004) proposed a semantic metadata management system based on ontologies which enables the basis for spatial data interoperability.

According to Wayne (2005a), organizations are slowly realizing the value of metadata. However, it may take a while before organizations may consider implementing advanced metadata systems. Furthermore, manual metadata creation is still most utilized globally. Norheim *et al.* (2000) acknowledged that metadata generation was amongst the least of company priorities and an area of low funding preference. As a result, Norheim *et al.* (2000) conducted research on organizations less aware of metadata benefits. A metadata expert team was sent to educate the spatial data community in the Peninsula region about metadata software, metadata standards and metadata benefits promotion. One of the project objectives was to enable metadata flow amongst the organizations. This exercise enlightened the metadata experts on the importance of high level creativity involved in promoting metadata development amongst organizations. Subsequently, financial resource availability successfully sustained future monitoring on metadata developments and the building of a clearinghouse for the Peninsula spatial data community.

Even though there have been considerable research investigations on various metadata elements in other countries, not much publication is available on South African metadata issues. Since the establishment of the National Spatial Information Framework (NISF) (South Africa SDI initiative) in 1997 (Gavin, 2001 and, Thabethe, 2008b), much concern has been drawn on issues hampering further SDI developments. For instance, Thabethe (2008b) stated that, South Africa currently faces policy differences across government departments, thus creating a barrier to efficient flow of data sharing. Additionally, some municipalities lack adequate spatial data to prepare appropriate spatial development frameworks (SDF). Other municipalities have the spatial data but lack technical resources and expertise, therefore hindering them from SDF plans. Furthermore, Thabethe (2008b) reported that, lack of funding and uncertain institutional arrangements are some of the milestones that hinder effective collection, monitoring and management of datasets.

The NSIF also runs an internet based national clearinghouse known as the Spatial Metadata Discovery Facility (SMDF). Furthermore, the SMDF functionalities extend to a metadata database that facilitates searchable spatial data holdings through metadata documentation (Gavin, 2001 and, Thabethe, 2008b). Also, under the NSIF initiative, the South African SDI Act addresses organs of state to capture, publish and maintain metadata for their spatial data (UNECA, 2004a). Further, data custodians are obliged to update their spatial information and provide electronic copies of the updated records within 30 days (UNECA, 2004a). Nonetheless, the degree of compliance with the above obligations regarding all state organs remains questionable.

1.2 Problem Statement

Spatial data has become extensively involved in impact analysis of socio-economic and environmental developments (Deng and Di, 2008). A country with a lack of orderly spatial data collection negatively impacts on decision-making and development. According to UNECA (2007), most national development priorities are directed to social, economic and environmental issues. Therefore, governments in most parts of the world have regarded spatial data as an asset that requires management for national interest (Masser, 2005). In particular, Gavin (2001) emphasizes that South Africa is one of the few African countries to invest resources in data asset management such as SDIs. Thabethe (2008b) notes that organizations involved in spatial information have made investments in spatial data collection. However, spatial data is an expensive resource to create and maintain (Thabethe, 2008b).

Tulloch and Fuld (2001) assert that metadata, clearinghouses and standards are some of the main SDI concepts considered important to facilitate accessibility, sharing and reuse of spatial data. The metadata concept is particularly important in providing information about the content of datasets. As a result, this may assist the user to understand the data better and evaluate the fitness of use of the documented dataset (Duff *et al.,* 2003). By ensuring the availability of metadata through clearinghouses, various users can discover data to access and use. Moreover, organizations utilizing the clearinghouse could enable them to share data collection activities thereby reducing duplication of efforts (FGDC, 2000).

Over the years, the NSIF has encountered problems with the functionality of the SMDF (clearinghouse) (Osei, n.d). Hence, spatial data users have been unable to appropriately and conveniently link with other spatial data users to share and discover relevant data. On the other hand, even if the SMDF were operational, there is the problem of insufficient and inappropriate metadata available for the clearinghouse. Also, metadata problems impact on efficient spatial data use. The following are examples of problems associated with spatial metadata:

- Metadata records are absent or incomplete for some datasets. In such cases, if the contents or structure of the acquired data is difficult to understand, the user could be limited in effective use of the data. For instance, the metadata may contain missing elements such as: spatial reference information, scale, data currency and data originator contact details. If such relevant information like spatial reference is missing, this could delay or prevent further application of the data.
- Outdated metadata is another example of the shortcomings of metadata. Metadata should be as current as the data. In other words, when data is created or edited, its metadata should immediately be created or updated to reflect data changes. However, according to ESRI (2008), creating and updating metadata requires a substantial quantity of work and time. For this reason, data holdings are largely left unchecked

for their appropriate age and structure to verify which data should be maintained, updated or deleted (ESRI, 2008). Therefore, institutional spatial data memory could be lost through inappropriate storage of metadata records. Furthermore, outdated metadata could misinform and confuse users about the data.

- Inconsistency of metadata standards prevails among the spatial data producers and custodians. Organizations involved in spatial data holdings continue to use metadata standards appropriate to their internal uses. To sustain a large-scale clearinghouse for data sharing with diverse spatial users, it is essential to maintain a common metadata standard. The reason being, the standard ensures that the spatial data community has the same understanding of the terminology used. Furthermore, non-standardized metadata could impede on the interoperability within the SDI.

It is significant to investigate the causes and effects of metadata problems to determine an approach that may reduce impacts on efficient spatial data use. Figure 1.1 identifies some of the assumed causes and effects of metadata problems in the South African context. This is illustrated in a problem tree (figure 1.1).

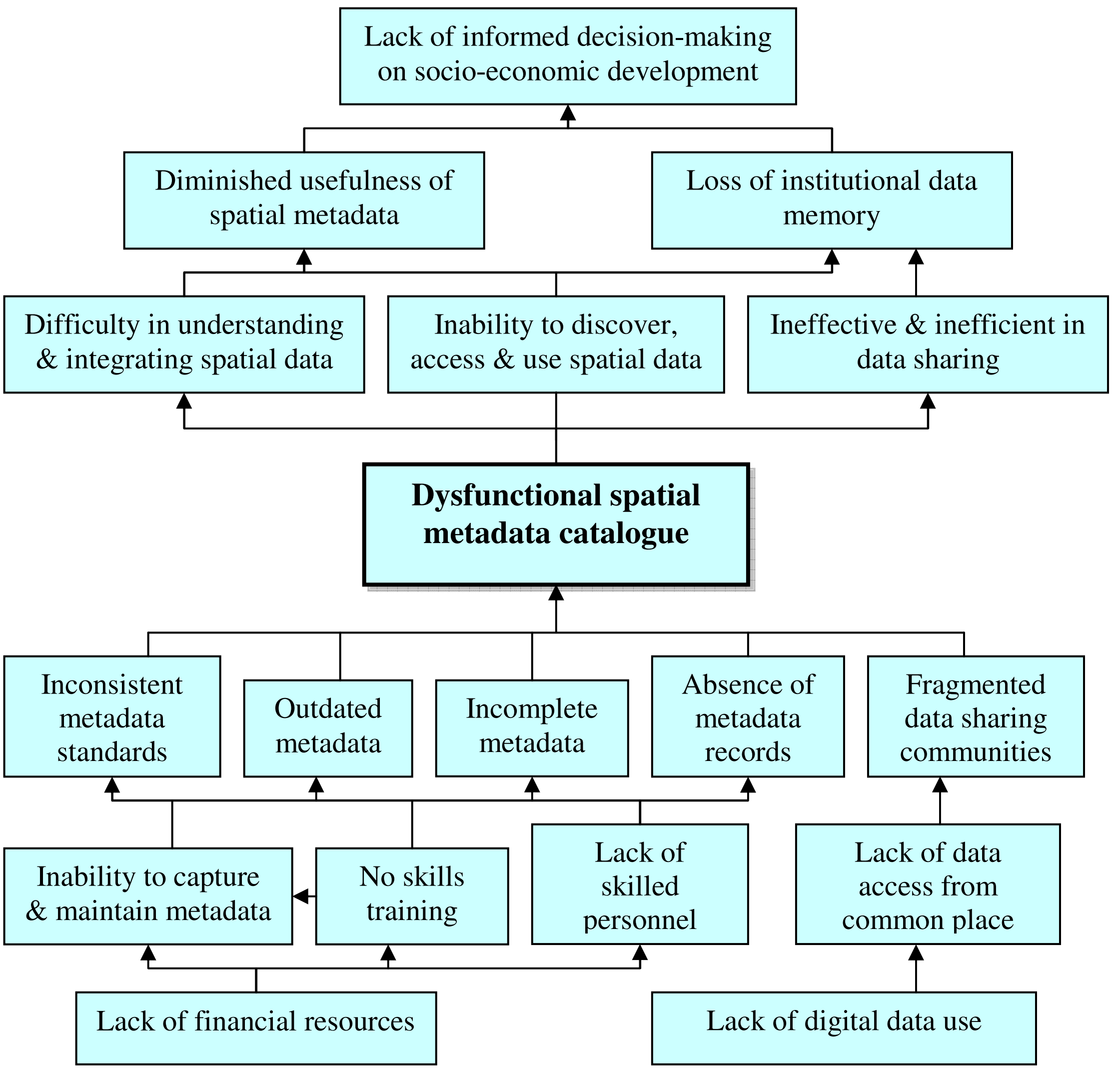

Figure 1.1: Problem tree for spatial metadata in South Africa

In figure 1.1, lack of skilled personnel, inabilities to create and maintain metadata and less skills training are assumed to stem from lack of financial resources in the country. Organizations with lack of financial resources could impede insufficient employment of skilled metadata personnel. Less metadata skilled personnel can also affect inefficient metadata creation and maintenance. Lack of metadata creation could result in the absence of metadata documentation. Also, lack of metadata maintenance could result in outdated and incomplete metadata. Furthermore, lack of skilled personnel and metadata maintenance may lead to the use of inconsistent metadata standards. In another case, hard copy exchange seems to be more common than digital data distribution. For that reason, there is lack of common place to access the digital data. Hence, fragmented data sharing communities start to develop limiting the sharing of data to a wider community. An accumulation of various metadata problems could result to a centered problem of a dysfunctional spatial metadata catalogue.

Spatial data users may find it difficult to understand and integrate data with metadata that is either: absent, incomplete, outdated and/or non-standardized. Further effects may be the inability of users to discover, access and use spatial data. Consequently, it might become difficult for spatial data producers, custodians and users to share their data with others in the industry. This could cause duplication of data collection. The difficulty to understand, integrate and access data in an organization may cause loss of institutional data memory. Consequently, diminished usefulness of spatial metadata may arise and could negatively impact on informed decision-making.

1.3 Research Aim and Objectives

1.3.1 Aim

The aim is to assess the metadata challenges faced by producers and users of spatial data.

1.3.2 Objectives

- To find out experiences on metadata use by the spatial data producers and users.
- To assess the level of compliance with SDI policies and standards on metadata.
- To establish reasons for metadata problems.

1.4 Research questions

- Who are the various users and how do they acquire data?
- Who is responsible for metadata maintenance?
- What issues are associated with metadata updates and or maintenance?
- What inconsistencies exist within metadata standards?
- What are the constraints to efficient metadata use?
- How can metadata be improved?

1.5 Study Design

The intention of this study was to explore metadata challenges within the SASDI. Importantly, the causes and outcomes of metadata problems were based on the experience of spatial data producers and users. A sample of the study population was selected from the spatial data community of South Africa. The primary data was collected through questionnaires emailed to prospective respondents. The secondary data was collected by comparing some of the existing metadata for South African framework datasets. Both data collection methods contributed towards the exploration of causes and effects of metadata problems.

1.6 Thesis Structure

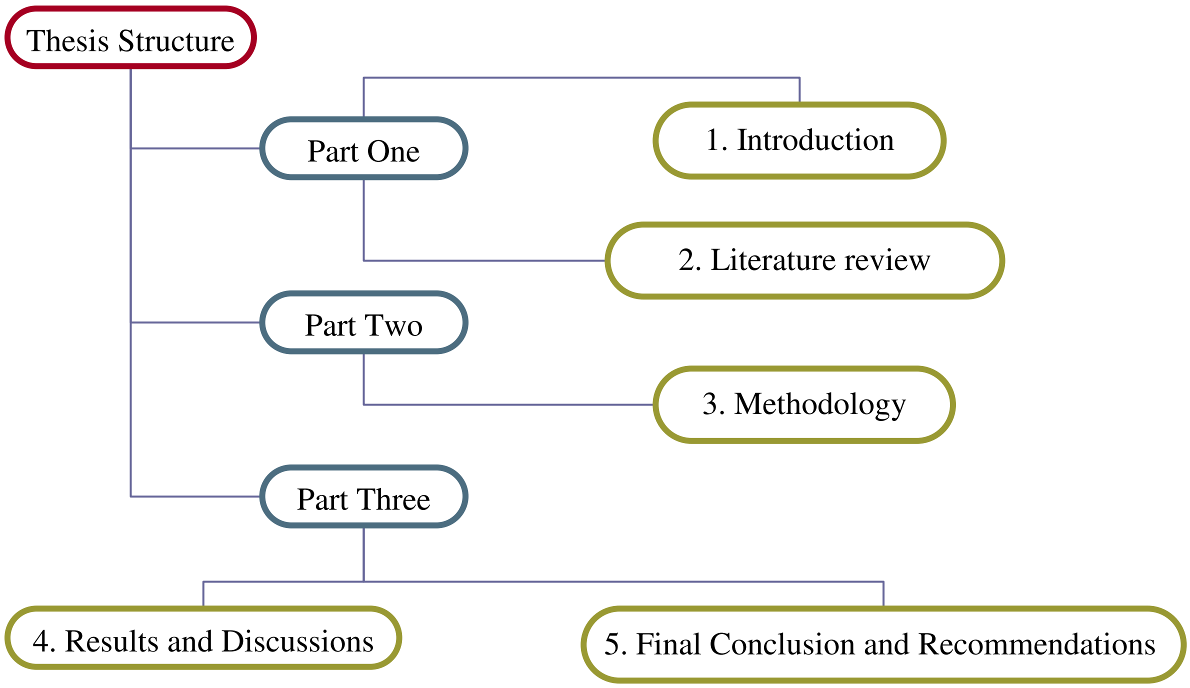

Figure 1.2: Structure of thesis

The figure above describes the structure of the thesis. It is divided into three parts. Part one consists of chapter one presenting the theoretical background, previous findings in the area of study, the research problem, objectives and research questions. Chapter two reviews literature relevant to SDI and related components. Furthermore, in particular, attention stems to the metadata component elaborating on concerned issues. Part two constitutes the methodology chapter. This section firstly explains the conceptual framework structure and secondly expands on the data collection and data analysis methods. The third part represents chapter four and five. Respectively, findings of the results will be described and discussed, and the overall conclusion will mainly result from chapter four findings. Also, recommendations will be presented.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Obtaining information about an existing dataset is of great benefit to spatial data users. For instance, with a detailed data background, a user may avoid further enquiries about the dataset; also, dataset producers and custodians save time by explaining the missing information. Detailed documentation about spatial data is maintained as metadata, describing the content and quality of the dataset. Spatial information has increasingly become important for successful public task performance, such as economic and social development (Kok and van Loenen, 2005). Through data sharing, spatial data from various disciplines could be integrated, providing easy access to city planners and stakeholder developers. Metadata has become a key element to build an effective Spatial Data Infrastructure (SDI) (Limbach *et al.,* 2004). Well-documented metadata may facilitate a better understanding of spatial data, thus leading to better decision-making.

This chapter will:

- Examine the concept of SDI.
- Identify and explain the different SDI components. The focus of this review is a discussion on metadata.
- Provide an overview of South African Spatial Data Infrastructure (SASDI).

2.2 Spatial Data Infrastructure Concepts

There has been a gradual shift from a single-purpose perspective to a multi-purpose perspective for organizations administering land (Thellufsen *et al.,* 2008). Thus, the goal for such multi-purpose land administration systems has evolved from focusing on data handling in isolation, to spatial information available to other government organizations. Thellufsen *et al*. (2008) describe this goal as changing from “silo thinking” to service-oriented, modern organizations delivering land information to a wider society. Such multi-purpose perspectives have evolved into the building of Spatial Data Infrastructures.

The Spatial Data Infrastructure, also termed Geospatial Data Infrastructure (GDI), has been variously defined. The variety of SDI descriptions results from the many different purposes for its use (Rajabifard *et al.,* 2003), its multifaceted character (De Man, 2007), and because SDI is understood differently by stakeholders from different backgrounds. For example, SDI can be defined as:

- "the facilitation and coordination of the exchange and sharing of spatial data between stakeholders in the spatial data community" (Crompvoets *et al.,* 2004);
- "an initiative intended to create an environment in which all stakeholders can cooperate with each other and interact with technology" (Williamson, 2003, p.3);
- "a term to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data" (Nebert, 2004, p.8); and
- "the totality of technology, policies, standards, human resources, and related activities necessary to acquire, process, distribute, use, maintain, and preserve spatial data throughout all levels of government, the private and non-profit sectors, and academia" according to the US Federal Geographic Data Committee (FGDC) definition (Maguire and Longley, 2005, p.5).

Despite the non-existence of a universal standard definition of an SDI, the above definitions depict an approximation of different SDI aspects. In other words, it is the dynamic, functional SDI components that aim to facilitate and coordinate exchange, sharing, accessibility and spatial data use (De Man, 2007).

To understand an SDI, its components are usually defined. Thus, SDI components set the underlying framework of the system "infrastructure" (Rajabifard *et al*., 2003). Coleman and McLaughlin (1998) as cited by Rajabifard *et al.* (2003), explain infrastructure as (i) people, (ii) laws and (iii) the education to use systems. For example, in a highway infrastructure, a motorist tends to think of the bridges and the highway condition. As mentioned by Coleman and McLaughlin (1998), cited in Rajabifard *et al.* (2003), however, the infrastructure also includes the highway laws, drivers' licenses, petrol stations, the people cutting the grass along the highways and all other supporting systems.

Therefore, SDIs are like other forms of infrastructure, such as roads and power lines (Ezigbalike *et al.,* 2000) because they are made up of several components.

2.3 SDI Components

A Spatial Data Infrastructure, whether at a local, regional, national or global level, comprises at least five constituent components (Figure 2.1):

- Spatial data;
- People;
- An institutional framework;
- Technology; and
- Standards (ANZLIC, 2006; FGDC, 2006; PCGIAP, 2003 cited in (Dushimyimana Simbizi, 2008).

These constituent components of an SDI are a key prerequisite for the effective collection, management, access and use of spatial data at different administrative levels (Nedovic-Budica *et al.,* 2004). Furthermore, SDI components and their segments are strongly related and have dynamic inter-relationships influencing the development of an SDI framework (Warnest *et al.,* 2005).

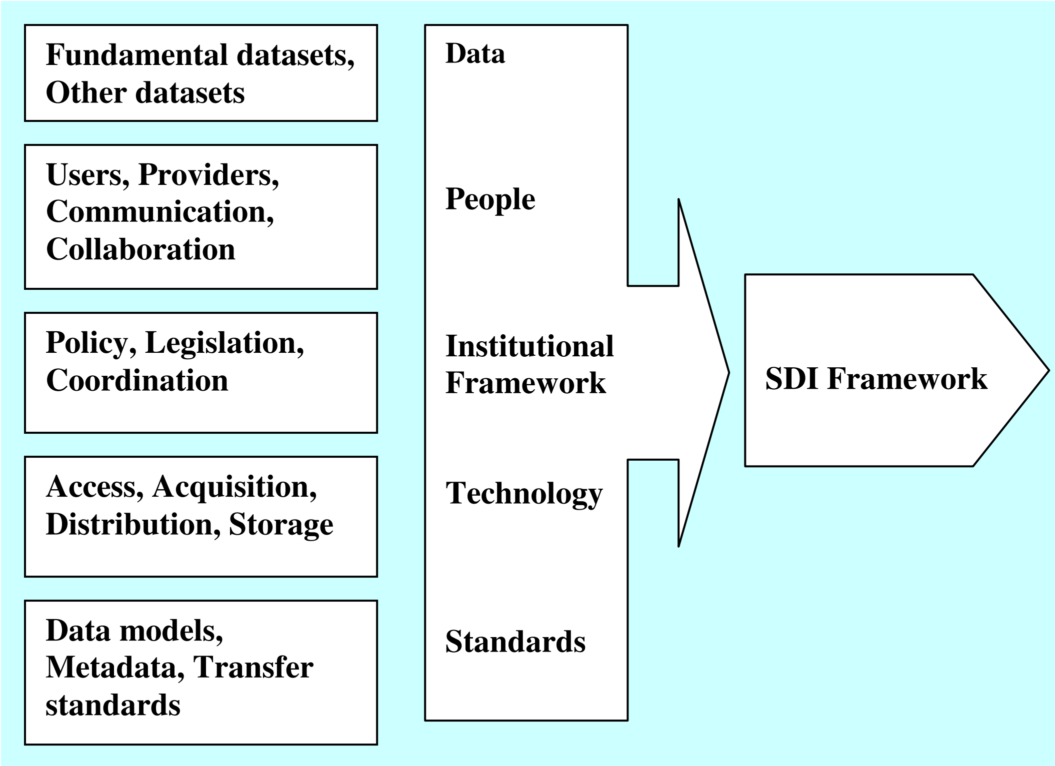

Figure 2.1: Core components of an SDI Framework (Adapted from Warnest *et al.,* 2005, p.4)

2.3.1 Spatial Data

Large amounts of spatial data have been collected over the years. Latitude and longitude or a national co-ordinate system reference such data (Williamson *et al.,* 2003). Bishop *et al*. (2000) suggested that to build an SDI, the basic step starts with collecting spatial data to avoid redundancy. In response, Williamson *et al*. (2003) suggested the sharing of existing data through the facilitation of SDI; the use of spatial data is an integral part of SDI development (Crompvoets *et al.,* 2004). Some examples of spatial data are topographical, geographic features, place names, height data, land cover and hydrography, property boundary information, administrative boundaries and environmental and natural resources (Rajabifard *et al*., 2003, p.17). However, spatial data not in digital format is incompatible for utilisation in the SDI context. It is in digital format that datasets can be stored, managed and analysed by spatial data software (Rajabifard *et al*., 2003).

Spatial data exists in a wide variety of data formats. Vector and raster formats are the most common, but can be complex because of the various ways coordinates, attributes and display information are stored. Generally, information often included in digital maps contains:

- Geographic information providing the position and shapes of spatial features;
- Attribute information providing additional non-graphic information about each feature; and
- Display information describing how the feature appears on screen (GeoCommunity, 2007).

It has been widely estimated that about 80 percent of all governmental data is in the form of spatial data (Grant and Williamson, 2003, Rajabifard *et al.,* 2003). Gyamfi-Aidoo *et al.* (2007), however, suggest the type of spatial data required for proper decision-making and development-planning requires a systematic approach to the collection and maintenance of the dataset. An important example is framework data: this is represented in a consistent manner and referenced to a common georeference allowing data layers to be integrated with other datasets (Holland and Borrero, 2003).

Even though there is no universally accepted definition of framework datasets (fundamental or core datasets), UNECA (2007, p.1) recommends the following definition be adopted, particularly for the African context:

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

Framework datasets form part of a national asset essential for country developments. The range of framework datasets may differ from country to country because of various interests and data collection priorities (Rajabifard and Williamson, 2001). A comparison of framework datasets for various countries and programmes is summarized in Table 2.1.

Table 2.1: Comparison of framework datasets for various countries and programmes (Adapted from UNECA, 2007, p.29)

Datasets	**South Africa**	**Botswana**	**Namibia**	**Nigeria**	**Mexico**	**Western Australia**	**US**	**Global map**
Transportation	♦	♦	♦	♦		♦	♦	♦
Administrative boundaries	♦	♦	♦	♦	♦	♦	♦	♦
Hydrography	♦	♦	♦	♦	♦		♦	♦
Settlements		♦						♦
Topography		♦		♦	♦	♦		
Elevation	♦	♦		♦	♦		♦	♦
Vegetation								♦
Land cover	♦		♦	♦				♦
Land use		♦	♦	♦				♦
Geodetic control		♦		♦	♦	♦	♦	
Cadastre and Tenure	♦	♦		♦	♦	♦	♦	
Imagery	♦	♦		♦	♦	♦	♦	
Place names					♦	♦		
Geology			♦	♦				
Demography				♦				
Coastlines					♦	♦		
Property street address						♦		
Freehold tenure						♦		
Electoral boundaries						♦		
Baseline/Territorial Sea-lanes						♦		
Utility networks	♦							

To achieve appropriate framework datasets collection, the relevant data should be easily accessible to users for further spatial data discoveries (UNECA, 2007). It is important that the spatial data community should be in agreement on common terminology and descriptors for every piece of a dataset (UNECA, 2007). Already, some common elements have been agreed upon by data producers, custodians and users in the sub-regions of Africa. Thus, when capturing metadata for fundamental datasets, the following compulsory fields should be included:

- Originator of the data set;
- Publication date;
- Title of the data set;
- Format of the dataset;
- Description of the data set;
- Purpose of the data set;
- Date of completion;
- Status of data set (e.g. completeness);
- Contact details of custodian;
- Accuracy of attributes;
- Accuracy of spatial data;
- Scale of maps;
- Projection/coordinate system;
- Datum;
- Ellipsoid;
- Access constraints;
- Use constraints;
- Distribution information; and
- Spatial boundary extent (UNECA, 2007).

2.3.2 People

In general, people appear as the key to transaction processing and decision-making (Feeney, 2003, Rajabifard *et al.,* 2003, Williamson, 2002). The people involved in establishing an SDI are: spatial data producers, custodians, users, and value-adding agents (Chan *et al.,* 2001).

Data producers are the original developers of datasets according to their various business requirements and demands (Nebert, 2004). Data producers may distribute the data to the public or hand over the responsibility to data custodians.

Custodians are important in handling spatial data (Thompson *et al.,* 2003). They have the responsibility to ensure that spatial datasets are collected and managed on behalf of the wider spatial data community (CSI, 2004). Though custodians are the main information trustees, they still need interaction with users to collect and maintain useful datasets for planning and developments purposes.

Williamson (2003) has observed that users are giving more input on spatial data requirements and access than in the past, as more forms of spatial data use are being explored. For example, some users have become more interested in improving the response to emergencies such as bushfire threats and other natural disasters (Williamson, 2003). Improved knowledge on various spatial data uses could guide spatial data users to demand better data content, quality and accessibility. Williamson (2003) predicts that in the future, users will be the main drivers of effective SDI development.

Value-adding agents interact between custodians and users. According to Groot and McLaughlin (2000), cited in Dushimyimana Simbizi (2008), value-adding agents make use of framework data and other datasets to prepare and supply application spatial data to end-users.

SDIs are recognised as potentially facilitating the relationship between people. However, Feeney (2003) emphasises that it can only be achieved if human capital is developed. This may be achieved through training and skills for the improvement of spatial data access and use (Feeney, 2003).

2.3.3 Institutional Frameworks

An institutional framework refers to coordination and collaboration of organizations involved in building and maintaining a national SDI (Ezigbalike *et al*., 2000). Williamson

(2002) emphasises that SDI development is a matter of cooperation and partnerships between all stakeholders. In addition, Clarke *et al*. (2003) state that an institutional framework defines the policy and legislations maintenance for standard applications and common spatial dataset procedures.

The increase of the spatial data community has resulted in diverse spatial data uses that require consensus building through a participatory approach to achieve an efficient SDI (UNECA, 2004d). Proper government and private sector partnership may essentially form a range of resources within an SDI framework to support all partners on spatial data use (UNECA, 2004e). To remain coordinated and focussed, therefore, a country should establish an institutional arrangement. Table 2.2 shows an example of a general organisational structure comprising the following:

- Ministry in charge;
- Lead agency;
- Forum of data producers and users;
- Steering committee; and
- Technical working group.

Table 2.2: Summary of a general organizational structure (Source: UNECA, 2004d, p.2)

Organs in charge	Roles
Ministry in charge	Office of the president in charge of the main spatial planning sectors (e.g. environment, Mining, Defence). Ensures political, administrative and financial advocacy.
Lead agency	An institution acting in a secretariat role for the stakeholders below. Facilitates administratively the use of SDI by managing its resources and integration with other SDI initiatives.
Forum of producers and users	An instrument for participation of the spatial data community in SDI operation. Also, a decision-making body regarding SDI development and management. However, decisions must first be approved by government. Can act as a pressure group to move SDI developments forward.
Steering committee	It represents a sample of the stakeholder community. They analyse the outcomes of activities taken by the technical

	working group and make recommendations to the forum.
Technical working group	The working group are expertise that focuses on particular SDI problems. They facilitate in suggesting solutions in areas such as drafting standards, policies and capacity building programmes.

Institutions or organisations may experience risks to data sharing. For example, some institutions are reluctant to participate in data exchange in order to avoid revealing poor data quality. Such organisations run the risk of damaging their data ownership reputation, and could consequently be left out of data-sharing arrangements by other organisations. In addition, some institutions fear data ownership loss because other data users might not acknowledge copyrights (UNECA, 2004a).

Nonetheless, in general an SDI aims to provide easy access to data sharing. Therefore, SDI policies should be designed to avoid potential risks such as the examples above. UNECA (2004e) identified the following policies to support SDI development:

- Policies and legislations related to the right to access information, for example, South Africa's Promotion of Access to Information Act, Act 2 of 2000; such rights promote the sharing of data.
- Pricing policies that provide low costs on data that have been collected using public funding. This could protect third parties from high inconsistent data costs that could lead to data-sharing barriers.
- Policies regarding ownership of copyright on spatial data. Such policies could influence the use and reuse of spatial datasets. In South Africa, the Copyright Act, Act 98 of 1978, protects the spatial information owned by the government. All other government organs may use the data without copyright permission. However, any third party (such as private sector clients) must acknowledge state copyright (CSI, 2002).

2.3.4 Technology

Information and communication technical advances, particularly digital technology to capture spatial data (UNECA, 2004e), are essential to SDI evolution (Sanga *et al.,* 2005, Williamson, 2003, Williamson, 2002). Technology is generally assumed to be a catalyst for other SDI component improvements (Ting, 2003). For example, the nature of spatial metadata depends partly on communication technologies to facilitate access and discovery of spatial data for wider usage (Budhathoki and Nedovic-Budic, 2007).

Geographic Information System (GIS) technologies are computer systems capable of assembling, storing, manipulating and displaying spatially referenced data (UNECA, 2004e). GIS technologies have had a positively impact on the spatial information domain (Williamson, 2003). For example, in the past hard copy maps containing errors had to be re-created; however, with the contemporary use of GIS, spatial data can easily be edited, allowing producers capabilities to reuse existing digital data (UNECA, 2004e). Moreover, communication technologies such as the Internet have revolutionized spatial data discovery (Ting, 2003, Williamson, 2003). Nonetheless, SDIs can only function effectively if such communication technologies provide the foundation for interoperability of heterogeneous software and hardware platforms (Kainz, 2000).

Knowledge about which data exists, what the data characteristics are, and under what conditions the data is to be accessed is likely to reduce duplication of effort, and improve efficiency and decision making. The term commonly used to describe an electronic facility to search, view, transfer, advertise and disseminate spatial data from several sources via the Internet is termed a spatial data clearinghouse (Bernard *et al.,* 2005). FGDC defines a clearinghouse (One-Stop-Geoportal) as a software system and an institution that facilitate digital spatial data downloads and discoveries (Kainz, 2000). Usually, a clearinghouse consists of servers on the Internet that report on discovered metadata (Kainz, 2000, Maguire and Longley, 2005).

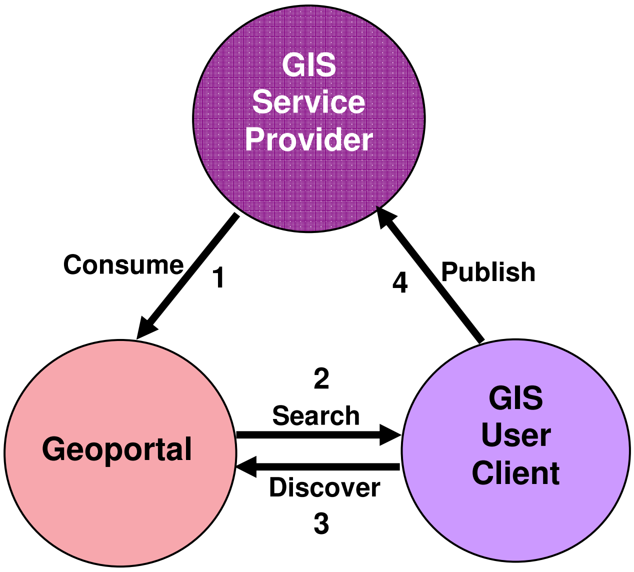

Figure 2.2: The role of a clearinghouse in an SDI (Adapted from Maguire and Longley, 2005 p.8)

Figure 2.2 illustrates how a simplified version of a clearinghouse works. In operation 1, spatial data providers feed the geoportal database with metadata documentation of published spatial data products. In operation 2, with an Internet connection, spatial data users are able to query the metadata against the database from a web client desktop. In return, this enables users to discover the available spatial services and products (operation 3). Accordingly, services can be applied towards spatial data projects, planning, decision-making and other output applications. In operation 4, the output applications could be converted into inputs for new service operations.

However, although the above clearinghouse system reflects a means to an end for spatial data access and dissemination, Nedovic-Budica *et al*. (2004) are concerned about the widespread of these SDI-related network and communication channels. They conclude that spatial data clearinghouses are in the formative stages and scarcely used by the majority of potential spatial data users. In essence, an increased level of awareness among spatial data users could be a way to obtain spatial information such as data type (what), location (where), quality (accuracy, currency, completeness), and ownership (whose).

2.3.5 Standards

Eagleson and Escobar (2003) identify standards as guidelines that enable efficient and up-to-date use of spatial data. Standardized content of information permits resourceful use of data by people and information systems (Campbell, 2008). Moreover, data sharing within an organization and between organizations necessitates the adoption of common standards derived from a clear consensus (Jacoby *et al.,* 2002). However, Crosswell (2000) argues that, adopting standards alone is not a sustainable approach. Therefore, the challenge faced by an institution may entail opting for appropriate standards that can easily be implemented by spatial data stakeholders.

At an international level, the International Organization for Standards (ISO) has focused on designing a wide range of standards that affect SDI (Nebert, 2004, Crosswell, 2000). The GIS committee within ISO, which was formulated in 1993 (Nebert, 2004), comprises various participants from developed and developing countries. The working group addresses standards that can be applied from diverse levels within an SDI *(Crompvoets et al.,* 2004, Crosswell, 2000).

Standards are required for:

- Reference systems;
- Data quality;
- Data models;
- Data transfer; and
- Metadata.

2.4 Metadata

Originally, metadata was first used within the digital library context in the 1960s *(Riall et al.,* 2004). In its traditional context, metadata was described as information used to arrange, file and facilitate access to library or museum resources (Riall *et al*., 2004). In the spatial information field, metadata was only introduced in the 1990s (Vermeij, 2001). Taylor (2004) notes that the major difference between metadata for spatial data and other metadata sets for libraries and elsewhere, is emphasis on the spatial component.

2.4.1 Metadata Description

Metadata is data about data. It describes characteristics of original data regarding the content, quality, accessibility, lineage and other features (Crompvoets and Bregt, 2003, Woldai, 2002, and FGDC, 2000). Additionally, WyGISC (n.d. p.2) describes metadata as a text documentation that attempts to answer the following questions about spatial data:

- **Who** collected the data and **who** distributes the data?
- **What** is the projection of the data?
- **When** was the data collected?
- **Where** was the data collected?
- **Why** was the data collected?
- **How** was the data collected?
- **How** should it be used?
- **How** much does it cost?

Metadata also describes different elements of spatial data, such as information on data identification, data quality, spatial reference and organization, entity and attribute, and distribution (FGDC, 2000). In addition, these elements are provided as categories for data content within the Federal Geographic Data Committee (FGDC) content standard. Figure 2.3 illustrates an example of metadata contents.

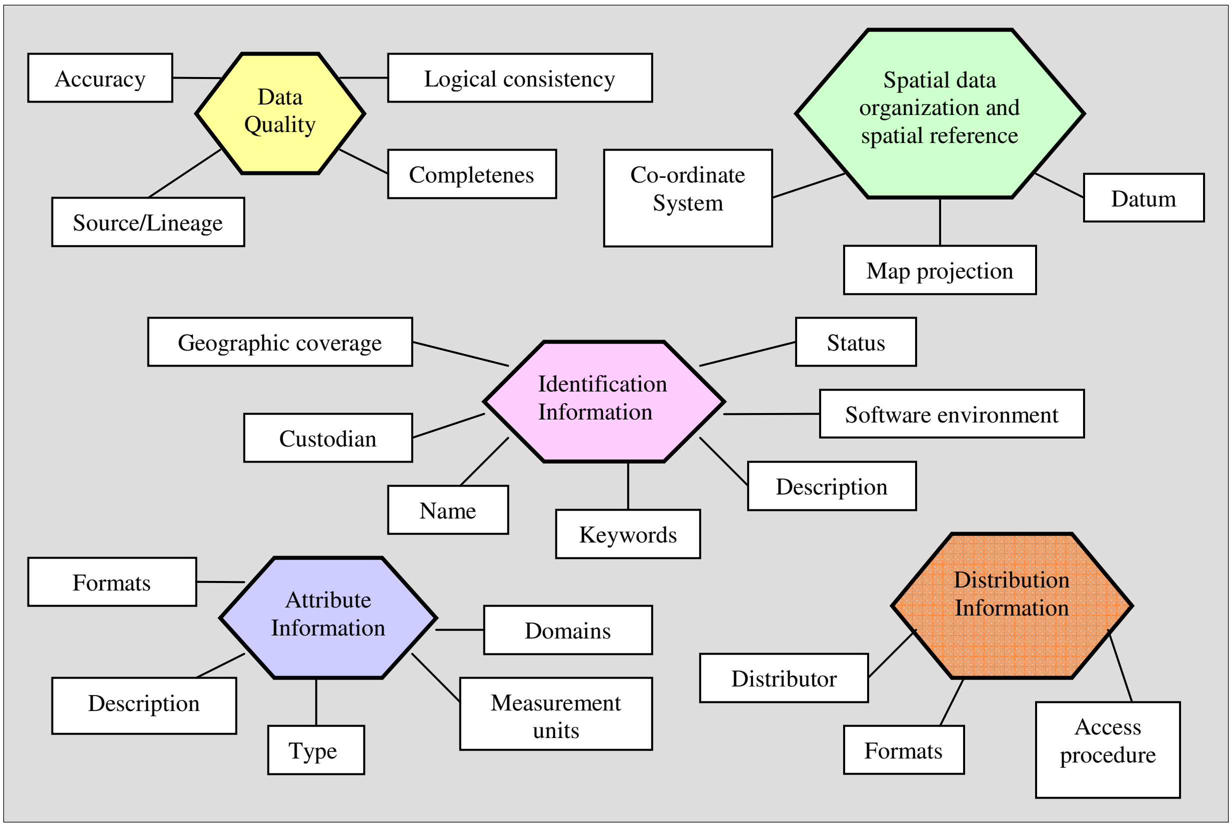

Figure 2.3: Example of the metadata content. (Source: Crosswell, 2000, p.19)

2.4.2 General Metadata Standards

A common understanding of metadata enables various disciplines to integrate the datasets (Bishr and Radwan, 2000). Therefore, standardized metadata could facilitate the spatial data community to interpret data in similar terms. In other words, to guarantee a wide acceptance in the national and international spatial data market, metadata specifications should be linked to standards supported by international systems (Brox *et al.*, 2002). There are three internationally recognised standardization initiatives for spatial data which aim to promote the standards and use of metadata (Taylor, 2004, Woldai, 2002) :

- European Committee for Standardization (CEN)-CEN/TC 287 for European level;
- International Organisation for Standardization (ISO)-ISO/TC 211 applied at a global level; and
- Federal Geographic Data Committee's content standards for Digital Geospatial Metadata (FGDC, CSDGM) of the United States of Geological Survey.

CSDGM and ISO are known as "content standards", which serve the purpose for entering metadata for spatial data (WyGISC, n.d.). Moreover, content standards describe a common set of terminology and definitions for concepts related to digital spatial data (FGDC, 2000). However, there are other metadata standards such as the Dublin core that do not apply to spatial information (Taylor, 2004). Nonetheless, they may serve as valuable references to link non-spatial resources into a spatial framework (Taylor, 2004).

Although metadata standard creation ensures valuable spatial data consistency, some questions have been raised about the complex tools. For instance, Parekh *et al.* (2004) report that the CSDGM is a difficult standard to use, as the standard is very complex with 334 different elements and a further 119 which contain other elements. Consequently, the standard appears to be confusing. Furthermore, Woldai (2002) argues that metadata standardization seems complex because the tasks require geographic feature classifications and keeping track of evolving standards at global scales. Conversely, WyGISC (n.d.) state that metadata standards are like any other multifaceted system which requires time and effort to comprehend: in the end it is a worthwhile investment.

2.4.3 Metadata Importance

Metadata may serve as data workload reductions because the element contents can be applied to locate and retrieve data resources. Such contents include keywords, time period, contacts, data type, and attributes. Therefore, metadata could channel spatial data consumers to decide on how best to use data (Nebert, 2004). Metadata could enable users to acquire information about the data without conducting inquiries to the data originator (Hunter, 2003). Equally, data producers would be faced with fewer attendance to inquiries (Wayne, 2005a).

Metadata may benefit data-producing organizations in many ways. According to Deng (2002), some spatial data professionals view metadata as a way to analyze and organize the underlying data. Similarly, Wayne (2005a) observes that an increasing number of spatial software and analysts rely on metadata to assimilate and display data. In addition, metadata contributes towards data management efforts through data resource organisation, location, maintenance and update (Batcheller, 2008). Thus, this might lead to reduction in data search duplication efforts. As more datasets are collected, storage space is reduced (Wayne, 2005b). Eventually, there could be difficulties to distinguish between data that requires maintenance or dispensation. However, incorporation of metadata processing elements into data documentation may well facilitate identification of data to be updated, archived or removed.

Metadata is also a means of preserving data investment values (Wayne, 2005a). This is particularly beneficial to local government and other organizations experiencing personnel change. For example, as staff members change or time passes, data information may get lost or lose value. As a result, workers, particularly new staff, may experience confusion and difficulty with understanding the existing database content and its uses. In that case, workers may develop sceptical attitudes towards the existing data. Therefore, complete metadata content and accuracy descriptions for data could encourage appropriate data use (FGDC, 2006).

A proper metadata record may also describe what data is not (Wayne, 2005b). In other words, metadata serves as a means of declaring data limitations by indicating use constraint statements like data scale or geographic limitations. From a user perspective, liability statements could indicate and guide appropriate and misappropriate data use. From a producer perspective, dataset limitations included in the metadata could avoid the trouble of responding to queries or lawsuits (Wayne, 2005b).

Often data of one organisation may be useful to another. Therefore, organisations can create easy data exchange on available metadata through clearinghouses. Accordingly, participation could promote the name of the company. As stated by Wayne (2005a), metadata is the basic way of discovering available spatial data resources through the Internet. However, the metadata should comply with international standards to allow for participation in global spatial data clearinghouses.

2.4.4 Metadata Problems

Constraints to metadata practices remain. Batcheller (2008) reports that a large number of the spatial data community complain that metadata creation is a monotonous, time-consuming and labour-intensive process. The spatial data community suggested that this was because spatial dataset documentation is a manual process. This could contribute to the avoidance of metadata creation (Mathys, 2004). Further, Wayne (2005a) indicates that many organisations capture spatial metadata after data development is complete. Hence, the resulting metadata process becomes cumbersome and may lead to inaccurate and incomplete documentation. Furthermore, Stvilia and Gasser (2008) hold the view that inaccurate or missing metadata may result in failed data searches. It is recommended that metadata should be captured during the data development process (Wayne, 2005a).

In some cases, metadata creation and maintenance is insignificant. For example in Jamaica in 2006, the Ministry of Agriculture conducted a questionnaire to detect the status of their metadata implementation guidelines. The results revealed that only 14 out of 34 stakeholders were actively capturing and maintaining metadata (Campbell, 2008). As a result, the Ministry decided to investigate the culture of metadata management among the organisation stakeholders. Interestingly, findings of the project showed some key limitations to metadata production. Such constraints included restricted access to certain spatial data, which reduced the number of metadata documentations and lack of skilled metadata personnel, and a number of organisations had difficulties with creating accurate and up-to-date metadata (Campbell, 2008). Some other major barriers to metadata production were identified in an interview with metadata workshop members executed by Mecklenburg County, North Carolina. The survey group revealed that:

- "Metadata standards were too expensive and difficult to implement";
- "Metadata production requires time and other resources"; and
- "There were few tangible benefits and incentives to produce metadata" (Wayne, 2005a, p.1).

Stvilia and Gasser (2008) suggest that other factors influencing metadata problems include inaccurate, incomplete or inconsistent mapping. For instance, Odongo and Rodrigues (2007) argue that it is difficult to document metadata for non-spatial referenced data with absent coordinates. In an additional example, according to research findings by Odongo and Rodrigues (2007), the Kenyan Ministry of Health has large quantities of data with explicit locations but which are not spatially referenced. Subsequently, metadata may not accurately represent the actual information object, thus affecting metadata quality.

Norheim *et al*. (2000) undertook a research project to encourage a spatial data community to develop metadata. Concerns raised by the potential metadata developers and the research team's responses are given in Table 2.3.

Table 2.3. Concerns raised by potential metadata developers (Norheim *et al.,* 2000)

Metadata developers were concerned that…	The researchers responded by…
Dataset documentation was a waste of time.	Educating them about metadata benefits.
Metadata publications on the National Spatial Data Infrastructure (NSDI) would increase workloads as dataset demands grew.	Enlightening them about metadata creation requirements for local governments.
Software packages for metadata creation were not user-friendly.	Helping the contributors to select software that matched their platform and their budget and was easier to use.
The 119 elements for the FGDC metadata standard appeared complex.	Raising awareness on standard benefits and using good software that could ease metadata standard approachability.
Their sensitive data was too valuable to be shared with other spatial data users.	Explaining that posting metadata on the NSDI did not require free data sharing; data owners had full control over the dataset receiver.

2.4.5 Metadata Incentives

Many organisations resist distributing their data information (Wayne, 2005b). Batcheller (2008) suggests that incentives could be introduced as people consider the implications on metadata negligence. The key is to encourage a change in mentality towards data documentation processes (Batcheller, 2008). Similarly, Vermeij (2001) points out that metadata would no longer be perceived as a necessary evil once people became aware of the benefits of metadata. Furthermore, a country requires a solid infrastructure based on policy, administration arrangements and reliable access points for spatial data. Therefore such strong frameworks could serve as driving forces for wide metadata use (Woldai, 2002).

In the United States, some factors that have heightened the metadata content initiative are the availability of financial resources, metadata knowledge, geographic information skills, and software tools provided by the FGDC (Taylor, 2004). Importantly, governments should recognise and prioritise resource incentive provisions for effective and sustainable use and management of metadata. Metadata development encompasses various organisational aspects that involve funding and material resources. Therefore, a substantial portion of the costs has to be covered by the government. However, in the context of a developing country, the problem is commonly a lack of resource incentives. Publicity mechanisms that will reach data managers and top national leaders with influence on metadata improvements should be boosted (Woldai, 2002).

2.5 South African SDI Overview

As stated by Bishop *et al*. (2000), the underlying SDI practices in developed countries are more advanced and inappropriate to adopt in developing countries. This may be because of differing histories and geography layouts between various countries. Bishop *et al*. (2000) point out that spatial information in developing countries is either poor or non-existent. By contrast, South Africa, as a developing country has proven otherwise despite challenges within the spatial data environment.

2.5.1 The National Spatial Data Framework (NSIF)

NSIF is the main body for the South African SDI. It was initiated by the Department of Land Affairs (DLA) in 1997 (Figure 2.4).

NISF also established a sub-directorate – Standards Development and Implementation – that aims to adopt and adapt standards and policies in the South African spatial data family. The NSIF in conjunction with the South African Bureau of Standards (SABS), are members of ISO. Therefore, South Africa is a participating member of the ISO TC211 (GISSA, 2006).

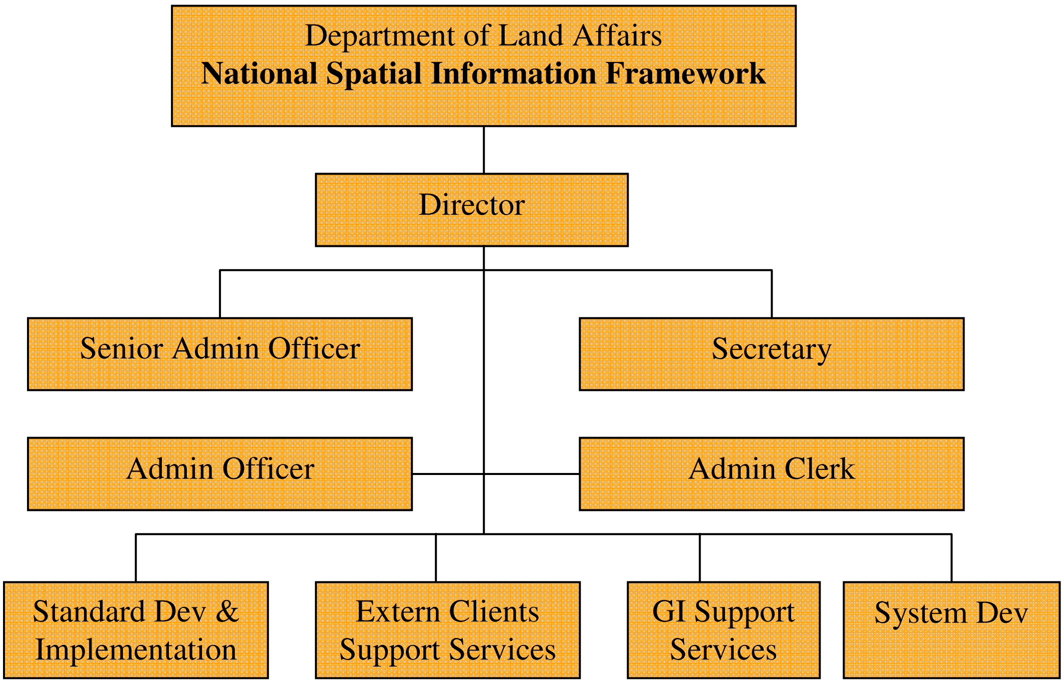

Figure 2.4: The organizational structure of NSIF (Adapted from UNECA, 2004d, p.7)

Another SDI achievement was the publication of the Spatial Data Infrastructure Act, No. 54 of 2003. However, the regulations relevant to the Act are yet to be published. The objectives of the Act include:

- "To establish the South African Spatial Data Infrastructure (SASDI);
- To establish the Committee for Spatial Information (CSI);
- To provide for the capturing and publishing of the metadata as well as develop an electronic metadata catalogue;
- To determine standards and prescriptions with regard to facilitating sharing of spatial data; and
- To avoid duplication of spatial data capturing" (DLA, n.d. p.6).

In general, SASDI aims to manage and supply spatial data, promote sharing and access to spatial data, and to avoid data collection duplication (Thabethe, 2008b, UNECA, 2004b). The CSI aims to facilitate the achievement of SASDI objectives. Thus the CSI plays a key role to provide advice to the Minister, the Director-General and any organ of state dealing with spatial data (GISSA, 2006). The committee members are representatives from all departments of state and provincial government, one rural municipality, one urban municipality, one GIS association, one GIS tertiary education institution, Public Finance Management experts, and State data custodians (UNECA, 2004d). Despite the well prepared structure of the SASDI, in reality government departments currently face financial and resource constraints on dataset collection (Thabethe, 2008b). Thus, this calls for focus on data investment and preservation of existing data.

2.5.2 South Africa Framework Datasets

The SDI Act defines framework datasets as "those themes of information which have been captured or collected by data custodians" (DLA, n.d. p.3). After scheduled workshops, the South African spatial data community identified the framework datasets and their producers/custodians summarized in Table 2.3.

Table 2.4: Framework datasets and their producers/custodians identified by the South African spatial data community (From (Thabethe, 2008b))

Producers	Framework datasets
Chief Surveyor General	**Cadastral Boundaries** ▪ (Urban & Rural- 1:50 000)
Surveys and mapping	**Hydrographic** ▪ Perennial rivers (1: 50 000) ▪ Non-Perennial rivers (1: 50 000) ▪ Dams (1: 50 000) **Services and utility** ▪ Roads (1: 50 000) ▪ Power lines (1: 50 000) ▪ Railways (1:50 000) ▪ Pipelines (1: 50 000) ▪ Airports **Elevation** **Digital Orthophoto Images (1: 10 000)**
CSIR Environmentek	**Land Cover (1: 250 000)**
Department of Water Affairs and	**Hydrographic**

Producers	Framework datasets
Forestry	▪ Catchment areas ▪ Lagoons ▪ Lakes, Vlei ▪ Canals ▪ Perennial & Non-Perennial Pans ▪ Dam areas & walls
Human Science Research Council	**Services and utility** ▪ Post offices ▪ Schools ▪ Hospitals ▪ Clinics ▪ Police Stations
Municipal Demarcation Board	**Administrative Boundaries** ▪ International, provincial, Magisterial Districts ▪ Local Authorities ▪ Enumerating areas
Statistics SA	**Administrative Boundaries** ▪ Voting District ▪ Electoral Wards
IEC	**Administrative Boundaries** ▪ Tribal Authorities
Public Land Inventory	**Administrative Boundaries** ▪ Village Boundaries

According to Thabethe (2008b), in the SASDI context, the framework datasets provide a base to apply other themes and to attach geographic attributes. In addition, the framework datasets may act as reference points to overlay datasets produced elsewhere. Moreover, framework datasets as outputs can also be used as inputs for other data. However, geographic data themes in Table 2.3 are not an end to themselves, but recognized as dynamic and will continue to evolve in future (UNECA, 2007).

2.5.3 The SASDI Metadata

In compliance with the SDI Act, producers and custodians are obliged to publish their metadata and ensure that it is included in the electronic metadata catalogue (clearinghouse). Furthermore, in 1998 the NSIF technical framework team implemented the Spatial Metadata Discovery Facility (SMDF), previous called the Spatial Data Discovery Facility (SDDF) (Thabethe, 2008a, pers. comm.). The SMDF aims to improve spatial data access by enabling the data community to publish their metadata and search for data published by others (Osei, n.d.). The SMDF has a large amount of metadata documentation existing within its system. However, one of the main challenges faced by the NSIF is getting custodians and other people to capture, update and standardize their metadata (Thabethe, 2008b, UNECA, 2004b).

2.6 Conclusion

Project developers and other stakeholder types continue to use existing spatial data collected from previous projects. Organizations have gradually evolved from thinking in isolation to integration of information with different disciplines. The development of an SDI has been recognized as facilitating and coordinating exchange, sharing, access and use of spatial data.

The main components of an SDI are spatial data, people, institutional frameworks, technology and standards. These components and their various segments, including metadata, are inter-linked – one weak component may affect the entire operation of the SDI negatively. Spatial data is one of the key components that should exist at the start of the SDI development. People on the other hand, use data to support appropriate decision-making. Therefore, as data becomes a concern in socio-economic issues, a defined relationship between data and people should be developed. Information and communication technology improvements have lead to the wide use of digital spatial data and better access and sharing of information. Clearinghouses linked to other websites, together with hardware, software and networks, facilitate data sharing ease and reduce duplication of efforts.

Metadata describes the information of spatial data in depth. As a result, a user can easily understand what the data is about, without spending time inquiring about the missing information. According to previous research findings, some organisations have raised concerns about the use of metadata. Some point out that metadata is costly, time consuming, labour intensive and the metadata standards complex. As a result, such misconceptions could discourage the use of metadata in organisations. It is therefore important to educate the spatial data community about the benefits of metadata and to raise awareness about its importance. Another important point recognized from previous research findings is that a strong government can facilitate in the provision of relevant resources, knowledge, skilled metadata personnel and software tools.

In South Africa, the NSIF has faced challenges that include metadata problems. It has been observed that producers and custodians of spatial data are reluctant to comply with the SDI Act in terms of developing and maintaining their metadata. The following chapters continue to explore reasons behind the challenges encountered within spatial metadata.

CHAPTER 3: METHODOLOGY

3.1 Conceptual framework

The investigation concentrates on finding answers to the research questions. A significant guidance emerged from previous literature reviews related to SDI and metadata issues. Particularly, findings from previous research in other countries on metadata production and its limitations provided awareness on what could be expected for this study. However, the basic consideration was to ask questions associated with metadata challenges to a group of people in the spatial data community of South Africa.

The choice of selecting a sample population was dependent upon purposive sampling design. Purposive sampling strategies differ from probability or random sampling strategies in that they involve non-random selections. This allows researchers to apply their knowledge on selecting individuals, groups or organisations that are representative of the phenomenon under study (Kumar, 2005). Purposive sampling offers researchers a certain degree of control to the preferred group rather than relying on random selections, particularly if the group of interest is difficult to access (Barbour, 2001). In this research, it was relevant to select individuals from organisations that would provide the greatest insight into the research questions (Topp *et al.,* 2004). Producers and users of spatial data were the prospective population of interest, which were identified from various organisations associated with spatial data. In particular, the National Spatial Information Framework (NSIF) comprised of 13 producers for their framework datasets (See table 2.4, Chapter 2). Most importantly, the (NSIF) was identified to be the core provider and guidance for selecting the rest of population size for the data collection. The population size was geographically distributed around the country. Physical access to the sample population was unfeasible. Therefore, this influenced the type of primary data collection method. E-mailed questionnaires were more appropriate to execute for this research (Frazer and Lawley, 2000).

3.2 Data collection methods

Two approaches were used to gather information on metadata challenges, namely primary and secondary data collection. Primary sources provided first hand information gained from individual opinions. Secondary sources provided second hand information found in literature reviews (Kumar, 2005) and from analyzing metadata from some available framework datasets. Below is a discussion on the data collection process.

3.2.1 Primary data collection

Before data collection, the first step was to decide on sampling. According to Frazer and Lawley (2000), the type of sampling often influences the questionnaire administration approach. For example, the sample population could be widely distributed in a large area, hence imposing a mail or telephonic questionnaire. For this study, as explained in the conceptual framework, a non-random sampling design particularly purposive sampling was appropriate because controlled selection of the representative group was undertaken. The selection of the sample group, both producers and users, were identified through the NSIF. Organisations linked to NSIF most probably are associated with spatial data. Therefore, the respondents from the selected organisations possibly hold the relevant information require for the research questionnaires.

The sample population size was 152. The 152 prospective respondents were contacted by e-mail. Their e-mail addresses were obtained from the NSIF website of which provided further linkage to Surveys and Mapping website and eThekwini municipality website. However, the actual sample size settled to 27, precisely 14 data producers (including 4 framework dataset producers from NSIF) and 13 data users. Table 3.1 illustrates the various organizations from which the sample size was obtained.

Table 3.1: Sample Survey

Organization/Institution	**City**	**Data producers**	**Data users**
Business Connexion Microsoft Competency (BCX)	Johannesburg	√	
CSIR_Satellite Application Centre	Pretoria	√	
Department of Water Affairs and Forestry	Pretoria	√	
GISCOE	Midrand	√	
Knowledge Factory	Cape Town	√	
Municipal Demarcation Board	Pretoria	√	
Surveys and Mapping	Cape Town	√	
eThekwini Municipality_Environmental Management Branch	Durban	√	
eThekwini Municipality_Geology	Durban	√	
eThekwini Municipality Corporate GIS	Durban	√	
Land Resource International (LRI)	Pietermaritzburg	√	
University of KwaZulu-Natal_Geography Department	Pietermaritzburg	√	√
NCT Forestry Co-operative Limited	Pietermaritzburg		√
Land Affairs_Spatial Planning and Information	Pietermaritzburg and Mafikeng		√
QuarterX	Pietermaritzburg		√
uMsunduzi Municipality	Pietermaritzburg	√	√
Dept Of Local Government and Traditional Affairs	Pietermaritzburg	√	√

In this study, e-mail questionnaires were an appropriate and logical survey method to obtain information from a widely dispersed spatial data community. Two types of questionnaires were formulated, purposely for data producers and data users (See Annex 1). Both questionnaires contained open and closed questions designed to answer the research questions. De Vos *et al.* (2002) explain that open questions allow respondents to give extra information, yet in closed questions respondents are conditioned to select from a list of answers. Both questionnaires consisted of more closed questions than open questions. Consequently, the intention was to heighten responses. According to Kumar (2005), closed questions may ensure obtainable information required by the researcher and minimize time consumed to complete the questionnaires. Interestingly, three of the respondents expressed interest in the research topic and requested for an elaborate discussion on metadata issues. Furthermore, the respondents equally explained that their extra opinions were too broad to be presented on the questionnaires. Subsequently, a face to face contact was established with two of the respondents, who were based in Durban and Pietermaritzburg. Unfortunately the third respondent was based in the North West province. Therefore, it was unfeasible to execute the discussion due to financial limitations for the project in terms of conducting a face to face interview. A telephonic interview would have been an alternative. However, the discussion over the telephone would be limited an unstructured interview which is best conducted in a person to person interaction (Kumar, 2005)

Data was collected over a period of 3 months. Particularly, questionnaires were e-mailed from mid-November 2008 and responses were awaited until the end of February 2009. Unfortunately, approximately 54 e-mail addresses entailed permanent fatal errors. Thus, inquiries on current e-mail addresses were traced telephonically from some prospective organizations. An alternative to obtain responses for organisations based in Pietermaritzburg was to physically visit their offices because the research was being undertaken in the same area. As a result, table 3.1 illustrates more Pietermaritzburg respondents.

3.2.2 Secondary data collection

Preliminary information was gathered through literature reviews such as books, journals, conference proceedings, thesis, reports and other internet articles. Search for relevant literature sources was obtainable from the University of KwaZulu-Natal, Pietermaritzburg library and the internet. Previous literature sources provided guideline for a suitable methodology to conduct the study. In addition, internet sources provided contact information for the sample survey. Framework metadata provided by the Department of Land Affairs (DLA) was accessible on the NSIF website. Therefore, some of the framework metadata were used for comparison. Furthermore, compared metadata facilitated in the identification of existing gaps, consequently providing answers to some research questions.

3.3 Data Analysis

Answers to open and closed questions were coded for computer analysis. SPSS, a programme with good analytical abilities was used to analyze the completed questionnaires. Conversely, comparative analysis was used for the framework metadata. Both types of analysis offered insight about the conception of metadata challenges with regard to its use and management.

3.4 Problems encountered

Outdated e-mail addresses were an inconvenience. Time was wasted through the process of investigating for the correct and current contact details. Similarly, low questionnaire responses affected the time plan. Therefore, additional time was necessary to increase the number of respondents. Possibly, the data collection overlapped with end of year holiday period. Nonetheless, Kumar (2005) indicates that a small sample size could provide a good estimate if it includes a homogeneous population and low variation characteristics. Another difficulty was the inaccessibility of some framework metadata through their URLs. As a result, not all metadata for the different framework datasets were included in the comparative analysis.

3.5 Conclusion

This chapter highlighted the conceptual framework, data collection methods and data analysis used for the study. Initially, the conceptual framework stemmed from the general

SDI theoretical study and focused on the metadata component. The conceptual framework formed the basis of the research problem with focus on answering the research questions. Non-randomization was used to select the sample population. Hence a purposive sampling design was most appropriate. Secondary information was helpful in clarifying the theoretical aspect of the study area through previous literature reviews. Another secondary data collection involved the comparison of various framework metadata provided by DLA. Even though limitations were encountered, the data collected afforded adequacy for data analysis and ensured validity to objectives and answers to the research questions.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Introduction

Findings obtained from answered questionnaires were analyzed using SPSS. The majority of the questions required more than one answer from individual respondents. This required an SPSS multiple response analysis. Multiple responses were particularly obtained from closed-ended questions because most respondents chose multiple answers for one question. In general, the data was limited for statistical significance but findings were supported by previous studies in the literature review chapter. In addition, the comparative analysis for metadata framework dataset was displayed in tabular form. Overall, the results were displayed and interpreted with reference to the research questions.

4.2 Spatial data users and means to access data

4.2.1 Spatial data users

Fourteen producers from various organizations provided spatial data to 8 user categories: decision makers, institutions or organizations involved in spatial data, commercial users, academic community, Non Government Organizations (NGOs), consultants, donors and the media. The 8 various user categories are illustrated in figure 4.1.

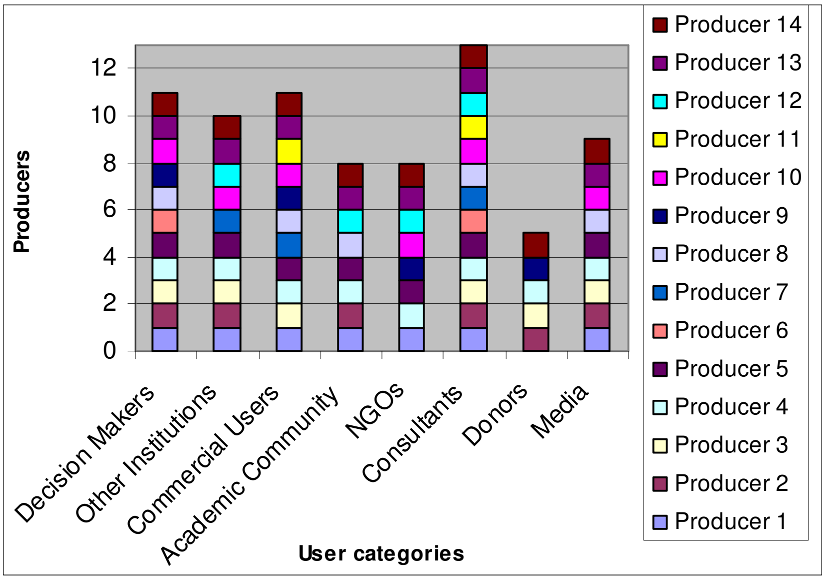

Figure 4.1: Spatial data provided by producers to various users categories

Multiple responses were obtained from the 14 producers illustrated in figure 4.1. Results suggest that consultants appeared to be the main users of spatial data. According to Buthelezi (2009, pers. comm.), most government sectors hire private consultants to gather their spatial data. In general, it may be assumed that private consultants aim to earn high profits to sustain company operations. For that reason, private companies are likely compelled to render efficient services to government spatial data projects. 11 of the producer respondents presented decision makers and commercial users as their second prominent users of data. Probably, decision makers in most development planning brackets require support from spatial data. Kok and van Loenen (2005) have noted that spatial information has significantly contributed to public task successes. Based on observations made by Foote and Lynch (1995), relatively, there has been a great range of spatial data types utilized. Therefore, the involvement of commercial users could indicate their awareness of the broader range of data uses that meets their needs. Other organizations involved in spatial data and the academic community were identified by 10 and 9 data producers respectively. The media and NGOs were shown to be among the few users because only 8 of the producers were their data providers. The donor category was signified as the least users. Most likely, donor agency interests tend to remain favorably on other social and economic problems rather than on direct data uses.

4.2.2 Partnership with other institutions/organizations

Furthermore, the 14 data producers were asked if they had any partnership with other institutions or organizations involved in the spatial data sector. 10 producers agreed to have partnership while the other 4 claimed otherwise.

A further survey was channeled at the 10 producers regarding the basis of collaboration with other organizations involved in the spatial data sector. Multiple responses were acquired for these results.

The partnership consisted of: data distribution, data capture, technology, technical assistance, metadata creation, metadata maintenance, funding, and license agreement (Figure 4.2).

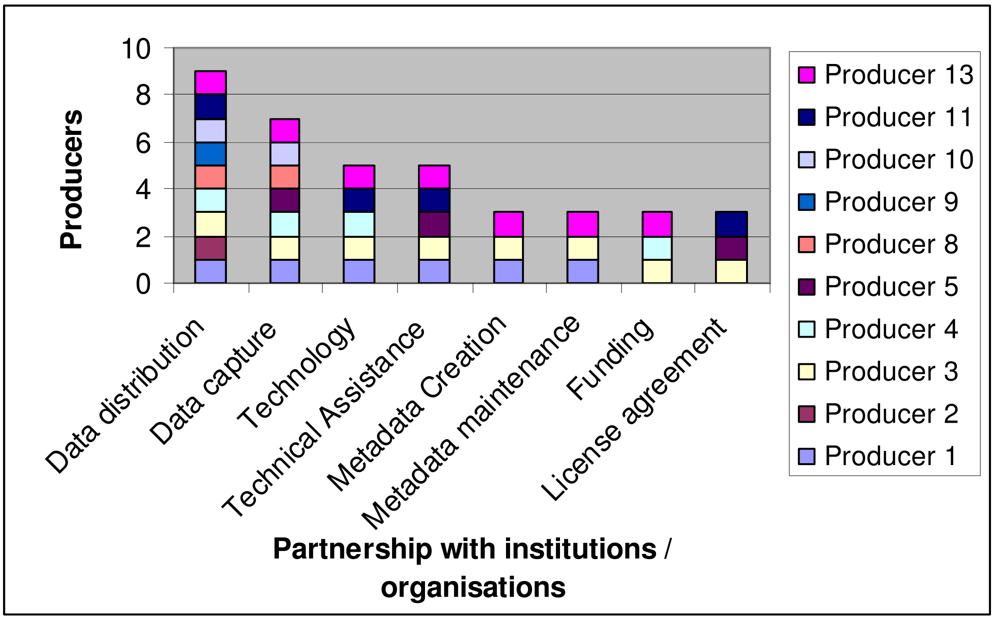

Figure 4.2: Basis of partnership between data producers and other Institutions / Organizations

According to UNECA (2004e), in most conditions as a single agency, it is almost improbable to possess all resources and knowledge required for data activities. Thus, it is important for SDI building activities to entail partnership and cooperation of various organizations. In figure 4.2 illustrations, 9 producers had a data distribution relationship with other organizations. Based on informal observations made by Buthelezi (2009, pers. comm.), the GIS industry is likely to survive because of data distribution relationship with one another. Further, in the discussion held with one of the producers, their company regularly exchange spatial data with other known small spatial data companies. This however, could create a small network community of isolation. Consequently, newcomers in the industry could experience difficulties in locating the relevant organizations to share data with.

Seven producers indicated data capturing as another common basis of partnership collaboration. However, Thabethe (2008b) mentioned that there overall, inefficient data collection still prevails in South Africa. Similarly, Buthelezi (2009, pers. comm.) has speculated duplication of data collection in the GIS industry. Furthermore, during a discussion, a producer mentioned that their company realized the data collected for a particular project had already been collected by another consultancy company involved in a different project. Therefore, a common place to archive data by multiple capturers could be a solution to less duplication of data.

Technology and technical assistance were used as collaborative activities at equal rates (5 producers). The results suggest that organizations provide technical services for the very technologies they are in collaboration with. In a previous study, Ezigbalike *et al.* (2000) noted that, effectiveness of data capturing partly depends on geospatial-information and communication technologies (Geo-ICT). Therefore, the degree of communication technology facilitating spatial data discoveries in South Africa is questionable because data collection duplication seems to persist, as highlighted in the previous paragraph.

Other important collaboration arrangements included: metadata creation; metadata maintenance; funding; and license agreements, however, each were least practiced (3 producers). In the case of metadata creation and maintenance, there is a slight possibility that: data originators are likely to create and maintain their own metadata; and perhaps, companies that assist in metadata creation and maintenance are uncommon in the spatial data industry. Also, during discussions with a respondent, it was mentioned that thorough metadata was unnecessary. Hence, such perceptions identify collaboration with metadata experts as needless.

In the view of several literature sources, lack of financial resources prevails in most developing countries. Depicted results on funding (figure 4.2) could assume that South Africa has few spatial data donors. Perhaps, as stated by (UNECA, 2004c), the spatial data market in developing countries could still be a concept in its infancy stage. Therefore, the spatial information market is yet to draw more attention to the relevant data investors from donor countries. Apart from few donor countries, for some reason, the annual budget from the government may be perceived inadequate for the SASDI to solve critical issues. Therefore, it becomes a bigger challenge if the country lacks financial resources and less support from outside donors.

Only a few organizations seem to be in license agreement collaboration. Most probably the copyright terms and conditions are emphasized when data is accessed.

4.2.3 Data access system

The various spatial data users in the 8 categories depicted in figure 4.1, access their data either through restrictions, non restrictions, authorizations or both through restriction and authorization (figure 4.3). 5 of the surveyed producers stated that an authorization was required prior to users accessing data. 3 of the producers mentioned access to their data was restricted. As stated by one of the respondents, perhaps their restricted data could only be view and not edited. Such data may be highly confidential to be used by the public users. 4 other producers stated that their restricted data had to go through the authorization process as well. Only 2 producers allowed free data access.

In general, once data has become a consumer service, producers of such data may highly be protective over their interests. According to Moyo (2008), in Malawi for instance, movement was towards unrestricted data exchange. Consequently, data originators were hesitant about their intellectual rights since users could easily acquire the data indirectly from other individuals. In South Africa the Copyright act of 1978 protects state owned spatial information and services. The Act mandates that individuals and the private sector may use spatial data products without specific authorizations, however, the state copyright must be acknowledged (CSI, 2002).

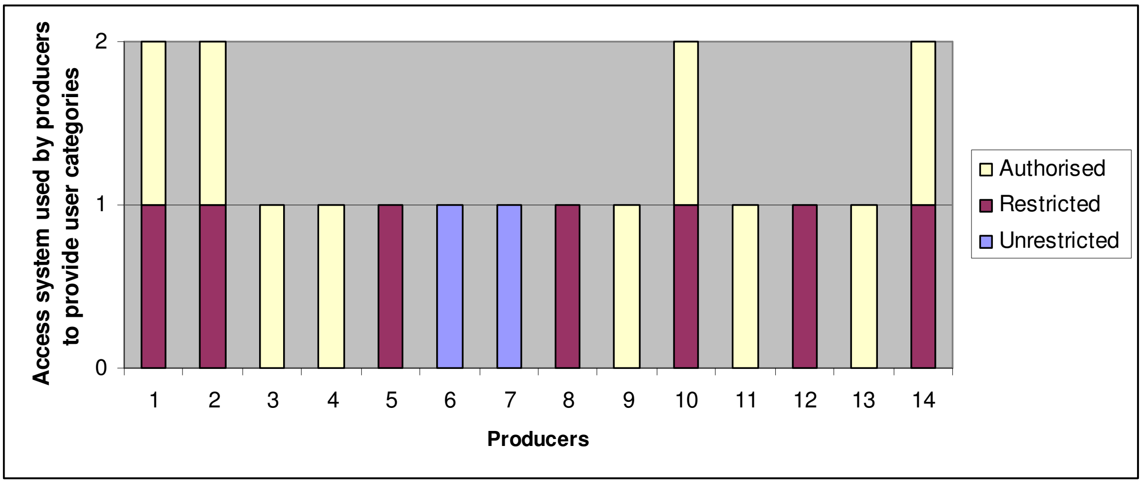

Figure 4.3: Access to available spatial data

4.2.4 Producer experiences in access conditions

The assessment of data access conditions was centered on data producer experiences. Permission to data access was based on charges at the office and on websites, depending on the license disclaimer of the organization. Results of the multiple responses were shown in figure 4.4.

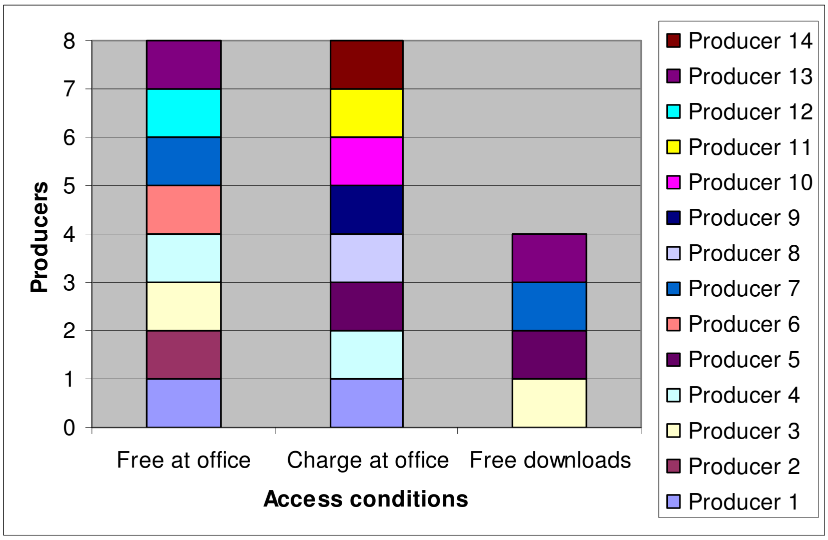

Figure 4.4: Producers experiencing data access conditions

Eight producers either charged or did not charge data when requested at the office. In these particular findings, the majority of the public sector producers did not charge data on request at the office. This may imply that the public sector complies with the Promotion of Access to Information Act of 2000. The Act requires information to be provided to other public institutions on a non-profit basis (CSI, 2002). Further, private users are supposed to be charged on what it costs the public data producer to produce the data for them (CSI, 2002). However, it has been noted that charges to private users may depend of the institution and the type of spatial data requested. For instance: one of the public sector producers indicated that charges were free to any individuals but the medium of the data was charged; and another respondent specified that private users were charged according to hardcopy paper size requested. Findings also revealed that all producers from the private sector charged their data on request, perhaps mainly because they operate on a profit system for datasets production.

On the other hand, a low response rate of 4 producers maintained free data downloads via their websites. This could mean most data users have to travel to the relevant offices to collect their data or it could be couriered. Yet, if more data is accessed via websites, this may result in reduced access time consumptions.

4.2.5 User experiences in access conditions

Thirteen users were surveyed on the basis of their data access conditions. 11 of these users had more than one answer for the closed-ended question. Figure 4.5 illustrates the results. More than half of the respondents 7 ordered from the dataset originator and received hardcopy data for free. In addition, a low response rate of (5) ordered datasets from the originator and paid for the hardcopy data. As previously discussed in figure 4.4, the findings in figure 4.5 closely suggest that, to promote data sharing, access should be received free or at affordable prices (CSI, 2002). Therefore, as one of the SDI policy goals, the public sector in particular, is obliged to promote data sharing (Thabethe, 2008b). Furthermore, from the 6 users that specified other data access conditions, it was stated that, data was traded for required data with other GIS users. This could suggest that even in the private sector, there may be existence of agreements to free data access between involved parties.

Similar to figure 4.4 findings, figure 4.5 results also illustrated that data downloaded online was far less common than data accessed as hardcopies. Further, a user respondent noted that data obtained online were mostly aerial photographs, which have to be in digital format for use. Often, it is in digital format that spatial data can be stored, analysed, managed and utilized by GIS tools (Rajabifard *et al.,* 2003). Thus, if much of the current data is still accessed as hardcopy format, this could hamper the growth of data sharing, particularly in clearinghouses.

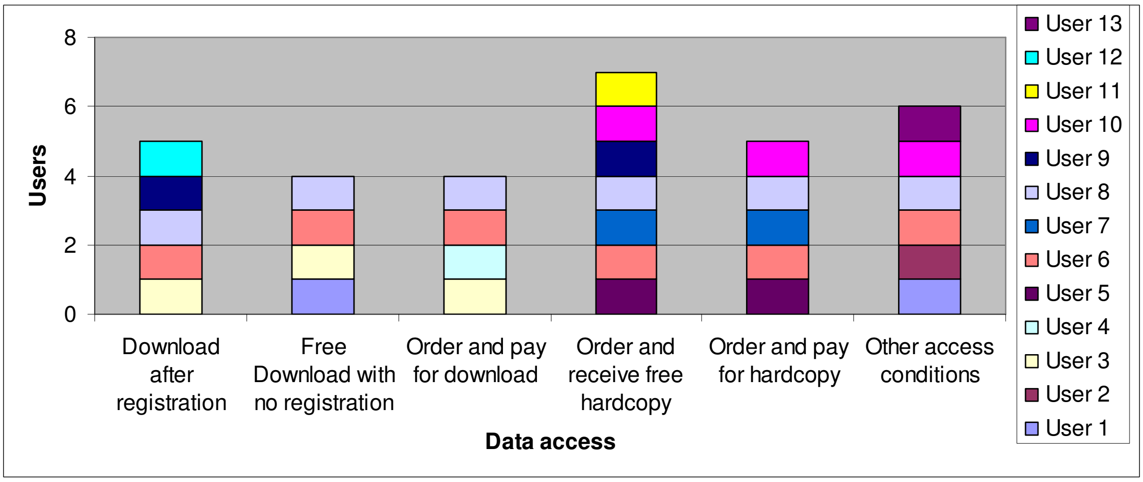

Figure 4.5: Users experiencing data access conditions

4.3 Metadata maintenance responsibilities

4.3.1 Metadata maintenance skills personnel

From the 14 surveyed producers, the diagram below illustrated that: 9 organizations had skilled personnel to maintain their metadata; and 5 indicated unavailable skilled personnel to maintain their metadata.

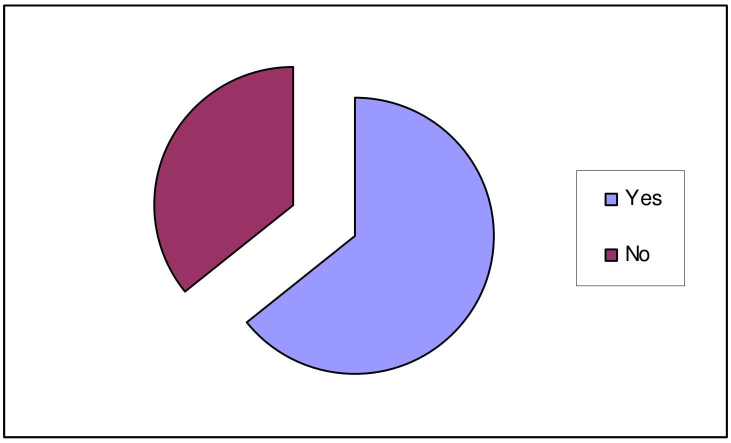

Figure 4.6: Condition of Metadata skilled personnel in the various organizations

Findings suggested that within the spatial data industry, metadata was in existence. As a result, more than half of the response rate indicated the availability of skilled personnel for metadata maintenance. Nonetheless, one of the producer respondents from the government sector realized metadata skilled workforce was not enough. Interestingly, none of the producer respondents from the private sector complained about unavailability of skilled metadata maintenance staff. One of the public sector respondents stated that their department commonly hired private consultants to maintain their metadata. Therefore, the respondent could have implied that it was unnecessary to retain internal metadata maintenance personnel. The results indicating the few public sectors lacking skilled metadata officers concerns their overall metadata quality. In other words, lack of skilled metadata personnel in an organization may lead to absent, incomplete and/or outdated metadata for their spatial datasets.

4.3.2 Metadata human resource qualification

A follow up was conducted on the qualification held by the metadata developer. Figure 4.7 showed that, 5 of the respondents that lacked metadata maintenance officers (as represented in the previous figure), were therefore unable to identify the skills required for developing metadata.

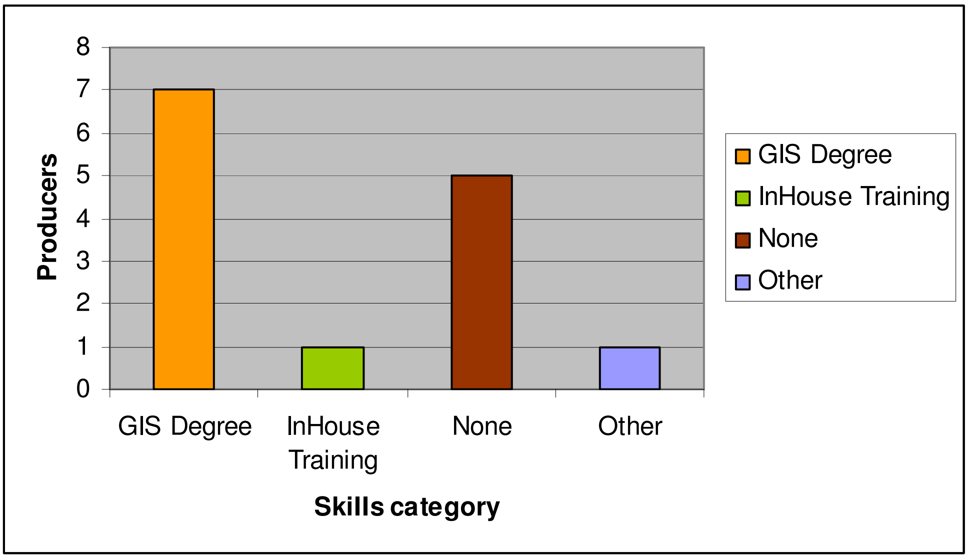

Figure 4.7: Qualification held by the metadata developer per organization

Nonetheless, from the figure above, half (7) of the producers stated that their metadata development and maintenance officers required a GIS degree. Only one respondent mentioned their organization offered in-house training for their metadata officers. One other producer (other skilled category) specified that their organization required qualifications ranging from a bachelor of commerce, actuarial science and a diploma in land surveying.

In general, the value of metadata is likely to be noticeable to GIS practitioners because they habitually handle spatial data (Norheim *et al.,* 2000). Therefore, it may be assumed that a GIS expert could require less intense training than an individual with no GIS background. However, individual knowledge about the dataset is vital. Thus, besides being a GIS expert that maintains metadata, a research phase with the spatial data originator may ensure accurate and comprehensive metadata development (Campbell, 2008).

4.3.3 Metadata currency

Spatial data producers were asked about their metadata update rate. Results in figure 4.8 showed that, 4 of the respondents regularly updated their metadata in less than 3 months. Perhaps this group of organizations had datasets that required regular modifications. The other 4 organizations updated their metadata only when necessary. As noted by one the respondents during a discussion, their metadata updates depended on the type of dataset. For example, datasets like dams did not involve frequent updates, yet water pipe datasets expected continual updates. Three of the organizations updated annually whilst one organization maintained metadata updates every 6 months. Two respondents (other update category) did not specify their update rates. Nonetheless, one of the respondents claimed to update metadata only when time allowed, whilst the other produce confirmed their department did not partake in metadata developments.

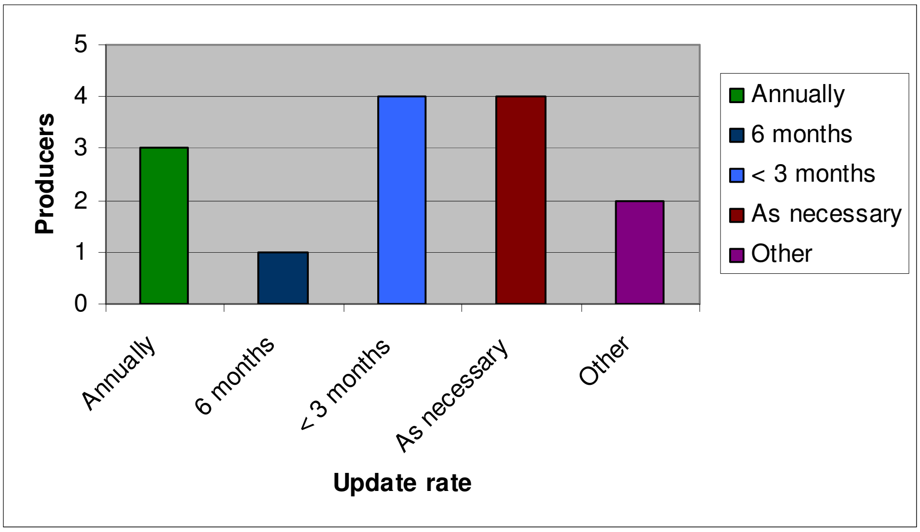

Figure 4.8: Metadata update frequency rate per organization

Interestingly, nearly all the organizations that experience lack of skilled metadata personnel in the previous findings (figure 4.6), mentioned their metadata was updated when necessary. Also, two respondents that specified other update rates from the above figure, revealed lack of metadata skilled workers in their organisations. A cross tabulation of the availability of metadata skilled personnel with the metadata update rate is displayed in the table below.

Table 4.1: Metadata skilled personnel with metadata update rate

		Metadata skilled personnel	
		Available	Not available
		Count	Count
Metadata update rate	Annually	3	0
	6 months	1	0
	3 months	0	0
	< 3 months	4	0
	As necessary	1	3
	Other	0	2

Alternatively, findings in table 4.1 may imply that, organizations that lacked metadata skilled workforce might not have sufficient capacity and time to regularly update their metadata. Consequently, organizations with shortages of metadata personnel may be producing datasets with outdated metadata.

4.4 Metadata standard inconsistencies

As aforesaid in chapter 2, the NSIF is mandated to the SDI Act to facilitate integration, access and sharing of spatial data. Previously, while international metadata standards were being developed, NSIF opted to use freely accessible metadata capturing tools that conformed to the FGDC content standard. On the other hand, According to (GISSA, 2006) the South Africa Bureau of Standards (SABS) adopted the ISO standard 19115 to be the official metadata standard for South Africa. Additionally, Osei (n.d.) presented that, the NSIF has participated in activities of the ISO TC211 standard developments. Furthermore, the NSIF and SABS jointly developed standards to supplement ISO Standards such as the national metadata standard (Osei, n.d.). The national metadata standard is acknowledged to be the South African implementation of the ISO metadata standard. Moreover, the spatial data community is obliged to capture metadata based on the national metadata standard (Osei, n.d.). The figure below illustrates some of the metadata standards used by organizations involved in spatial data.

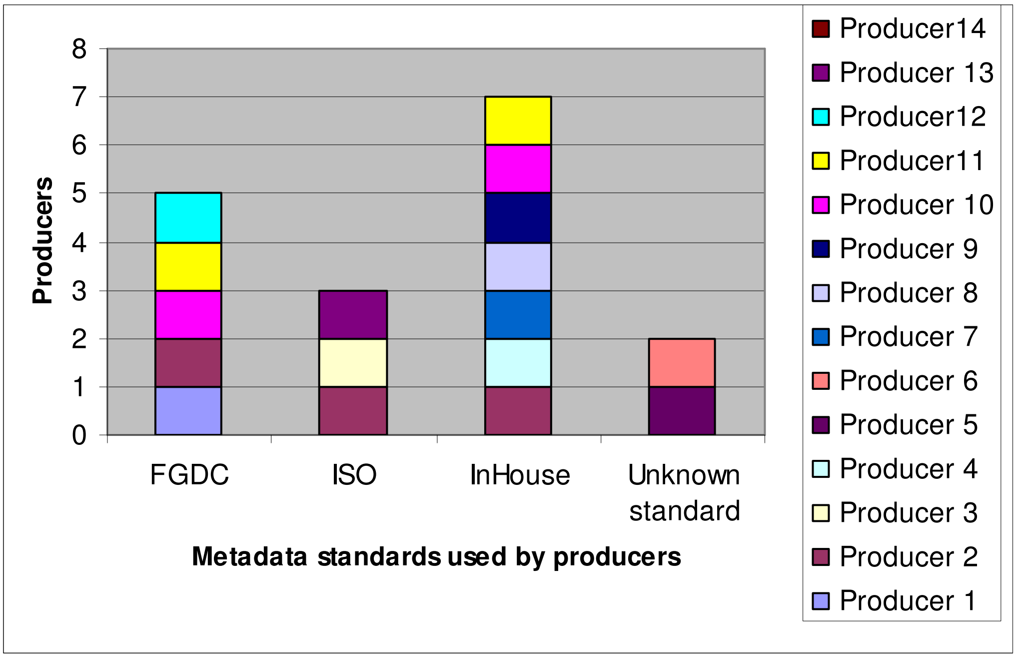

Figure 4.9: Metadata standards employed by various spatial data producers

Multiple responses in figure 4.9 showed that, unspecified in-house metadata standards were found to be used by 7 of the organizations. 5 organizations complied with the FGDC standard. ISO standards were the least used by 3 organizations, while 2 producers did not know the metadata standard type utilized by their organization. The findings may suggest reluctance of cooperation within the spatial data community to use a common metadata standard. As a result, a few organisations probably comply with the SDI Act. Ezigbalike *et al.* (2000) observed in broad spectrum that metadata structures in developing countries were still in rudimentary stages. As a suggestion, the public sector should represent abidance to the official metadata standard to gradually achieve common use with other spatial data members.

Table 4.2 below details the main similarities and differences depicted from 8 out of 13 metadata framework datasets. Due to retrieval problems on the NSIF webpage, 5 out 13 framework metadata could be compared.

FGDC was noted to be the main metadata standard for the framework datasets in table 4.2. The digital orthophoto images applied the 1.0 1998 FGDC CSDGM version whilst the other framework datasets represented the 1.0 June 1994 version. Further, if FGDC standard was commonly used for all framework datasets (depicted in the NSIF webpage), assumptions could be that most framework dataset producers were using FGDC standard.

With regard to metadata updates, the digital urban and rural cadastre datasets claimed to be updated continually. The digital Orthophoto images were maintained when necessary. On the contrary, the findings showed the datasets were neither updated continually nor as necessary. The reason being that, metadata creation dates for the 8 framework datasets ranged from late 1997 to early 1999. Further, the metadata maintenance and update content for the 1996 census enumerator areas were not specified. Lastly, the national land cover and catchment boundaries metadata maintenance were stated to be unplanned. Conversely, the land cover dataset producer respondent claimed their metadata was generated as necessary. Hence, it is highly probable not only for land cover but for other framework datasets that their metadata contents may be outdated.

Table 4.2 also depicted that, data presentation for the framework datasets is in map form except for the digital orthophoto images that are digital images. Furthermore, the maps are in vector formats and the digital images are in raster format. The projection parameters for the digital urban and rural cadastre maps are in Lambert Spheroid (modified Clarke 1880), whereas the digital Orthophoto images are in Gauss Conform projection. The rest of the framework datasets had missing reference systems. As pointed out by Thabethe (2008b, p.3) "Some strategic fundamental datasets have quality problems, which require urgent intervention". Thus, these problems are clearly revealed in the metadata for framework datasets. Overall, the freely available metadata framework datasets from the NSIF website seem to represent an update gap. This is an inconvenience for the data user since outdated metadata information could mislead decision-makings.

Table 4.2: Metadata for framework datasets, similarities and differences

Originator	Title	Data presentation Form	Spatial Reference Method	Projection parameters	FGDC CSDGM version	Maintenance and update frequency	Metadata Date	Metadata Contact Person
Surveyor General: Bloemfontein	Digital Urban and Rural Cadastre	Map	Vector	Lambert Spheroid: Clarke 1880 (modified)	1.0, June 8, 1994.	Continually	June, 1998	Nic Scheepers
Surveyor General: Cape Town	Digital Urban and Rural Cadastre	Map	Vector	Lambert Spheroid: Clarke 1880 (modified)	1.0, June 8, 1994.	Continually	June, 1998	Nic Scheepers
Surveyor General: Pietermaritzburg	Digital Urban and Rural Cadastre	Map	Vector	Lambert Spheroid: Clarke 1880 (modified)	1.0, June 8, 1994.	Continually	June, 1998	Nic Scheepers
Surveyor General: Pretoria	Digital Urban and Rural Cadastre	Map	Vector	Lambert Spheroid: Clarke 1880 (modified)	1.0, June 8, 1994.	Continually	April, 1998	Nic Scheepers
CSIR, Environmetek	National Land Cover	Map	Vector		1.0, June 8, 1994.	None planned	August, 1998	Rose Smith
Research Commission (WRC)	Catchment Boundaries	Map	Vector		1.0, June 8, 1994.	None planned	November, 1997	Rose Smith
Statistics South Africa (Stats SA)	1996 Census_Enumerator Areas	Map	Vector		1.0, June 8, 1994.		May, 1999	Sharti Laldaparsad
Chief Directorate: Surveys and Mapping	Digital Orthophoto Images	Digital image	Raster	Gauss Conform Clarke 1880 (modified)	01, 1998	As needed	June, 1998	Nic Scheepers

4.5 Metadata maintenance constraints

4.5.1 Metadata problems experienced by data producers

The following multiple response results highlight some of the metadata problems experienced by the data producers. Five of the producers did not share their views on metadata constraints. Perhaps this group of producers had not encountered metadata problems or they may be unaware of the metadata system. In particular, a data producer from the uMsunduzi municipality stated in the questionnaire that, their department did not maintain any form of metadata. Hence, their department was unaware of metadata constraints.

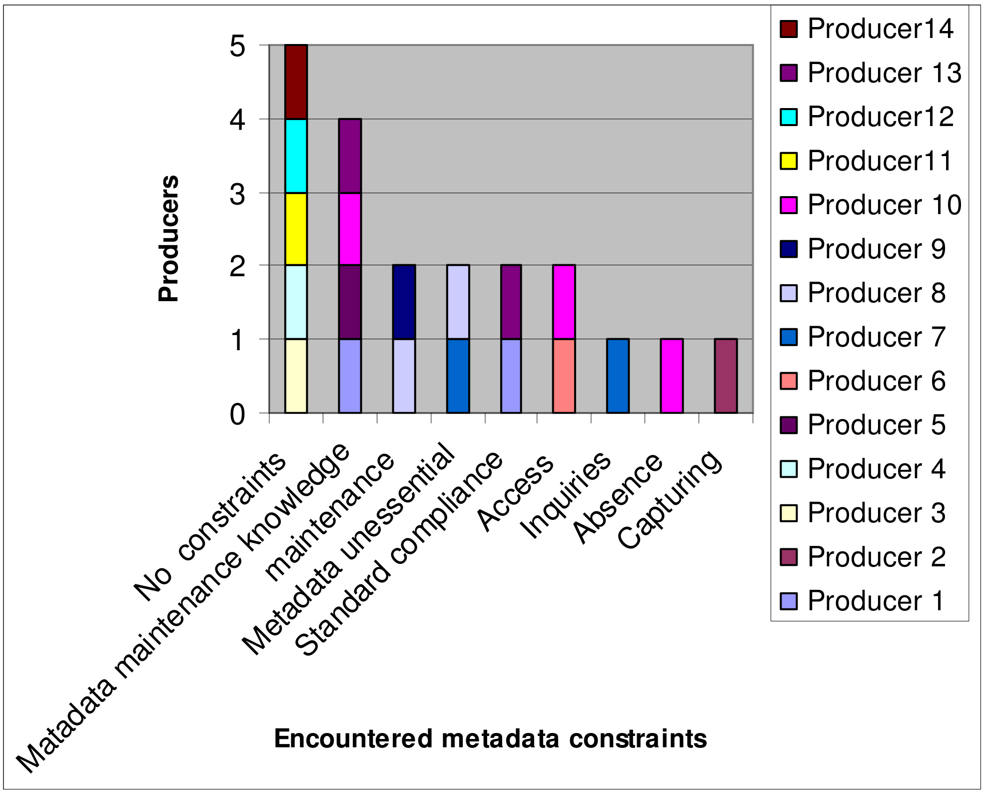

Figure 4.10: Main metadata constraints experienced by data producers

Lack of knowledge on metadata maintenance was recognized by 4 producers in figure 4.10. In other words, the producers presumed there existed more unskilled than skilled metadata maintenance personnel in organisations. Five producers in the previous figure 4.6 noted their organisations lacked metadata skilled personnel. However, from the 5 producer group in figure 4.6, only one producer recognized metadata constraint on lack of knowledge in figure 4.10. A cross tabulation of lack of knowledge on metadata maintenance with metadata skill personnel is referred to in the table 4.3.

Table 4.3: Lack of knowledge with metadata skilled personnel

		Metadata skilled personnel	
		available	not available
		Count	Count
Lack of knowledge on metadata maintenance	not recognized	6	4
	recognized	3	1

Table 4.3 depicted that most of the organizations which lacked metadata maintenance skilled staff did not recognize that lack of metadata knowledge was a constraint. Yet, it could be assumed these organizations were directly affected by this constraint.

Nonetheless, other producers were aware of shortcoming skills. For instance, one of the producers emphasized that there were shortages of good GIS experts in the public sector. In addition, a producer from one of the government sectors complained of being the only GIS staff personnel in their section. Hence, some GIS tasks became a challenge due to unavailable team work assistance. Another respondent noted that other departments within their municipality were reluctant to familiarize themselves to the upgraded version of ESRI (ArcGIS). It was perceived that several staff members were comfortable to continue utilizing ArcView 3.3. However, it is the upgraded ESRI software (ArcGIS) that includes ArcCatalogue which comprises of metadata creation (ESRI, 2003). Consequently, the lack of exposure to metadata creation tools could create a barrier to efficient metadata creation and maintenance.

Still in figure 4.10, two of the producers from 4 different categories criticized on the following:

- Developing metadata was seen as unessential in their organizations. Further, a data producer respondent perceived that, most of their data user clients seldom required information about the data. Therefore, it was pointless to capture metadata for their datasets. For this reason, it could be suggested that some spatial data producers lack incentives to create metadata. Also, lack of awareness on metadata benefits could be a driver to needless metadata creation in some organizations. According to Wayne (2005b), often the value of metadata is realized by staff members who later use data that was previously originated by the former data producer. In that case, it is important that metadata developers become aware of the long term metadata benefits to avoid loss of institutional data memory.

- Ensuring metadata maintenance was cumbersome, exceptionally for multiple datasets published on regular basis. It may be assumed that these respondents could be short staffed, thus they are unable to provide much attention to metadata systems. Wayne (2005a) observed that, a majority of organizations tend to focus intently on documenting existing data. However, metadata captured after data development appears to be cumbersome because metadata producers have to re-establish the data production steps and recall specific data values.

- Complying with metadata standards seemed difficult. In the view of some experts, metadata standards are perceived to be overwhelming because of the extensive number of elements. For example, Norheim *et al.* (2000) revealed from their findings that, many surveyed data holders raised concerns on the 199 elements, thus creating metadata standard complications. However, it was realized that many other elements were unnecessary for most datasets (Norheim *et al.,* 2000). Also, during a discussion, a respondent suggested that appropriate standard software with a few relevant elements should be adopted in the South African context. As identified in figure 4.9, half of the surveyed respondents employed in-house metadata standards. These in-house standards were perhaps more approachable for their organization instead of international standards that entails many elements (See example of eThekwini municipality metadata in Annex 2).

- There is lack of easy common access via the internet and between GIS members. The NSIF implemented the internet based SMDF aimed to improve access to spatial information. Simultaneously, the SMDF intend to facilitate the spatial data community to publish their metadata and to search for metadata held by others at a common place (Thabethe, 2008b). However, Osei (n.d.) emphasized that, over the past few years, the NSIF encountered problems with the functionality of the SMDF due to staff turnover. Therefore it may seem that the common access crisis is yet to be solved.

Also, three different respondents from figure 4.10 each critically pointed out the following:

- Inquiring to users about the same data information was an irritant. Most possibly the complaining organization held unreliable metadata for their datasets. Further, the same producer respondent formerly stated that their organization lacked skilled human resources for metadata maintenance. In addition, the organization was yet to determine on compiling comprehensive metadata. Thus, these could be some of the influencing factors that cause incomplete metadata for their datasets;
- Another producer complained some of the data they accessed did not hold metadata. There is a good possibility that data producers realize metadata necessity when they become users to confusing datasets. Therefore, awareness of metadata benefits is a critical subject to be addressed;
- Finally, one of the producers mentioned difficulty in obtaining sufficient information on metadata creation. It could be assumed that inadequate education on metadata creation persists. Thus, organizations in partnership through technical services should promote knowledge on the appropriate metadata capturing tools. Complicated metadata creation tools may result in high resistance to metadata capturing.

4.5.2 Users encountering problems with acquired data

Multiple response results in figure 4.11 showed that, 10 surveyed data users mainly encountered problems on acquired data with incomplete metadata. Additionally, 9 users expressed concerns on some of the acquired data with no metadata. Another group of users surveyed (7) recognized some of their accessed data consisted of non standardized metadata.

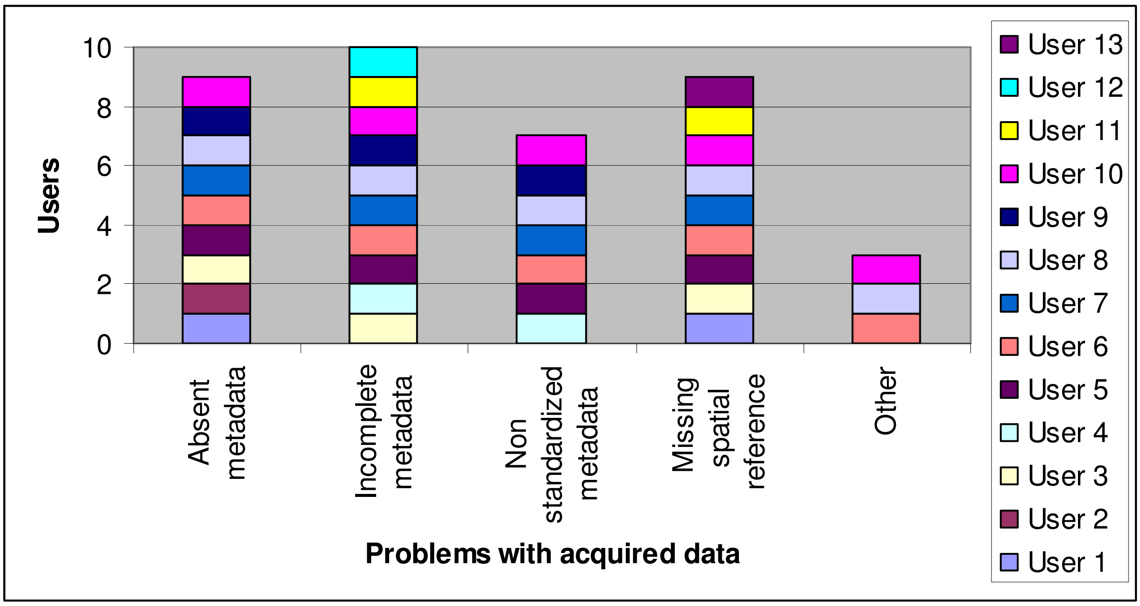

Figure 4.11: Common data problems encountered by spatial data users

The represented complaints suggest that there may be an overall use of insufficient metadata in the spatial data industry. Some data producers mentioned in the previous results that, metadata was unnecessary to develop. As a result, the metadata system could be the least of their priorities compared to other activities in the organisations. One strong assumption could be that, data holders are not pressurised to comply with the SDI Act of 2003. With respect to section 12 (1) of SDI Act No.54, 2003, "A data custodian must capture and maintain metadata for any spatial information held by it". Further, section 12 (2) of the SDI Act states that, "A data custodian must ensure that metadata is available to users". Also, section 14 (3.b) states that, "A data custodian or a data vendor supplying spatial information must provide the relevant metadata together with the spatial information".

Another 9 of the users was concerned about data that excluded appropriate information on spatial references. A discussion held with one of the users underlined that, most coordinate system errors are detected when the acquired datasets are overlaid on base maps. Consequently, it creates a duplication of effort to be checking data location errors.

In addition, it may be perceived that some data collectors might hold brief knowledge on precision collection of spatial points with GPS. Nevertheless, the data originator is responsible to check projections and location accuracies before publishing the data. A respondent who participated in a discussion commented that, data inaccuracies affect metadata accuracies. Therefore, the user respondent noted that it was critical to resolve data issues before centralizing concerns on metadata.

Other valuable opinions stated by very few users on encountered data problems included: poor understanding of data; incomplete capturing of vector datasets; and acquiring data with inaccurate metadata.

4.5.3 Suggestions on metadata improvements

Findings in figure 4.12 displayed diverse suggestions on ways to improve the use of metadata. Nine of the 13 surveyed users proposed metadata should entail compulsory fields to capture, particularly spatial reference details. As revealed in figure 4.11 results, a strong area of concern was about data with inaccurate spatial references. Therefore, metadata maintenance officers should be aware of the common metadata fields that are most relevant for data users. The national metadata standard for spatial data however, seems to be adequately appropriate for the South African context. For example, according to the technical framework presentation by Osei (n.d.) in the NSIF video, the national metadata standard comprises of the following elements:
Dataset title, Reference data, Custodians contact list, Geographic coverage, Language, Topic category, Spatial resolution or scale, Abstract, Distribution format, Spatial presentation type, Reference system, Quality statement, File identity and Date stamp.

A user proposed that the SDI Act requirements for metadata capture should be simplified. For this reason, there might be a few people in the GIS industry who find the SDI policy complex to follow. Consequently, the aim should be to promote awareness about the SDI policy to the data community. Results showed that 5 user respondents emphasized that all GIS members should conform to the SDI Act of 2003. Therefore, other GIS members are familiar to the SDI policy requirements. Another 5 respondents pointed out that, the GIS community should have common access via internet services. This suggestion links to the metadata constraint in figure 4.10. As aforesaid in figure 4.10 discussions, the SMDF has been facing functionality problems (Osei, n.d.). However, the SMDF is currently undergoing testing. Gradually with time, many spatial data members would possibly be registered as members and efficiently exchange data on the clearinghouse. In consequence, a successful common data access point might result in fewer data fragmented network communities

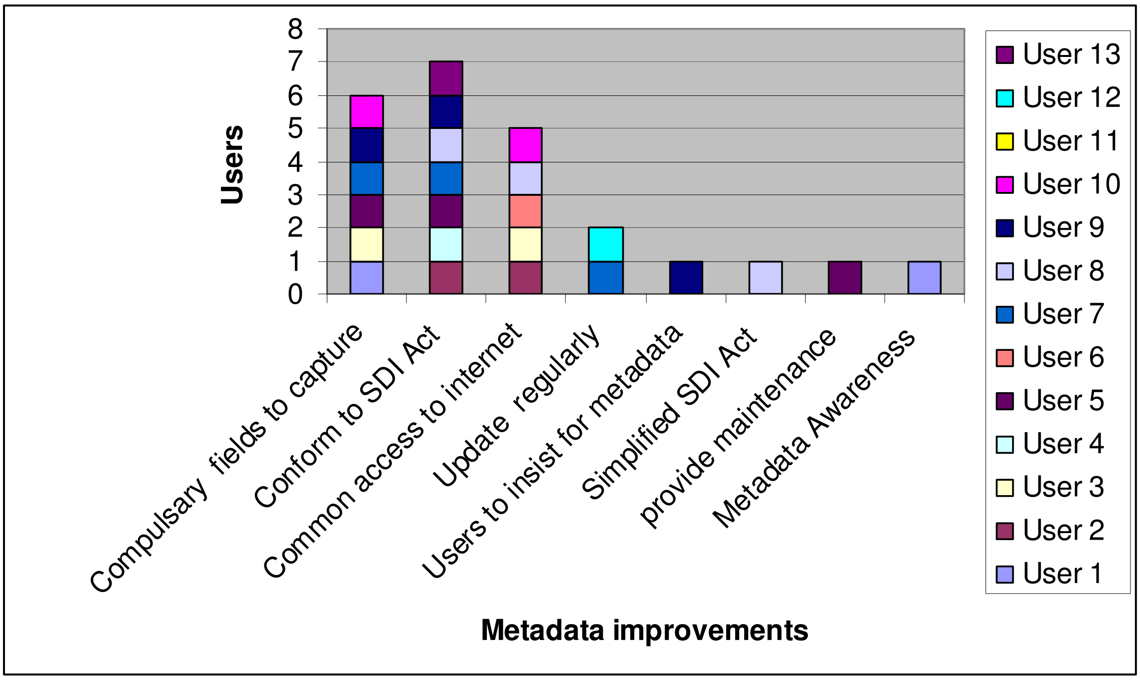

Figure 4.12: User opinions on metadata improvements

Still in figure 4.12, two of the users suggested that datasets should be regularly updated. One respondent proposed that users should insist for metadata if their acquired data has absent metadata. In that case, an attitude of interest and concern for metadata improvements must grow within spatial data producers and users. It was suggested by another respondent that, the national government should provide leadership on data maintenance. Additionally, a respondent during a discussion proposed that, the NSIF should establish a team to regularly monitor GIS organizations through door to door inspections. Further, such regular inspections could possibly lead to effective involvement in the national SDI. However, if the NSIF for example, continues to request data via e-mail from the various GIS member organizations, it could result to some low level of cooperation. Nonetheless, the question still remains on the GIS expert staff turnover in the public sector to lead the role. Finally, a crucial suggestion that was noted only by one user was to raise metadata awareness.

4.6 Conclusion

The findings of the surveyed producers and users provided an overview of the gaps and shortcomings of metadata use. To begin with, results indicated that very few data users accessed their data via the internet. This creates ineffective data sharing because one has to travel to the office to collect data. Thus, this could be time consuming. If most data was accessed from a common place, not only will it be quicker to access but could lead to increased digital data use. Hence, the use of digital data could be manipulated with GIS software. For that purpose, perhaps more metadata could be captured because the metadata creation tools will be available on the software.

Very few organisations were involved in funding collaborations. According to Thabethe (2008b), currently SASDI faces challenges such as, lack of effective collection and datasets maintenance, which require adequate funding. Therefore, lack of financial resources can significantly impede on the efficient use of spatial data which could further lead to poor decision-makings on socio-economic developments.

Findings suggested that, there was lack of metadata maintenance skilled personnel, particularly in the public sector. Shortage of staff consequently implies that the available metadata maintenance officer may update metadata only when time allows. Hence, the data might be outdated for a period of time. Outdated datasets together with their metadata may inconvenience and mislead data users.

It was found that inconsistency of metadata standards prevail in the SASDI. A number of the surveyed organizations opted to create in-house metadata standards appropriate for their use. The NSIF has finally developed a national metadata standard which is based on the ISO standard. Therefore, all organisations involved in spatial data should adopt the national metadata standard to promote interoperability within the SASDI context. At present, the framework dataset metadata available on the NSIF website are still in the

FGDC version. These framework datasets contain outdated metadata content. Therefore, this creates a duplication of effort and time consumed because users would have to further enquire for updated data information.

Data producers and users were aware of some metadata problems. Results showed that individuals had various perceptions on metadata constraints. Few of the producer respondents recognized there was lack of knowledge on metadata capturing and maintenance. As suggested before, there may be a shortage of staff members with good GIS background. Most data users either acquired data with incomplete or absent metadata. Another concern was acquiring data with missing spatial reference systems. It is suggested that data suppliers should be encouraged to maintain their data with complete metadata documentation

Lastly, results showed various ideas on metadata improvements. A good number of user respondents suggested metadata should entail compulsory fields. However, metadata compliant with standards such as FGDC and ISO contain metadata elements relevant to data users. Therefore, concern should be on employing more effective metadata capturers. Another critical suggestion raised was that, GIS users should have common access via the internet. It has been reported that the NSIF is in the process of officially launching an electronic metadata catalogue. Therefore, the system aims to facilitate the sharing, dissemination and exchange of spatial. Equally important was emphasis on the spatial data community to comply with the SDI policy. It is however, imperative for an awareness campaign by the NSIF on the newly functioning of the SMDF, SDI in general, and SASDI metadata standards.

CHAPTER 5: FINAL CONCLUSION AND RECOMMENDATIONS

5.1 Final Conclusion

The overall research explored the aim and objectives in support of research questions. Results revealed that the private sector consultants seemed to be the main users of data. It was deduced that government departments commonly hired consultants to capture their data. However, duplication of data capturing was noted to be a major expense for several projects which could lead to major waste of resources. This requires an internet-enabled spatial server that facilitates people to discover data they need. The NSIF has recently completed the testing of the Spatial Metadata Discovery Facility (SMDF). Thus, the directorate is currently preparing an official launch of the clearinghouse system.

Findings suggested that hard copy data exchange was more common than digital data distribution and dissemination. Therefore, less data was accessed via the internet. Consequently, data users would have to travel to the relevant providers to collect their data. The inexistence of a common data access point has lead to fragmented data sharing communities. Therefore, inexperienced data users or newly developed organizations would initially struggle to build data distribution partnership with other organizations involved in spatial data.

For successful dataset integration into the SMDF, the metadata should be well evolved towards documentation of digital spatial data. Results showed that shortage of metadata maintenance skilled personnel persisted in the public sector more than in the private sector. The NSIF in particular, had past problems with the functionality of the SMDF due to staff turnover (Osei, n.d.). Lack of metadata maintenance officers in an organization often leads to insufficient metadata use. The metadata officers or data producers can only update their dataset metadata when time allows. Therefore, datasets that require continual metadata updates could remain outdated for a long period of time. It is crucial for the metadata maintenance officers to at least hold a GIS degree. Some organizations provide in-house training and others hire personnel initially with less background on GIS. Individuals with less GIS skills tend to know little about metadata. Therefore, the metadata maintained by such individuals could end up incomplete and inaccurate.

Technical resources have become a major setback to sufficient metadata use. For instance, upgrading from the most commonly used software could still be a challenge for some municipalities. The ESRI upgraded ArcGIS version contains ArcCatalog that supports metadata management (ESRI, 2003). Therefore, software that contains metadata records is a step ahead to metadata use than a software version that does not identify metadata elements. Overall, the country is stated to lack financial resources (Thabethe, 2008b). Thus spatial data budget problems might prevail in most government departments. Therefore, budget problems disable organizations to improve on their technical resources and skills training.

Findings showed that data users commonly encountered spatial reference system errors from their acquired data. This common complaint could imply that some data locations are inaccurately being collected. Lack of quality data impedes on appropriate spatial developments and effective decision-making. Moreover, dataset errors automatically display metadata with errors. As a result, the metadata becomes inadequate and can easily mislead the data user.

To foster effective data documentation, a metadata standard commonly recognized by the spatial data community should be adopted. The use of similar metadata standards increases interoperability of datasets for easy distribution and dissemination. Hence, the operation of the SMDF will require standard metadata documentation of published spatial datasets. Therefore, all participants as spatial data providers would have to adopt the metadata standard appropriate for the clearinghouse. However, most metadata standards utilized by many organizations are at rudimentary stage. This indicates that the level of compliance with the SDI policy is underdeveloped within the spatial data community. Perhaps also, in-house metadata standards are less complex and more approachable to use. The public sector, particularly the NSIF division should move from focusing on concept to practice. For instance, the metadata framework datasets are still displayed on the NSIF website as the FGDC version. Yet, the NSIF technical framework team stated that, national metadata standards supplementing ISO standards have already been developed (Osei, n.d.). The presentations of the metadata for framework dataset also signify inadequate metadata updates. Perhaps the NSIF division has shortages of metadata maintenance staff. Framework datasets are part of a national assert important for the realization of socio-economic and environmental developments. Also, as base maps for other datasets, framework datasets should be referenced to common spatial reference systems. Furthermore, decision makers and other data users may rely on framework datasets for further spatial data discoveries. Partly for the above reasons, it is central for framework dataset metadata to maintain good quality and constantly be updated.

It is assumed that, to a certain degree the spatial data producers and users lack metadata benefits awareness. According to opinions from some data producers, metadata is unnecessary to maintain. The reason being, most data users seldom insist for metadata. Gained knowledge on metadata benefits could guide users to demand for better data which includes metadata. Supposedly, organizations reluctant to maintain metadata are yet to realize its importance. In the long run, undocumented data could get lost and new staff may experience difficulty in understanding the existing spatial products. Some people tend to avoid metadata creation due to workload regarding datasets that entail frequent publishing. Additionally, perceptions about metadata problems could encourage spatial data managers to take metadata management low priority in relation to other activities.

5.2 Limitations

The results were based upon 27 respondents (14 producers and 13 users) out of 152 potential respondents. The small response size was partly due to administering questionnaires via e-mail. Therefore, it was difficult to continuously encourage the unwilling group to take part as they could simply choose to ignore the e-mails or forget to respond. This could signify that, if people fail to respond to e-mails, perhaps they are overworked with other activities. Therefore, this could result to unattended or pending matters such as metadata that require updates. Face to face and telephonic interviews would have been an alternative data collection method. However, the research project was financially constrained. Also, time constraints did not allow for face to face interviews because of the distances where some respondents were based.

5.3 Recommendations to the spatial data producers and users

The goal of the SASDI is to improve the availability, accessibility and applicability of spatial data. The SDI comprises of components that enable the system to achieve its objectives. The metadata component in particular, describes characteristics of the original data, thus used to facilitate access and spatial data sharing. One of the SDI Act objectives is to provide for the capturing and publishing of metadata as well as develop an electronic metadata catalogue. To reach the above goals, the following strategies need to be pursued:

- Promotion of the national metadata standards implemented for use in the electronic metadata catalogue. The promotion will encourage maximum consistency of standards adoption by multiple spatial data participants. This calls for awareness campaigns by the NSIF with assistance from other sectors such as the Spatial Planning and Information (SPI)

- With the upcoming official launch of the SMDF, the NSIF should encourage and allow many stakeholders to be involved in the clearinghouse. In the long run, the idea is to have as many members as possible at all levels within the government and the private sector. The outreach activity should not only be via e-mails but also through workshops and conferences.

- One of the challenges faced by the NSIF is getting spatial data producers and custodians to publish their metadata. The SDI Act regulations are currently under compilation. There will be clearly defined policies regarding the spatial data and metadata. However, policy implementations are useless without imposing solid and severe measures. A suggestion is to encourage spatial data providers maintain their metadata. Additionally, there should be a dedicated policy monitoring team for each region that constantly visits organizations for inspections. Consequently, the organizations involved in spatial data might be less reluctant to maintain their metadata.

- In the SASDI context, GIS capacity building persists to lack. GIS training from all education levels should be increased. The introduction of GIS from secondary school level is an important awareness initiative. The mission is to increase interested candidates to pursue spatial related courses at tertiary level. In the work environment, continuous training should be at equal pace with commonly used technology programmes such as ESRI.
- Lack of financial resources impedes SDI improvements. Thus, there is need to pursue in outside funding. However, it requires considerable effort to interest donors on spatial data investments. Ideally, raising metadata awareness is the key factor. Possibly, metadata benefits should be publicized to top leaders and decision-makers since they have more influence in accessing financial resources.

- The spatial data community plays a key role in enhancing efficient SDI and decision-making. However, it takes strong cooperation for the whole community to build towards SDI goals. Strong cooperation between spatial data members could for example, guide them to demand better data and insist for missing metadata from acquire data.

REFERENCES

BARBOUR, R. S. (2001) 'Checklists for improving rigour in qualitative research: a case of the tail wagging the dog?' *BMJ,* 322, 1115 - 1117.

BATCHELLER, J. K. (2008) 'Automating geospatial metadata generation—An integrated data management and documentation approach'. *Computers & Geosciences,* 34, 387-398.

BERNARD, L., KANELLOPOULOS, I., ANNONI, A. & SMITS, P. (2005) 'The European geoportal - one step towards the establishment of a European Spatial Data Infrastructure'. *Computers, Environment and Urban Systems,* 29, 15-31.

BISHOP, I. D., ESCOBAR, F. J., KARUPPANNAN, S., SUWARNARAT, K., WILLIAMSON, I. P., YATES, P. M. & YAQUB, H. W. (2000) 'Spatial data infrastructures for cities in developing countries: Lessons from the Bangkok experience' *Cities,* 17, 85-96.

BISHR, Y. & RADWAN, M. (2000) GDI architectures IN GROOT, R. & MCLAUGHLIN, J. (Eds.) *Geospatial Data Infrastructure: Concepts, Cases and Good Practice.* Oxford, Oxford University Press.

BROX, C., BISHR, Y., SENKLER, K., ZENS, K. & KUHN, W. (2002) 'Toward a geospatial data infrastructure for Northrhine-Westphalia'. *Computers, Environment and Urban Systems,* 26, 19-37.

BUDHATHOKI, N. R. & NEDOVIC-BUDIC, Z. (2007) *Expanding the Spatial Data Infrastructure Knowledge Base*. United States, Esri Press.

BUTHELEZI, S. (2009). Pietermaritzburg, Spatial Planning and Information.

CAMPBELL, T. (2008) *Fostering a Culture of Metadata Production*. GSDI-10. St. Augustine, Trinidad.

CHAN, T. O., FEENY, M., RAJABIFARD, A. & WILLIAMSON, I. (2001) 'The Dynamic Nature of Spatial Data Infrastructures: A Method of Descriptive Classification'. *Geomatica,* 55, 65-72.

CLARKE, D., OLAF, H. & WATKINS, W. (2003) Development of the Australian Spatial Data Infrastructure. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

COMMITTEE FOR SPATIAL INFORMATION (CSI) (2002) Policy on the pricing for and copyright of spatial information products and services. NSIF, South Africa.

COMMITTEE FOR SPATIAL INFORMATION (CSI) (2004) Committee Spatial Information: Policy Sub Committee. South Africa, NSIF, South Africa.

CROMPVOETS, J. & BREGT, A. (2003) 'World Status of National Spatial Data Clearinghouses' *URISA Journal* 15, 43-50.

CROMPVOETS, J., BREGT, A., RAJABIFARD, A. & WILLIAMSON, I. P. (2004) 'Assessing the worldwide developments of national spatial data clearinghouses'. *International Journal of Geographical Information Science,* 18, 665-689.

CROSSWELL, P. L. (2000) The role of standards in support of GDI. IN GROOT, R. & MCLAUGHLIN, J. (Eds.) *Geospatial Data Infrastructure: Concepts, Cases and Good Practice.* Oxford, Oxford University Press.

DE MAN, W. H. E. (2007) Are Spatial Data Infrastructures Special? United States, ESRI Press.

DE VOS, A. S., STRYDOM, H., FOUCHE, C. B. & DELPORT, C. S. L. (Eds.) (2002) *Research at grass roots: For the Social Sciences and Human Service Professions,* Pretoria, Van Schaik.

DENG, M. & DI, L. (2008) *Enhancing Uses of Spatial Data through Building Online Learning and Research Environment*. GSDI-10. St. Augustine, Trinidad.

DENG, Y. (2002) *The Metadata Architecture for Data Management in Web-based Choropleth Maps* College Park, University of Maryland.

DEPARTMENT OF LAND AFFAIRS (DLA) (n.d.) *Spatial Data Infrastructure: Act Made Easy,* Pretoria, South Africa, Department of Land Affairs.

DUFF, W., HOFMAN, H. & TROEMEL, M. (2003) Getting What You Want, Knowing What You Have, and Keeping What You Need. Marburg, Germany, ERPANET Training Seminar.

DUSHIMYIMANA SIMBIZI, M. C. (2008) Developing a Spatial Data Infrastructure for Rwanda: Case Study of Land Administration Sector. *CEAD.* Pietermaritzburg, University Of KwaZulu-Natal.

DUVAL, E. (2001) 'Metatada Standards: What, Who and Why'. *Journal of Universal Computer Science,* 7, 591-601.

EAGLESON, S. & ESCOBAR, F. (2003) Administrative Boundary Design in Support of SDI Objectives. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

ESRI (2003) *ESRI Profile of the Content Standard for Digital Geospatial Metadata*. New York, ESRI.[Online] www.esri.com Accessed 4 May 2009

ESRI (2008) *Creating and Maintaing Metadata* Using ArcGIS Desktop ESRI. [Online] www.esri.com/metadata/esriprof80 Accessed 4 May 2009

EZIGBALIKE, C., SELEBALO, Q. C., FAÏZ, S. & ZHOU, S. Z. (2000) *Spatial Data Infrastructures: Is Africa Ready?*. Fourth Spatial Data Infrastructure Conference. Cape Town, South Africa.

FEENEY, M. F. (2003) SDIs and Decision Support. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

FEDERAL GEOGRAPHIC DATA COMMITTEE (FGDC) (2000) *Content Standard for Digital Geospatial Metadata Workbook version 2.0.* Washington D.C., Federal Geographic Data Committee.[Online]www.fgdc.gov Accessed 11 October 2008.

FEDERAL GEOGRAPHIC DATA COMMITTEE (FGDC) (2006) *Geospatial metadata Federal Geographic Data Committee.* [Online]www.fgdc.gov/metadata/index Accessed 27 October 2008.

FEDERAL GEOGRAPHIC DATA COMMITTEE (FGDC) (2008) *Standards.* [Online]www.fgdc.gov/standard Accessed 15 October 2008.

FOOTE, K. & LYNCH, M. (1995) *The Geographer's Craft Project*. Boulder, University of Colorado

FRAZER, L. & LAWLEY, M. (2000) *Questinnaire Design and Administration: A Practical Guide,* Brisbane, John Wiley and Sons Australia, Ltd.

GAVIN, L. (2001) *Building SDI in Africa: A South African Expereince*. GSDI 5. Cartagena De Indias, Colombia.

GEOCOMMUNITY (2007) *GISDataDepot: GIS Data Formats* Mindsites Group. [Online] http://data.geocomm.com/helpdesk/formats Accessed 11 November 2008.

GEO-INFORMATION SOCIETY OF SOUTH AFRICA (GISSA) (2006) *SDI Overview*. Plone foundation. [Online] www.gissa.org.za/government/sdi Accessed 8 November 2008.

GRANT, D. & WILLIAMSON, I. (2003) State SDI Initiatives. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

GYAMFI-AIDOO, J., SCHWABE, C. & GOVENDER, S. (2007) *Determination of the Fundamental Geospatial Datasets for Africa through a User Needs Analysis* Addis Ababa, Ethiopia, UN Economic Commission for Africa.

HARRIS, P. (2008) *The Role of Spatial Information in Rebuilding a Sustainable Future for Aceh*. GSDI-10. St. Augustine, Trinidad.

HOLLAND, P. & BORRERO, S. (2003) Global Initiatives. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

HUNTER, J. (2003) *Working towards MetaUtopia - A Survey of Current Metadata Research.* Australia, University of Qld.

JACOBY, S., SMITH, J., TING, L. & 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, 305 — 322.

KAINZ, W. (2000) The foundation technologies. IN GROOT, R. & MCLAUGHLIN, J. (Eds.) *Geospatial Data Infrastructure: Concepts, Cases and Good Practice.* Oxford, Oxford University Press.

KOK, B. & VAN LOENEN, B. (2005) 'How to assess the success of National Spatial Data Infrastructures?' *Computers, Environment and Urban Systems,* 29, 699–717.

KUMAR, R. (Ed.) (2005) *Research Methodology: A Step-By-Step Guide for Beginners,* London, Sage

LIMBACH, T., KRAWCZYK, A. & SUROWIEC, G. (2004) *Metadata Lifecycle Management With GIS Context.* 10th EC GI & GIS Workshop. Poland.

MAGUIRE, D. J. & LONGLEY, P. A. (2005) 'The emergence of geoportals and their role in spatial data infrastructures'. *Computers, Environment and Urban Systems,* 29, 3-14.

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

MATHYS, T. (2004) *The Go-Geo! Portal metadata initiatives.* Proceedings of the Geographical Information Science Research UK 12th Annual Conference. University of East Anglia, Norwich, UK.

MOELLERING, H., BRODEUR, J., DANKO, D. M. & SHIN, S. (2006) *Towards a North American Profile of the ISO 19115 World Spatial Metadata Standard.* GSDI-9. Santiago, Chile.

MOYO, M. E. (2008) *Lack of Geospatial Data Copyright Protection and Liability: Can Malawi Really Afford This?* GSDI-10. St. Augustine, Trinidad.

NEBERT, D. (2004) *Developing Spatial Data Infrastructures: The SDI Cookbook*. Global Spatial Data Infrastructure (GSDI). [Online] www.gsdi.org/docs2004/cookbookV2.opdf Accessed 2 October 2008.

NEDOVIC-BUDICA, Z., FEENEYB, M.-E. F., RAJABIFARDB, A. & WILLIAMSONB, I. (2004) 'Are SDIs Serving the needs of local Planning? Case study of Victoria, Australia and Illinois, USA'. *Computers, Environment and Urban Systems,* 28, 329-351.

NOGUERAS-ISO, J., ZARAZAGA-SORIA, F. J., LACASTA, J., BEJAR, R. & MURO-MEDRANO, P. R. (2004) 'Metadata standard interoperability: application in the geographic information domain'. *Computers, Environment and Urban Systems,* 28, 611–634.

NORHEIM, R. A., PETERSON, D. L., ALCOCK, T. Z. & CHRISMAN, N. R. (2000) *Metadata SWAT Teams: A Success Story in Creating a Regional Clearinghouse Node*. Environmental Systems Research Institute 2000 User Conference. San Diego.

ODONGO, R. O. & RODRIGUES, A. J. (2007) *Metadata Models and Standards for Kenya National Spatial Data Infrastructure: A Case study of 12 Government Ministries*. Nairobi Kenya, University of Nairobi.

OSEI, S. (n.d.) *The Technical Framework*. NSIF, Department of Land Affairs. [Online] www.nsif.org.za Accessed 2 May 2009.

PAREKH, V., GWO, J.-P. & FININ, T. (2004) *Ontology based Semantic Metadata for Geoscience Data.* International Conference of Information and Knowledge Engineering. Baltimore

PIERKOT, C., MUSTIÈRE, S., RUAS, A. & HAMEURLAIN, A. (2006) *Using Metadata to Help the Integration of Several Multi-source Sets of Updates*. GSDI-9 Santiago, Chile.

RAJABIFARD, A., FEENEY, M. F. & WILLIAMSON, I. (2003) Spatial Data Infrastructures: Concepts, Nature and SDI Hierarchy. IN WILLIAMSON, I. P., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

RAJABIFARD, A. & WILLIAMSON, I. P. (2001) *Spatial Data Infrastructures: Concept, SDI Hierarchy and Future Directions* GEOMATICS'80 Conference,. Tehran, Iran

RIALL, R. L., MARINCIONI, F. & LIGHTSOM, L. F. (2004) *Content Metadata Standards for Marine Science: A Case Study* USGS.

SANGA, N., BIRNIEA, R. V., GEDDESA, A., BAYFIELDB, N. G., MIDGLEYC, J. L., SHUCKSMITHD, D. M. & ELSTONE, D. (2005) 'Improving the rural data infrastructure: the problem of addressable spatial units in a rural context'. *Land Use Policy,* 22 175–186.

STVILIA, B. & GASSER, L. (2008) 'Value-based metadata quality assessment'. *Library & Information Science Research* 30 67-74.

TAYLOR, M. (2004) Metadata: Describing geospatial data. IN NEBERT, D. (Ed.) *Developing Spatial Data Infrastructures: The SDI Cookbook* Global Spatial Data Infrastructure(GSDI).

THABETHE, A. (2008a) SDI Regulation Workshop. Durban, National Spatial Information Framework.

THABETHE, A. (2008b) The South African Spatial Data Infrastructure (SASDI): *Where are we since 1997. " The SDI that is need driven. The SDI in the South African Context"?* GSDI-10. St. Augustine, Trinidad.

THELLUFSEN, C., RAJABIFARD, A., ENEMARKA, S. & WILLIAMSON, I. (2008) 'Awareness as a foundation for developing effective spatial data infrastructures'. *Land Use Policy,* 665, 1-8.

THOMPSON, B., WARNEST, M. & CHIPCHASE, C. (2003) State SDI Development: A Victorian Perspective. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality* London, Taylor & Francis.

TING, L. (2003) Sustainable Development, the Place for SDIs, and the Potential of E-Government. IN WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality* London, Taylor & Francis.

TOPP, L., BARKER, B. & DEGENHARDT, L. (2004) 'The external validity of results derived from ecstasy users recruited using purposive sampling strategies'. *Drug and Alcohol Dependence* 73, 33-40.

TULLOCH, D. & FULD, J. (2001) 'Exploring County-level Production of Framework Data: Analysis of the National Framework' Data Survey. *Urban and Regional Information Systems Association,* 13, 11-21.

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2004a) *Data Policy. SDI-Africa: An Implementation Guideline.* Economic Commission for Africa. [Online] http://geoinfo.uneca.org/sdiafrica Accessed 6 November 2008

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2004b) *Getting Started. SDI-Africa: An Implementation Guideline.* Economic Commission for Africa. [Online] http://geoinfo.uneca.org/sdiafrica Accessed 6 November 2008

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2004c) *Funding Mechanisms for SDI Implementations in Developing Countries. SDI-Africa: An Implementation Guideline.* Economic Commission for Africa. [Online] http://geoinfo.uneca.org/sdiafrica Accessed 6 November 2008

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2004d) *Organisational Structure and Institutional Arrangements. SDI-Africa: An Implementation Guideline.* Economic Commission for Africa. [Online] http://geoinfo.uneca.org/sdiafrica Accessed 6 November 2008

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2004e) *Spatial Data Infrastructure (SDI) Explained. SDI-Africa: An Implementation Guideline.* Economic Commission for Africa. [Online] http://geoinfo.uneca.org/sdiafrica Accessed 6 November 2008

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA (UNECA) (2007) *Geoinformation in Socio-Economic Development: Determination of Fundamental Datasets for Africa by the Economic Commission for Africa*. Addis Ababa, Ethiopia, United Nations Economic Commission for Africa.

VERMEIJ, B. (2001) *Implementing European Metadata Using ArcCatalog*. The Netherlands, ESRI. [Online] www.esri.com/news/arcuser/0701/metadata Accessed 20 October 2008.

WARNEST, M., RAJABIFARD, A. & WILLIAMSON, I. (2005) *Collaborative Approach to Building National SDI in Federated State Systems: Case Study of Australia.* From Pharaohs to Geoinformatics GSDI-8. Cairo, Egypt.

WAUGH, A. (1998) 'Specifying metadata standards for metadata tool configuration'. *Computer Networks and ISDN Systems* 30, 23-32.

WAYNE, L. (2005a) Institutionalize Metadata Before it Institutionalizes You. GeoMaxim/Federal Geographic Data Committee.

WAYNE, L. (2005b) *Metadata in Action: Expanding the Utility of Geospatial Metadata*. GISPlanet.

WILLIAMSON, I. (2003) SDI: Setting the scene. IN WILLIAMSON, I. P., RAJABIFARD, A. & FEENEY, M. F. (Eds.) *Developing Spatial Data Infrastructures: from concept to reality.* London, Taylor & Francis.

WILLIAMSON, I., RAJABIFARD, A. & FEENEY, M. F. (Eds.) (2003) *Developing Spatial Data Infrastructures: from concept to reality,* London, Taylor & Francis.

WILLIAMSON, I. P. (2002) *Land Administration and Spatial Data Infrastructures – Trends and Developments*. FIG XXII International Congress. Washington, D.C. USA.

WOLDAI, T. (2002) *Geospatial Data Infrastructure: The Problem of Developing Metadata for Geoinformation in Africa.* Enschede, International Institute for Geo-information Science & Earth Observation (ITC).

WYGISC (n.d.) *Metadata Education Project*. University of Wyoming.[Online] wwwsdvc.uwyo.edu/metadata/what Accessed 16 October 2008.

ANNEXES

ANNEX 1

QUESTIONNAIRE SURVEY: DATA PRODUCERS

Dear Data Producer/Custodian

My name is Judith Alford, a student (student no. 201502636) at the Centre of Environment, Agriculture and Development (CEAD) at the University of KwaZulu-Natal, Pietermaritzburg Campus. I am conducting research for my Masters degree in Land Information Management entitled "**Challenges Facing Metadata Use by the Spatial Data Community in South Africa**".

My study aims to assess the problems encountered by producers and custodians of spatial data in developing, publishing and maintaining metadata.

I kindly ask you to complete the short questionnaire below to assist me with this task. I wish to emphasize that your participation will be greatly appreciated and that the information you provide will be used for the purpose of this research only in line with the policy of the University of KwaZulu-Natal.

I thank you in advance,

Judith Alford

QUESTIONS

1. **When do you update your metadata, database?**

- Annually:
- Every 6 months:
- Every 3 months:
- More frequently:
- Specify the last year updated if regular/irregular:

2. **What metadata standard does your organization employ?**

- FGDC standard, Content Standard for Digital Geospatial Metadata (specify version):
- ISO International Organization for Standards (specify version):
- In-house developed standard:
- *Specify other:*

3. **Do you have skilled human resources to maintain your metadata?**

- Yes:
- No:

4. **What qualification does the metadata development & maintenance officer hold?**

5. **Who are your spatial data users?**

- Decision makers:
- Other institutions/organizations involved in spatial data:
- Commercial users:
- Academic community:
- NGOs:
- Consultants:
- Donors:
- Media:
- *Specify other:*

6. **How do they access your data?**

- Unrestricted access:
- Authorized required:
- Restricted:

7. **What are the access conditions?**

- Free of charge on request at the office:
- With charge when requested:
- Free of charge on website:

- With charge on website:
- Indicate your URL:

8. **Do you have any partnership with other institutions/organizations involved in spatial data sector?**

- Yes:
- No:

9. **If yes, what is the basis of the collaboration?**

- Data capturing:
- Data distribution:
- Metadata creation:
- Metadata maintenance:
- Funding:
- Technology:
- License agreement:
- Technical assistance:
- *Specify other:*

10. **What are the main constraints in metadata?**

Would your organization be interested in receiving the outcomes of this study?

- Yes:
- No:

Your time and assistance in completing these questions is greatly appreciated.

Thank you!

QUESTIONNAIRE SURVEY: DATA USERS

1. **How do you access your data?**
   - Download online after registration:
   - Download online for free without registration:
   - Order from the dataset originator and then download with a charge:
   - Order from the dataset originator and receive as hardcopy for free:
   - Order from the dataset originator and receive as hardcopy with a charge:
   - Specify other:

2. **What main problems do you encounter with acquired data?**
   - No metadata:
   - Incomplete metadata:
   - Non standardized metadata:
   - Data without spatial reference:
   - Data with different spatial references:
   - Specify other:

3. **What are your suggestions for improving metadata?**

Would your organization be interested in receiving the outcomes of this study?

- Yes:
- No:

Your time and assistance in completing these questions is greatly appreciated.

Thank you!

ANNEX 2

ETHEKWINI

Corporate GIS MetaData

Dataset	Description	Contact Name	Contact No	Email	Owner Dept
Dams	Images of dams have been digitised from GIS images (except Nagle Dam which is beyond extent of images)	Felicity Gordon-Mckenzie	031-3117287	gordon-mckenzief@durban.gov.za	Coastal Stormwater & Catchment Management
bw2005	A MrSID mosaic made from aerial photography flown in May 2005.	Trevor Ireland	311 4196	irelandt@durban.gov.za	Corporate GIS
Coastal_Zone	New coastal zone being created. This shapefile is underconstruction.	Irene Strydom	3116363	strydomi@cesu.durban.gov.za	Drainage and Coastal Engineering - Records Section
Coastal_Erosion_Lines	Building control line along sea edge.	Felicity Gordon-McKenzie	031-3117287	gordon-mckenzief@durban.gov.za	Drainage and Coastal Engineering - Records Section
Fibreoptic_Bisasar	This is the (approximate) over and under ground pathway of a fibre-optic link between the Springfield	Danny Ramlahl	031-3024807	DannyRa@dmws.durban.gov.za	Durban Solid Waste

	Office complex and the Bisasar Road, Springfield Landfill Site				
Landfills	The location of all Solid Waste Landfills (Municipal and Private) throughout the Unicity (General and Garden Waste) whether open or closed, rehabilitated or not, and even those that are only planned - as at August 2001. This set rarely changes.	DANNY RAMLAHL	031-3024807	DannyRa@dmws.durban.gov.za	Durban Solid Waste
Recycling_facilities	Facilities where recycling is conducted, buy-back or drop off centres, reflecting what types of materials are accepted at each point. Compiled in Feb 2003; not fully comprehensive - but good coverage generally.	DANNY RAMLAHL	031-3024807	DannyRa@dmws.durban.gov.za	Durban Solid Waste
Roundsplits	The bounded areas served by a (DSW department) waste collection vehicle on a particular day or days of the week (both for domestic and commercial contract purposes), as at August 2001. Due to round balancing exercises in progress during late 2002 / early	DANNY RAMLAHL	031-3024807	DannyRa@dmws.durban.gov.za	Durban Solid Waste
Electricity_AOS	Electricity Area of Supply boundary.	Trevor Ireland / Denis Bodeker	031-3114196	irelandt@cesu.durban.gov.za	Electricity
Moss	A spatial representaion of the environmentally sensitive areas or open space system for eThekwini Municipality. It contains a general and detailed description of the habitats mapped and gives some guidance to the physical developability of the land.	Penny Croucamp	031-300 2517	croucamp@cesu.durban.gov.za	Environmental Management Branch - Development Planning Department
CCTV	Folder containing the cctv network : cameras, ducts, fibres, manholes, cables etc. Special program needed to view in full.	Darryl Thomas	031-3002340	thomasdh@cesu.durban.gov.za	eThekwini Traffic Authority and Electronic

					Services
Sewer_Manholes	Contains the position of Ethekwini Water sewer manholes	Denis Meredith	031-3024672	denisme@dmws.durban.gov.za	Ethekwini Waste Water Services - GIS Section
Sewer Pipes	Contains the position of Ethekwini Water sewer pipes	Denis Meredith	031-3024672	denisme@dmws.durban.gov.za	Ethekwini Waste Water Services - GIS Section
Cp_Cables	Contains the position of all the cathodic protection cables related to steel water mains, NB NB please can this data set be added to all plots to show position of the cables as if they are damaged for any reason this cost the Council thousands of Rands in	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Cp_Pnts	Contains the position of all Cathodic protection test posts, cross bonds, blue bonds and m points.	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Watermains	Contains all the water mains in the Unicity area. Please use AVL legend to view data.the data is collected via all the sources listed.	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Gpsmts	Contains the GPS position of water meters	Cathy Wood	031-3024753	cathyro@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Installations	Contains the GPS position of Ethekwini Water Services Installations for example (offices, depots, pay points, reservoirs, pump stations etc)	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS

					Section
Prvs	Contains the GPS position of Ethekwini Water pressure reducing devices	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Fittings	Contains the position of Ethekwini Water fittings related to water mains for Example (Valves, Ts etc)	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section
Proposed	Contains the position of Ethekwini Water proposed water mains for IMS use only The Proposed water mains are included in the watermains under \\Services\MetroWater\WATERMAINS\watermains.shp	Steve Pietersen	031-3024740	stevepi@dmws.durban.gov.za	Ethekwini Water Services - GIS Section