

The Challenges of Implementing Metadata in Developing Countries: A Case Study for Zimbabwe

David NJIKE and Kuda MUSHATI, Zimbabwe

Key words: Metadata, SDI, Standards, Geospatial data .

SUMMARY

The demand for spatial data is continuously increasing at a faster rate. With the rapid increase in the development of spatial data collection tools such as Global Positioning Systems (GPS) and remote sensing technologies, the amount of spatial data being collected and stored on computer network is becoming vast. Data collection is the most expensive part in any project, which uses spatial data, and it costs about 70% of the total cost of the project. Knowing what data is already available for an area is of great benefit to many spatial data users. This will reduce duplication of efforts and time for data collection. Metadata is descriptive data about data. If metadata is provided for all the data that is collected by different organisations, the cost of data collection will be reduced.

Metadata is an essential and vital vehicle in the economic and sustainable development of any country. For any development there is need for information to be available at the right time and in the right format. Without up-to-date and accurate information, proper decisions cannot be made. Decision-making, at all levels in society is only effective when the appropriate information is available. Metadata is the tool that can be used to improve the efficiency of collecting data by providing users with information about data.

This aim of this paper is to present the status of metadata development in Zimbabwe. The paper looks at the level of use of metadata and the challenges that organisations are facing in implementing metadata. The paper further discusses efforts that are being done to develop metadata in Zimbabwe. It also examines the challenges that Zimbabwe is facing in developing metadata. Finally the paper will discuss the different standards for metadata records that have been developed (nationally, regionally and internationally) and makes some recommendations for Zimbabwe. Recommendations on how to implement metadata in organisations taking into considerations technological and institutional issues are also going to be discussed.

The Challenges of Implementing Metadata in Developing Countries: A Case Study for Zimbabwe

David NJIKE and Kuda MUSHATI, Zimbabwe

1. INTRODUCTION

With the rapid increase in the development of spatial data collection tools, such as Global Positioning Systems (GPS) and Remote Sensing technologies, the amount of spatial data being collected and stored on computer networks is becoming vast. Numerous organisations, government agencies, the private sector and consulting firms create, handle and process vast amounts of data. The key to efficient data capture and transfer is to share and be aware, of what data is available and where and how can it be accessed. This reduces duplication of efforts of data capture which is time consuming and very costly. Descriptive data about the data sets are maintained as Metadata to provide such information to spatial data users.

Designing and developing Metadata in an organisation organises and maintains internal investment in the data. As personnel change or time passes, information about an organisations data will be lost and data may lose its value, this poses a great challenge in that later workers would have little understanding of the content and uses for digital dataset and may find it that they cannot trust results generated from these dataset. Geographic Information Systems (GIS) application normally require many data themes, from different sources and not all organisation are capable of creating all the data they need. Information about data should always accompany the transfer of a geospatial dataset because this will assist the receiving organisation to process and interpret data.

2. BASIC METADATA CONCEPTS

2.1 What is metadata?

Metadata is commonly defined as ‘data about data’. This is a set of data that describes a data set. It is used to describe the characteristics of datasets. This description is in a ‘structured’ manner. This concept is similar to that of libraries, where there are indexes that describe the information about books and their location within the library. Characteristics like the custodian, name of dataset, geographic extent of the data, currency of the data, storage format, data quality, contact information to inquire about the dataset are all described. Metadata is extremely important for spatial data as it allows a potential user of a dataset to determine whether the dataset is useful to them or not. Metadata systems can be established that allow users to search the Metadata records for the datasets located on a network. From the results they are able to determine if there are any datasets that may be of interest to them, how to gain access to them, any constraints on using them, etc. Metadata are therefore extremely important as they facilitate the more efficient use of spatial data. This is achieved by allowing potential users of spatial data to search for datasets that may suit their needs. They can look at the Metadata record for a dataset and see if it meets the criteria for use that they have set. The record will also tell the searcher the access rights/constraints of the dataset.

All this is very important, as it is usually cheaper to purchase that spatial dataset from another party that has produced it for another purpose than it is to reproduce the dataset oneself.

2.2 Importance of Metadata

The key to any decision-making process is the availability of up-to-date and accurate information. The information should be readily accessible. This is vital if sustainable development is to be realised. The importance of metadata in any organisation can be summarised as below:

- Protects investment in data. Agencies are not vulnerable to losing all the knowledge about their data when key employees retire or accept other jobs.
- Helps a user to understand data.
- Facilitates the browsing and transfer of data.
- Enables data discovery.
- Supports the creation of an inventory of the data holdings
- Supports the creation of a spatial data clearinghouse, where users can search to find the data they need.
- Metadata ensures that potential data users can make an informed decision about whether data are appropriate for the intended use.
- Metadata also ensures that the data holdings of an agency are well documented.
- Metadata can also be an effective marketing tool, which can increase traffic on websites as well as awareness of available resources.

2.3 Need for Metadata

The rapid advancement in spatial data capture technologies such as the Global Positioning System (GPS), satellite imaging and total stations have all made the capture of digital spatial data a relatively quick and easy process. This has meant that in the last decade or so the amount of digital spatial data in existence has increased significantly. There is now a vast amount of spatial data, stored by numerous organisations at various locations across the globe. Much of these data are not being used as effectively as they should. Referring to the LANDSAT images, the Vice President of the USA said:

In spite of the great need for the information, the vast majority of those images have never fired a single neuron in a single human brain. Instead, they are stored in electronic silos of data". (Gore 1998)

Recently there has been a greater focus on how to use the spatial data that are collected and stored in the expansive Geographic Information Systems (GISs) to their full potential. Metadata is a tool that can be used to effectively and efficiently access and understand spatial data.

2.4 Metadata Engine

A Metadata engine is an application that is used by database management systems (DBMS) to extract and display the results to a user's query. They work by parsing the query and then consulting the data dictionary for the database, which contains Metadata that outlines the

internal structure of the database. The Metadata engine works completely in the background with no direct interaction with the user of the database. The user of the database does not even have to know that the engine exists. All the user is concerned with is writing the query and then getting the right results returned. Metadata engines should not be confused with Metadata systems. A Metadata system is very similar to a search engine. They allow a user to search Metadata records, which have been produced to describe the characteristics of a dataset, and determine whether they wish to gain access to the dataset. Data directories and clearinghouses both use Metadata systems to allow users to search them. They both contain databases that hold the individual Metadata records for each dataset that is available. These databases are searched by keyword, geographical location, date, etc. and return the individual Metadata records to the user for them to view. By viewing these Metadata records the user is able to determine whether the dataset is of use to them, whether it meets their accuracy requirements, any access constraints, who to contact to gain access to the dataset, etc. In the case of a clearinghouse there is also the capability to download the dataset online. However there is no capability to query the dataset online, whereas a Metadata engine has this capability. At the present time there appears to be no true Metadata engines in existence that allow the distributed processing of spatial data over the WWW. Distributed processing is the term used when a distributed database is set up that allows the querying of autonomous databases that are located over a network. To the user of a system that allowed distributed processing of spatial data it would appear as if they were just using one integrated database. It should be transparent to them that the data that returns after they submit a query is actually returning from possibly two or more autonomous databases.

2.5 Metadata Standards

The Standard provides a common set of terminology and definitions for the documentation of geospatial data, including data elements. The objectives of the standard are to provide a common set of terminology and definitions for the documentation of digital geospatial data. The standard establishes the names of data elements and compound elements (groups of data elements) to be used for these purposes, the definitions of these compound elements and data elements, and information about the values that are to be provided for the data elements.

Several metadata standards have been developed at national, regional and international level. Some of the standards are the Dublin Core, Australian and New Zealand Land Information Council (ANZLIC), Federal Geographic Data Committee (FGDC) and ISO 15046-15. International Organisation of Standardisation (ISO) is working on a metadata standard ISO 15046-15 which can be adopted by any nation.

The FGDC standards specifies the elements needed to support three major uses of metadata:

- To maintain an organisation internal investment in spatial data.
- To provide information to data clearinghouses and
- To provide information needed to process and interpret data transferred from another organisation.

3. METADATA IN ZIMBABWE

The purpose of this research was to evaluate the status of Metadata in Zimbabwe. Data was collected from spatial data handling organisations by means of questionnaires, interviews, perusal of records and attending of workshops. The response from these organisations was very high. Most of the organisations were willing and ready to share their information. Over thirty organisations including the public and private sector, were interviewed.

3.1 A Sample of questions and responses from questionnaires

Question	Response	
Does your organization deal with spatial data?	Yes = 23	No = 7
Do you know anything about Metadata?	Yes = 11	No = 19
Do you think Metadata is important?	Yes = 10	No = 20
Do you have any Metadata repositories in your organization?	Yes = 5	No = 6
In what form is your Metadata?	Digital = 3	Analogue = 8
Do you know anything about SDI?	Yes = 11	No = 19
Are you a member of ZSDI?	Yes = 7	No = 23
Does your organization share data with other organizations?	Yes = 26	No = 4
What do you think are limiting factors in creating Metadata?		
i. Lack of resources (hardware and software).	11	
ii. Committed to other projects.	19	
iii. Lack of expertise.	19	
iv. Metadata not necessary.	7	
v. No support from the decision makers.	18	
vi. Institutional barriers.	22	
vii. No financial support.	16	

3.2 Metadata Development in Zimbabwe

Metadata systems have different degrees of usage and are at different stages of development in Zimbabwe. Very few organisations within the country have developed more integrated metadata systems. There is nothing in the public sector and a few from the private especially Non-Governmental Organisations (NGOs). One overriding factor has to do with priorities in the development of metadata systems. Most organisations were concerned more at getting their data and fulfilling their mandates.

It can be said that the full potential of metadata systems has not been fully realized in Zimbabwe due to the prevailing state of technological, political and economical development in most organizations. The ability of these systems is to account for spatial data. The main problem in most cases lies in the necessity to mobilize personnel and financial resources. This is particularly obvious in administrations for which the introduction of a metadata system is not a political priority.

From the results only 17% of the organisations interviewed have somekind of metadata records, which at the moment are just used as an internal spatial data management system. However the organisations that have metadata records are NGOs who have less data to share with the spatial data industry.

3.3 Spatial Data sharing

There is a problem of data sharing among different organisations. The most fundamental problems in data sharing are data accuracy, the scale used in digitising, the type of data classification systems used and the frequency in data updating (which is not normally known) The accuracy of data is measured in terms of quality, geodetic accuracy and the quality of data to be input by the receiving organisation. The scale used in digitising affects the resolution of the geographic data collected. It will be difficult to perform map overlay analysis if maps and images entered for different GIS are at different scales and resolutions. Data are often difficult to share among organisations because of differences in classification systems used in classifying and lack of metadata. Some organisations considered their data to confidential and sensitive and therefore are not in a position to share. Some policies in other organisations prohibits data sharing.

3.4 Data Transfer

The explosion in the collection and use of regular geographic data has led to the existence of many database which are of interests to a range of organisations in addition to that of own spatial data collection. To avoid unnecessary duplication of data collection there is need for a widespread data interchange between organisations with common information needs. In the current system the main obstacles to free data interchange are in reality more often administrative than technical with organisation reluctant to release their data. This may be for a variety of reasons. Technical issues, if they do arise, are relatively easy to deal with than institutional ones. Issues like incompatibility of systems can be solved much easier than issues like data security and organisational policies. Institutional issues cannot be solved in a day, because they involve a lot of people and a lot of consultations should be made before they are ammended. This can inhibit the transfer of data within and across organisations. The issue of 'empire building' has resulted in lack of data transfer within and among organisations.

The current practice of developing individual physical data and individual different data documentation schemes that are only useful to particular organisations works as long as the source data and target data retain the same documentation schemes. The existence of large quantity of desperate spatial data hinders organisations from exploiting their spatial data to their full advantage. There is too much spatial data and not enough metadata which is crippling many organisations in their effective use of spatial data.

Most organisations have large hidden data resource they could draw on if they only knew it existed. They need a way to explore these data to find their true meaning content and use

them to full advantage. The research however revealed that potential to share data in Zimbabwe is high as most of the organisations are now willing to share their data.

The major technical obstacles to data transfer reside in the lack of application of a national standard for digital spatial data, and the almost total absence of Metadata. Data documentation is generally poor. However, without proper Metadata, data sharing will always be a difficult and time-consuming enterprise. Efforts are underway to develop our own national standards. This effort is being coordinated by the Department of the Surveyor General with the assistance of all the stakeholders.

Efforts to harmonise the classification schemes used in different institutions (e.g. vegetation cover maps) have failed so far. As a result, each institution organises its data according to its own beliefs and knowledge resulting in mostly incompatible data structures and classification schemes.

3.5 Networking

The opportunity to share experiences in metadata development and application is a great stimulus for future development and to ensuring its effectiveness and efficiency. There is need to focus attention to the interchange of information and experience both within and between organisations. Such networking should have both formal and informal mechanisms. In Zimbabwe this type of networking of metadata development is done by an informal initiative committee (SDI Steering committee). This is a voluntary organisation with its activities sustained by external funding. The committee has a working group which looks at metadata issues and collects metadata records. A few organisations have send thier records to this committee. This committee has the following terms of reference:

- Develop an action plan for metadata activities
- Establish a clearinghouse node and serve metadata through this node.
- Develop a website for metadata activities on the ZSDI website.
- Participate in the annual ZSDI workshop
- Create an e-mail forum to discuss metadata development issues in the country
- Encourage the participation of more players by recruiting at least 1 organisation in every 2 months.
- Develop University-Industry research programs to allow students to work on metadata projects

3.6 Technology

The concept of Metadata according to the FGDC standards, can be used in a clearinghouse nodes, few spatial data handling organisations in Zimbabwe are fully computerised, they only use them for typing and if at all connected to the internet which is mainly used for email services. The economic situation in the country is not that good for many organisations, which means they will only access free data from the clearinghouse. Telecommunication facilities have always been a problem in Zimbabwe. The advent of Internet has played a

major role in promoting information exchange worldwide. However, due to telecommunication problems, the few institutions that have Internet or e-mail facilities complain that the facilities breakdown too often or that the services are too slow.

3.7 Responding to the need for spatial data infrastructures

In Zimbabwe, spatial data infrastructure steering committee was founded in April 2003 as a voluntary organisation. This committee stands as a key grouping for institutional development, introduction of technology change and competitiveness, focusing in the promotion of Metadata and clearinghouse development for the Zimbabwean spatial data infrastructure. Created a metadata working group intended to look at the metadata development issues. In spite of growing awareness regarding the importance of spatial data infrastructure, institutional participation in the ZSDI is still low.. The committee has managed to hold trainings and workshops on SDI and metadata issues. Meetings are done on a monthly basis and the number of participating organisations is increasing but at rather slow pace. Most of the organisations do not attend consistently with just around 10 organisations sending representatives every month. The committee is working towards its formalization.

4. FINDINGS

The following findings were made during the research:

4.1 Metadata development situation

Metadata development is still in the initiation stage, thus national metadata system does not exist in Zimbabwe. There are very few organisations which are maintaining a proper metadata in the country.

4.2 Metadata standards

There are no metadata standards in the country and organisations are using standards from other nations. The few organisations which have metadata records use the FGDC standard.

4.3 Metadata repository

It was discovered that organisation did not have any tradition of keeping information about their datasets some still used analogous means of recording information about spatial data in small booklets. However it was also discovered that few organisations created their metadata in digital and analogue formats.

4.4 Resources

Most of the organisations lacked resources to implement metadata recording systems. Some organisations have got potential, expertise and resources that permit the implementation of

metadata systems but lack of support from the top management. However it was discovered that soon many organisations were considering implementing metadata systems.

4.5 Education and training

The education institutions do not have any modules which cover metadata issues. Most of the students from these institutions do not have a sound knowledge on metadata issues. There are no trainings in terms of short courses, workshops or seminars that are offered that cover metadata issue except that which is done annually by the ZSDI.

4.6 The factors affecting metadata development in Zimbabwe.

The following were found to be the main factors affecting metadata development in Zimbabwe:

- Lack of expertise
- Lack of resources
- Committed to other pressing projects
- Top management and politicians does not support metadata development initiatives
- Lack of appreciation
- Metadata not part of work mandate in Zimbabwe

5. RECOMMENDATIONS

A number of recommendations can be made from the findings of this research.

- The SDI Metadata working group should increase awareness and understanding of the vision, concepts and benefits of developing metadata system through research and education.
- The SDI should consider the creation of a clearinghouse node so as to make metadata development a worth move by organizations.
- The SDI Metadata working group should identify and promote the attitudes and actions that help to develop Metadata in the Zimbabwean SDI.
- The institutions of higher education should consider training and conduct research on Metadata since in Zimbabwe there are no enough human resources in Spatial Data Infrastructure. They should also review their programmes and include topics that cover metadata issues.
- The SDI Metadata working group should come up with a National metadata policy: This should be a national metadata policy on the standardization of data entered in the metadata systems to ensure good data sharing and compatibility of data collected and to foster for metadata development. ZSDI members and all stakeholders should contribute to this metadata policy.

- ZSDI should be formalised and become recognised by as a legal framework.
- ZSDI invite all actors involved in the development of local, regional and global SDI projects.

6. CONCLUSIONS

Metadata development appears to be the dream of many spatial data handling organizations in Zimbabwe. However cost benefit analysis needs to be looked at. The goodwill and potential is there for the Metadata development by spatial data handling organization in Zimbabwe. The greatest challenges are the lack of financial resources, sound SDI, commitment from management, and decision makers and institutional barriers. The institutions, which were surveyed, expressed very positive commitments and goodwill to ensure that they produce Metadata for their datasets. The positive aspects are that many people expressed in a lot of interest in learning more about metadata. This was demonstrated by the way people volunteered in participating in the several meetings that have taken place so far.

REFERENCES

1. Best practice of EIS- the case study of Zimbabwe.
2. Contents standards for digital geospatial metadata workbook version 2.0 fgdc 200:
<http://www.scs.noaa.gov/metadata/index.intergraph.co.gis/smms>
3. Generating Metadata: saving money in the organizations :
<http://.anzlic.org.au/infrastructure ASDI.html>
4. Michael H. B. (1994) Data sharing using a common Architecture:
<http://www.gridano/els-ssa/products/Zimbabwe/zimeao08.html>
5. Michelle L A (2003) Development of GSDI Clearinghouse nodes
6. On the problem of finding the geographic data we are looking for:
<http://www.lmu.jrc.it/workshops/9ec-gis/papers/tabs-cantan.pdf>
7. Paul F. S. NSIF: <http://www.sadc-four.org.zw/rrsu/clrnhouse/workingsouthafrica>
8. Smith R. 1999, A consolidated effort of sharing Metadata in South Africa

BIOGRAPHICAL NOTES

Lecturer and Head of Department: Department of Surveying and Geomatics

Area of interests: GIS, SDI, Remote Sensing and Photogrammetry

Membership to societies: Survey Institute of Zimbabwe (SIZ) - Chairman of Education and Change committee, ZSDI – Chairman of the Education working group

Courses taught: GIS I-III, Surveying I-III, Remote Sensing, Photogrammetry I-III

Research in progress: Use of satellite imagery for map production and revision in Zimbabwe.

Conference Attended: 7th Science and Technology Symposium 1-3 September 2004, Harare International Centre.

CONTACTS

Mr. David Njike
Midlands State University
Department of Surveying and Geomatics
Private Bag 9055
Gweru
ZIMBABWE
Tel. +263 54 60464 or +263 91 238678 (mobile)
Fax + 263 54 60233
Email: davidnjike@yahoo.co.uk

Mr. Kuda Mushati
Midlands State University
Department of Surveying and Geomatics
Private Bag 9055
Gweru
ZIMBABWE
Tel. +263 54 60464
Fax + 263 54 60233
Email: kmushati@yahoo.com