Heterogeneous user requirements on GI standards – Implementation strategies in an Ethiopian perspective

Anders Östman*, Mehari Alemayehu, Eyob Teshome****

* Swedesurvey AB, SE-80182 Gävle, Sweden.
** Prime Consultants PLC, Addis Ababa, Ethiopia

Abstract. By tradition, national GI standards tend to be decided at a high level and implemented in a top-down approach. Such an approach has its advantages from a decision making point of view. However, experiences from several countries around the world also report severe problems in the implementation phase. Some of the problems are related to changing requirements. One such example is the metadata standards, where for instance regional governments in Sweden have faced five different metadata standards during the last 10 years. Other problems are related to the limited testing of the standards. The specifications developed by the OGC as well as the INSPIRE implementation are exceptions to this, although neither are actually developing formal standards. As a consequence, the benefits the standards are expected to provide may be difficult or costly to realize in practice. A third problem is caused by the heterogeneity of user environments, where the standards are to be implemented. In many African countries, the GI maturity varies a lot between different authorities and organizations, even within the same country. Specifying GI standards that are meaningful for such diversity of organizations is a challenge.

Swedesurvey has, in cooperation with their partner Prime Consultant Inc, specified an implementation strategy for GI standards in Ethiopia. The scope of the standard and guidelines being specified relates to metadata, database, spatial data, web services and GIS business strategies. The initial user survey showed clearly the heterogeneity within the potential target group, with respect to GI maturity, SDI readiness and level of IT maturity. Being well aware of experiences from other countries, a strategy for implementing GI standards was outlined, which address the problems of limited testing of standards as well as the heterogeneity within the user community. The key factor to consider when specifying this strategy was the benefits for user having various levels of GI maturity.

The problem of different levels of maturity within the target group is a problem that is less recognized within the GI standardization domain. Within IT industry there are however several different stack-based approaches for gradually complying with increasing complexity of requirements. The stack-based model here presented also has the advantage of being directly linked to benefits. In practice it means that less mature organizations can focus on making their data digital while more mature organizations can focus on sharing their information according to standardized XML/GML schemas. All this is achieved within one single model of standards. As a consequence, this also allows each individual organization to grow in maturity in their own speed.

In recent years, the question of linked data and RDF structures has become a part of the standardization discussions. Linked data may be seen as one way of achieving organizational interoperability. The upper levels of the 5-star model deal with linked data and associated standards such as RDF. As a consequence, the proposed strategy for the implementation of GI standards in Ethiopia is also scalable to such level of complexity. A study was also conducted in order to understand the level of GIS knowledge and application in the participating institutions. A Capacity Maturity Model (CMM) was applied to categorize them into five classes, each one addressing a certain standard compliance level. But for the time being, the focus is, as well as in other places such as Europe, on the lower levels and on having sustainable standards, agreements and institutional frameworks for data sharing.

Keywords: GI standards, heterogeneity, 5-star model

1. Introduction

Geographic Information/Spatial data plays a significant role in delivering effective government services, informed decision makings and creating business opportunities. At the same time, the capacity to meet such user needs and to deliver services and tools within the spatial information community has gone far beyond the ability of single organizations. That is, geographic information has transitioned from being the essence of national mapping organizations like the Ethiopian Mapping Agency (EMA), to being produced and utilized in different organizations and being the common commodity of consumers. Consequently, organizations move towards sharing spatial datasets and collecting different spatial data from different sources.

The diversity of methods taken by different stakeholders and providers make multi-sourced spatial data inconsistent, where most spatial data sets

do not comply easily and shared smoothly. In many cases, this leads to weak collaboration and cooperation among different stakeholders which hinders effective spatial data integration. Hence, without the development of a comprehensive standard and guideline effective integration and sharing of spatial data cannot be achieved

Consequently, the Ministry of Communication and Information Technology and the Information Networks Security Agency (INSA) have taken the initiative for the development of the National GIS Standard and Guidelines as the first step to bring on board the various stakeholders by assessing the level and usage of GI in the country and addressing international best practices and interests of the client and stakeholders.

1.1. Objectives

The goals of this project are to develop a set of standards and guidelines for coordinating the GI community to share their geo-data effectively, efficiently, timely and with minimum cost. They are intended to serve as a shared foundation, encouraging improved communication and collaboration amongst stakeholders and the community at large. It is also anticipated that these standards and guidelines will be updated as more and more organizations adopt its operating procedures for proper production, utilization and sharing of geo-data.

The objectives of this paper are to describe the current state of GIS maturity in Ethiopia, to describe the proposed guidelines and to describe the implementation roadmap.

1.2. International experiences

Many other standardization initiatives in the GIS domain only identify which standards to be implemented and leave the implementation to the stakeholders, without any guidance. The main contribution of this study is that it goes two steps further, by also identifying and analyzing appropriate guidelines for the implementation of the standards and also by proposing a governance structure for the implementation. The latter is absolutely crucial in order to reach the goal of improved cooperation among governmental agencies and other stakeholders. It should however be stressed that the guidelines, standards as well as the proposed governance structure are not tested in the Ethiopian context. As a consequence, modifications are probably needed during the implementation phase.

The scope of the Ethiopian GI standardization should also be seen in an international context. The most well known initiatives are the ones in Europe (European Commission, 2011), United States (FGDC, 2011) and Australia – New Zeeland (Anzlic, 2011). According to study by Makanga and

Smit (2010), the formal NSDI activity in most African countries is in general still in its infancy. As an example they mention that in 2003 there were two national SDI clearinghouses in Africa. Five years later there were three, but in 2010 two of them have ceased to be operational. As a consequence, the international outlook will concentrate on USA, Europe and Australia – New Zeeland.

In order to participate in and contribute to the national and international development within the GI sector, the long term goal is to implement the international standards of the ISO-19100 series (ISO, 2011) and industry standards such as the OGC specifications (OGC, 2011) and W3C standards (W3C, 2011). The strategies for implementing these standards differ however across the world. Europe has for instance launched the INSPIRE directive, aiming at the creation of a spatial data infrastructure (SDI) in Europe (European Commission, 2011). The directive entered into force in 2007 and the implementation phase is planned to be finalized in 2019. The implementation covers many different standards such as metadata, network services and pan-European data specifications. Australia has taken a simpler approach and concentrates mainly on metadata and search facilities (ASDI, 2011 and Office of Spatial Data Management, 2011). A similar approach has also been adopted in the USA (FGDC, 2011).

A long implementation phase has a large financial risk. In general, the expenses should arise as late as possible in the implementation, while the benefits should come as early as possible. The implementation must also be able to adapt to changing requirements and technological developments. As a consequence, a strategy similar to the agile development strategies is sought for. In this way, early benefits are sought, improving the overall cost-efficiency of the implementation

The problem of different levels of maturity within the target group is a problem that is less recognized within the GI standardization domain. Within IT industry there are however several different stack-based approaches for gradually complying with increasing complexity of requirements. One such model is the 5-star model of Tim Berners-Lee. Other successful stack-oriented models are the web service stack model of the World Wide Web Consortium (W3C) and the ISO/OSI model for data communication. The stack-based model here being developed in this project also has the advantage of being directly linked to benefits. In practice it means that less mature organizations can focus on making their data digital while more mature organizations can focus on sharing their information according to standardized XML/GML schemas.

1.3. Tim Berners-Lee 5-star model

One theoretical framework for an agile approach to implementation of GI standards is the 5-star model developed by Berners-Lee (2006). This approach has also been used by Lopez-Pellicer et.al (2011) in SDI development. The model is based on a step-wise improvement of the data, aiming at increased sharing and linking of data sets. The model consists of the following levels (Inkdroid, 2011)

1. Make your data available on the web (whatever format)
2. Make your data available as structured data (e.g excel instead of image scan of a table)
3. Use non-proprietary formats (e.g csv instead of excel)
4. Use URI's to identify things, so that others can link to your data
5. Link your data to other people's data to provide context

The first three levels concerns data sharing, while the remaining two relates to institutional cooperation. The latter require also data harmonization, a procedure not well recognized in the 5-star model. In addition, the 5-star model assumes a communication based on the Internet. In the Ethiopian context, such communication can not always be assumed. As a consequence, an entry level based on other types of communication tools, such as ordinary postal mail, must also be considered.

Considering these limitations, the methodology for producing the implementation roadmap of GI standards is based on a modified version of the Berners-Lee 5-star model, taken into account the data harmonization requirements for institutional cooperation and the limitations in Internet access.

2. Method

2.1. Assessment of current state

Assessment of the existing GIS practices in stakeholder organizations in the country is required in order to draft a realistic implementation plan. By assessing the current maturity of GIS among the stakeholders, the strengths and weaknesses and knowledge about what to do to reach to the next level of maturity has been identified. The assessment was made using questionnaires.

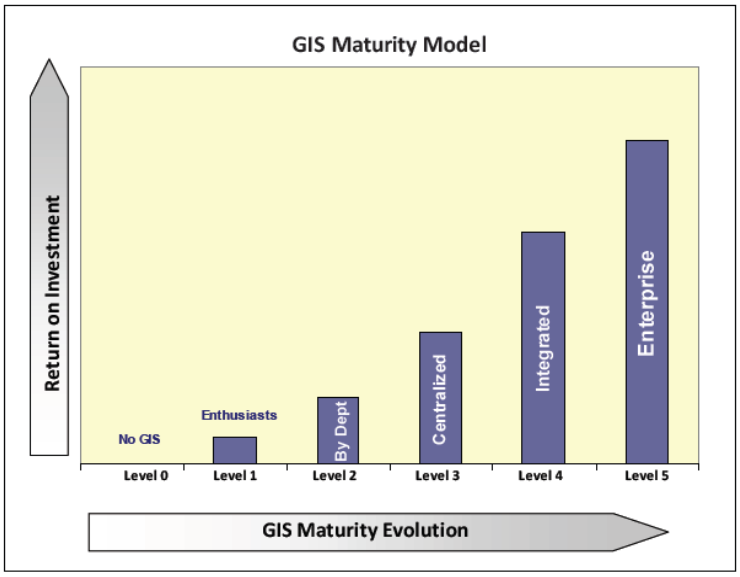

Figure 1. GIS Maturity Model

The As-Is study has focused on five major thematic subjects (metadata, spatial data, database, GIS web services and GIS business) to charaterize the GIS maturity levels of the participating organizations see *Figure 1*.

The GIS Maturity Model is based on a Capability Maturity Model (CMM) approach. This approach broadly refers to a process improvement approach that is based on a process model. This process model is a structured collection of practices that describe the characteristics of effective processes. The practices included are those proven by experience to be effective.

The CMM techniques implement appropriate scaling definitions and scoring logics to be used to classify the maturity into five levels (Enthusiast, Departmental, Centralized, Integrated and Enterprise level) according to its standardization of processes in the subject area being assessed.

2.2. Procedures for developing guidelines for producing data specifications, metadata profiles and web service specifications

The guidelines are aimed for the development of data specifications, the specification of a metadata profile and the specification of web services.

The main international experience of producing data specifications has been obtained within the INSPIRE development. A methodology for pro-

ducing data specifications has also been issued (European Commission, 2008). There are several experiences being drawn from this work, also reported during several INSPIRE conferences. In general, the result is that the procedures specified in the INSPIRE documentations works quite well, at least within a European context. The method applied for developing Ethiopian data specifications is to adapt the INSPIRE methodology to Ethiopian context, by specifying recommendations and instructions how to read and understand the INSPIRE method.

In contrast to guidelines for producing data specifications, there are a large number of metadata profiles being published. The most prominent ones are the one by INSPIRE (European Commission, 2010), the North American Profile - NAP (FGDC, 2011) and the one from Australia and New Zeeland (Anzlic, 2001). The method for developing the Ethiopian metadata profile is then based on a comparison and evaluation of these three major approaches.

Implementing rules for network services has been published in Europe (European Commission, 2011). These rules are based on the OGC specifications of web services, like Web Mapping Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Coordinate Transformation Services (CTS) and Catalogue Services for the Web (CSW). Similar initiatives have not been found in North America or Australia – New Zeeland. The European specifications do however extend the OGC specifications in some aspects, resulting in an uncertainty of a wide software support for these extensions. As a consequence, the main candidates for web service specifications were the OGC specifications, without extensions. However, the INSPIRE extensions will also be studied and evaluated, in order to assure their suitability to the Ethiopian context.

3. Results

3.1. Metadata guidelines

Metadata is usually defined as “data about data”. Metadata may serve several purposes such as data discovery, data assessment (determination of fitness for use), data access, data use, data transfer and data management (Anzlic, 2001).

One objective of this paper is to propose a draft metadata profile for Ethiopia. A metadata profile is here considered to be a subset of an international metadata standard, specifically adapted to the Ethiopian context. The profile may also include extensions as compared to the international standard.

The international standards that are of interest are the ISO TC211 standard suite and the Dublin Core metadata standard.

The as-is survey showed that the overall knowledge and application of metadata was below average (1.90) at country level, see *Figure 2*. Moreover, 75 % of the participants have confirmed that their organizations are at a very early stage of either planning or introducing of metadata services. Mere 3% of the participants have been found to have a proper knowledge of metadata services and are preparing to fully implement the scheme. MoWE, EMA, AAU-ES and INSA are found to be in progress of realizing appropriate metadata services in their organization.

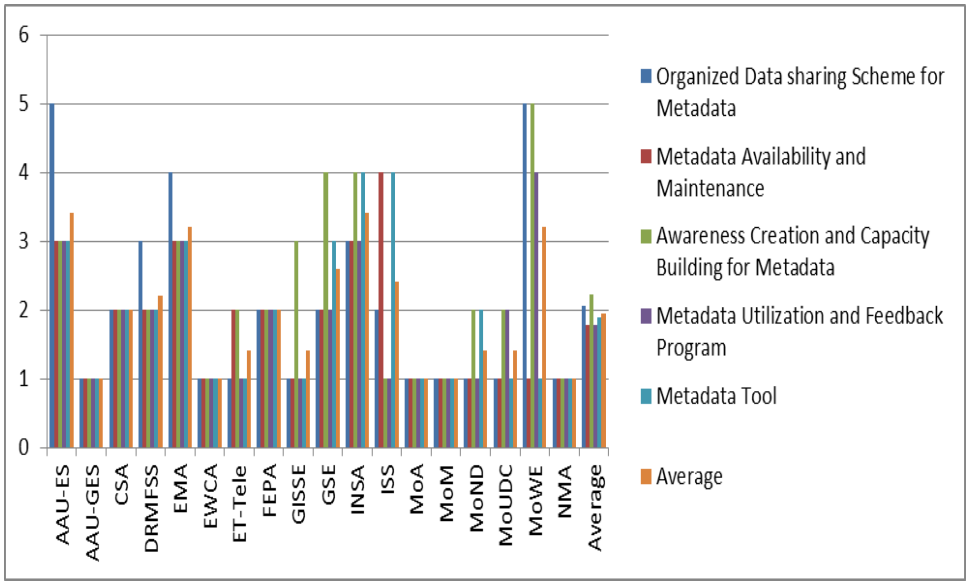

Figure 2. Assessment of current state of metadata in Ethiopia

The standards in the ISO TC211 standard suite that are relevant to this task are the ISO 19115 metadata standard, the ISO 19119 standard on services and the ISO 19139 Metadata XML schema implementation. The ISO TC211 suite of standards is specifically made for geographic information, in contrast to Dublin Core which is more general.

The Dublin Core set of metadata elements provide a small and fundamental group of text elements through which most resources can be described and catalogued. Using only 15 base text fields, a Dublin Core metadata record can describe physical resources such as books, digital materials such as vid-

eo, sound, image, or text files, and composite media like Web pages. Implementations of Dublin Core typically make use of XML and are, in contrast to the current versions of the ISO TC211 suite, based on the Resource Description Framework (RDF). Dublin Core is defined by ISO through ISO Standard 15836.

For the time being, there is no metadata profile for geographic information available, which is based on Dublin Core. As a conclusion, only profiles based on the ISO TC211 suite are considered, although Dublin Core is proposed in the Ethiopian interoperability framework. The reason for this is that there are vfew international experiences in utilizing Dublin Core for geospatial datasets.

When evaluating the different metadata profiles, special attention is given to their compliance with the ISO TC211 standards (19115, 19119 and 19139). Other criteria include the availability of tools and to what extent the profile has been used operationally.

The first version of the ISO 19115 standard was released in 2003 (ISO, 2003). However, several countries, such as USA, Canada and Australia, started to use metadata long before that. These national standards are still in use. In order to adapt to the international standards, mappings between the national standard and the ISO standards have been defined. This means that it is possible to use the metadata elements identified in these older national standards, but mapped into an ISO context. For some metadata elements, these mappings are not semantically correct, so approximate mappings are used.

The ISO 19115 metadata standard also has a set of core elements. These core metadata elements are required in order to identify a dataset, typically for catalogue purposes. Using the recommended optional elements in addition to the mandatory elements will increase interoperability, allowing users to understand without ambiguity the geographic data and the related metadata provided by either the producer or the distributor. A metadata profile compliant to the ISO 19115 standard shall include this core.

The FGDC standard does not support some ISO core elements, mainly related to languages and character sets. This means that tools that are tightly coupled with the FGDC standard are less suitable if full ISO compliance is to be achieved.

The Anzlic approach to metadata is consistent with the US FGDC approach (Anzlic, 2001). A mapping to the ISO 19115 standard has been developed. Otherwise the Anzlic standard has similar drawbacks as the FGDC standard.

The INSPIRE metadata profile consists of three levels. The basic level specifies metadata at a general level, also including metadata for services. The middle level specifies additional metadata to be applied for datasets (not services). Then additional thematic metadata may be specified for each theme.

Based on the evaluation of existing metadata profiles, it is recommended that the European metadata profile for INSPIRE (European Commission 2010) is selected as a draft Ethiopian metadata profile.

It should however be pointed out that the experiences of this metadata profile is limited only to the European context. It is therefore recommended that the European metadata profile for INSPIRE is tested with respect to the Ethiopian context.

3.2. Web services guidelines

One additional objective of this paper is to propose specifications of web services to be included in the National Spatial Data Infrastructure of Ethiopia. A web service is in this context defined as "a software system designed to support interoperable machine-to-machine interaction over a network" (W3C, 2004). Web services provide a standard means of interoperating between different software applications, running on a variety of platforms and/or frameworks. This means that focus is on machine-to-machine communication and to the utilization of other web standards and protocols such as HTTP, SOAP and WSDL.

Implementation of a proper GIS application operation, maintenance and replacement plan is critical in web service standardization. For the time being, none of the governmental organizations have fully functional replacement plan. Similar conclusion can be drawn about GIS Software maintenance. All-in-all, most of the organizations with exception of AAU and ISS needs to make more effort on GIS application portfolio operation and maintenance.

Since the basic foundations for successful GIS web services are lacking, it is not surprising that GIS Web Services are not currently implemented. However, most of the organizations are in the process of doing so, see *Figure 3*. This may indicate that this is a proper time to introduce more homogenous standards regarding geospatial data exchange.

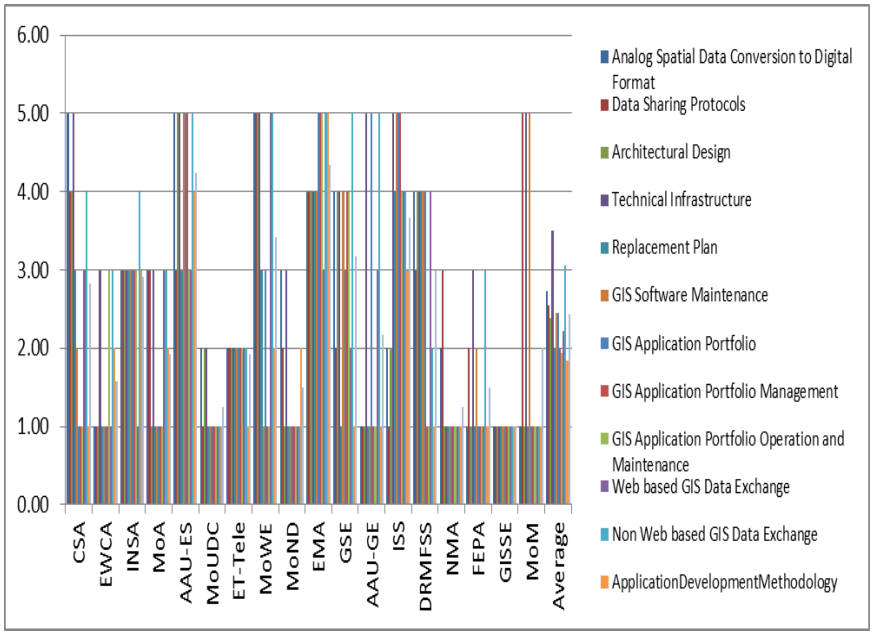

Figure 3. Current status of GIS web services and GIS application

Standards for web services in the GI domain are developed and specified by the Open Geospatial Consortium (OGC). Some of these initiatives have also been approved as ISO standards, for instance Web Map Service (ISO 19128), Web Feature Service (ISO 19142) and Geography Markup Language (ISO 19136).

The Ethiopian e-Government Interoperability Framework (PriceWaterhouseCooper, 2010) recommends the use of W3C standards for web services.

In an international context, there are a several types of web services that are in use, such as discovery services, view services, download services and coordinate transformation services. Other type of services which has been discussed in the INSPIRE context are the schema transformation services and invoke services. However, these services have currently limited support from the international standardization community. As a consequence, the first mentioned four services are proposed to constitute the initial backbone of services in the Ethiopian NSDI. A more detailed description of the web services is provided by OGC (2005).

3.3. Road map

The aim of the roadmap is to describe the major activities required for implementing the GI standards in Ethiopia. The roadmap shall also include major milestones, which serves as checkpoints for the monitoring of the implementation process. It is not possible at this stage of the development to specify the timing of the actions in detail. In order to do this, the financing, governance and training needs have to be known in detail. Instead, the roadmap presented here is instead to outline the required steps, without fixing dates and durations for each action.

The roadmap here proposed is based on the Berners-Lee 5-star model. This model has however been modified slightly, in order to fit the Ethiopian context and the nature of GI standards. In addition to this, the motivation for upgrading services in the 5-star model is mapped into a business motivation model developed by Zdravkovic and Östman (2008). This business motivation model states that the critical success factors for services are that they are

– *Fast*. The delivery of the service to the recipient shall be fast.

– *High quality*. The provided information shall be correct, relevant, and up-to-date.

– *Flexible*. The service should be customizable to the needs of the customer.

– *Low Cost*. The service should be cheap, that is, provided at a low cost.

– *Convenient*. The service shall be provided with high convenience.

– *Reliable*. The service shall be provided with a consistency, that is, it should be always completed.

Due to space limitations of this paper, the entire roadmap can not be presented here. Instead, the transition process from level 0 to level 1 is given as an example.

Level 0 means here that data and maps are available in electronic form. The data and the maps may be distributed by various media such as CD's or USB sticks sent by ordinary mail. Level 1 in the 5-star model means that the information is provided using simple web access. There are no requirements on the format or semantics being used, see *Table 1*.

Table 1. The change of standards when upgrading from level 0 to level 1

Standard component	Standards, level 0	Standars, level 1
Data	Data files	Data files
Maps	Printed or electronic	Image files
Metadata	Modified INSPIRE metadata standard	Modified INSPIRE metadata standard
Discovery service	Geoportal, based on OGC CSW	Geoportal, based on OGC CSW
View service	Delivery by mail, no standard	Web pages with files to download
Download service	Delivery by mail, no standard	Web pages with files to download
Coordinate transformation service	-	-

Compared to the services at level 0, the benefits listed in *Table 2* may be achieved for the clients performing the transition from level 0 to level 1

Table 2. Expected benefits when moving from level 0 to level 1.

Benefit	Outcome
Faster delivery	Yes
Higher quality	Yes, better usability of map images
More flexible service	No, perhaps less flexible
Provided at lower cost	Yes, probably
More convenient	Yes, 24:7
More reliable	Yes, less human errors

4. Conclusion

The conclusions of this paper may be summarized as follows

The proposed metadata profile is based on the INSPIRE metadata profile. However, some metadata tests have been carried out in Ethiopia and the results of these tests should be considered when setting the Ethiopian metadata profile.

The proposed standardized web services are discovery services, viewing services, download services and coordinate transformation service. For the first three service types, plain OGC specifications are proposed. The OGC version of coordinate transformation services is however outdated. The INSPIRE version is instead based on pure OGC Web Processing Services, which is more aligned with the general e-Government strategy of Ethiopia.

In order to optimize the costs and benefits of the implementation of the standards, a step-wise approach to the implementation is proposed. This strategy is based on the Berners-Lee 5-star model, but adapted to the Ethiopian context within the field of geographic information. The model also clarifies the operational benefits of moving from a lower stage to a higher.

A successful implementation of the standards also requires a proper testing strategy as well as other supporting actions like training and organizational development.

References

Anzlic (2001). Anzlic Metadata Guidelines: Core metadata elements for geographic data in Australia and New Zeeland. http://www.ga.gov.au/image_cache/GA9364.pdf

Anzlic (2011). Anzlic – the Spatial Information Council. http://www.anzlic.org.au/

ASDI (2011). Requirements and Standards. http://asdd.ga.gov.au/asdd/tech/requirements.html

Berners-Lee T (2006). Linked Data. http://www.w3.org/DesignIssues/LinkedData.html

European Commission (2008). Methodology for the development of data specifications. http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf

European Commission (2009). Draft Technical Guidance for INSPIRE

Download Services. http://inspire.jrc.ec.europa.eu/documents/Network_Services/INSPIRE%20Draft%20Technical%20Guidance%20Download%20(Version%202.0).pdf

European Commission (2010). INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119.

http://inspire.jrc.ec.europa.eu/documents/Metadata/INSPIRE_MD_IR_and_ISO_v1_2_20100616.pdf

European Commission (2011). INSPIRE. http://inspire.jrc.ec.europa.eu/

FGDC (2011a). The North American Profile (NAP) of the ISO 19115. http://www.fgdc.gov/metadata/geospatial-metadata-standards#nap

FGDC (2011b). Federal Geodata Committee. http://www.fgdc.gov/

Inkdroid (2011). The 5 stars of open linked data. http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/

ISO (2011). ISO/TC 211 Geographic information. http://www.isotc211.org/

Lopez-Pellicer F.J, Florczyk A.J, Nogueras-Iso J, Muro-Medrano P.R, Zarazaga-Soria F.J (2011). Linked Open Data for INSPIRE: From 3 to 5 Star Geospatial Data. Paper presented at the INSPIRE 2011 conference in Edinburgh. http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2011/abstracts/170.docx

Makanga P, Smit J (2010). A Review of the Status of Spatial Data Infrastructure Implementation in Africa. South African Computer Journal, No 45, July 2010. http://sacj.cs.uct.ac.za/index.php/sacj/article/viewFile/36/19

OGC (2011). OGC Standards and Specifications. http://www.opengeospatial.org/standards

Office of Spatial Data Management (2011). http://www.osdm.gov.au/

PriceWaterhouseCoopers (2010). E-Government Interoperability Framework for Ethiopia – Final Report.

W3C (2004). Web Services Glossary. http://www.w3.org/TR/2004/NOTE-ws-gloss-20040211/

Zdravkovic J, Östman A (2008). An Agile Method for Automated Provisioning of the Geographical Information in Public Services. Urban and Regional Data Management: UDMS Annual 2008. ISBN: 978-0-415-44059-2, Taylor and Francis, London.