Automatic Metadata Creation for Supporting Interoperability Levels of Spatial Data Infrastructures

M. Manso-Callejo, M. Wachowicz, M. Bernabé-Poveda
Topographic and Cartographic Engineering Department, Technical University of Madrid
ETSI en Topografía, Geodesia y Cartografía, Ctra. Valencia Km. 7,5 28031- Madrid (Spain)
{m.manso@upm.es, monicawachowicz@gmail.com, ma.bernabe@upm.es}

Abstract:
Interoperability in Spatial Data Infrastructures (SDI) is a full-grown subject and an objective with many shortcomings as far as definition of standards for geographic data transfer and exchange, different data types’ integration, and comprehensive semantic models is concerned. There is a vast literature available on interoperability models containing different interoperability levels, including technological, syntactic and semantic levels. However, very limited research has been carried out on the development of interoperability models for the implementation of Spatial Data Infrastructures (SDI). This paper provides a short review of the main advances in interoperability related to SDI. It also discusses the important role of metadata elements in the formalization of interoperability models for the implementation of SDI. We propose an integrated interoperability model based on the definition of a common template that integrates seven interoperability levels: technical, syntactic, semantic, pragmatic, dynamic, conceptual and organizational. The implementation is carried out by automatic production of ISO19115 metadata. Finally, the results outline the strength and deficiencies in terms of the dynamic interoperability level of SDI based on the elements of ISO19115 metadata.

Keywords: SDI, Interoperability Model, Metadata.

1 INTRODUCTION

Two usual definitions of SDI have been provided: 1) the SDI cookbook version 2.0 "Spatial Data Infrastructure" (SDI), which is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data’ (GSDI 2004 p. 8), and 2) SDI subsumes technology, systems, standards, networks, people, policies, organizational aspects, geo-referenced data, and delivery mechanisms to end-users (Georgiadou, Puri, and Sahay 2005, Williamson 2004).

SDI provides a basis for spatial data discovery, evaluation, and application for users, promoting a reliable environment to facilitate the access to geographic information and the agreements, organizations and programs needed to coordinate SDIs at different scales (Béjar et al. 2008; GSDI 2004).

Georgiadou, Puri, and Sahay (2005) consider that SDI must be dealt with from both a technical and a social point of view. They state that SDIs are a special case of Information Infrastructures (IIIs), specifically geared towards geographic information, according to Bernard et al. (2005). Béjar et al. (2008) propose other frameworks to support SDI research such as System of Systems (SoS) compositions, in the sense suggested by Maier (1996). Before reviewing the relationship between II and SoS, the main conclusion of this proposal is that these terms are used to refer to similar concepts from different perspectives, SoS being a broader term. This relationship provides a new conceptual framework to study SDIs.
As a result of the review carried out to define the state of the art of interoperability within the GIS context (Manso and Wachowicz 2009), we planned to research interoperability in the SoS, the use of non-hierarchical levels and the definition of a model allowing measurement of interoperability between systems. As an outcome of the research carried out on system interoperability, where we have dealt with SDI as an SoS, we have put forward an Interoperability Model Based on Metadata for SDI (IMBM-SDI) (Manso, Wachowicz and Bernabé, 2009). This model is made up of the technical, syntactic, semantic, pragmatic, dynamic and conceptual levels – as defined by the “Levels of Conceptual Interoperability Model”, LCIM) – to which an organizational level has been added, gathering the legal aspects, the data policies and responsibilities among other aspects. In addition to defining the levels the integrated model is made up of from the SDI perspective, the metadata items of ISO 19115 (2003) are analyzed as regards the interoperability they provide. This same idea has also been used by Tolk, Diallo and Turnitsa (2007) by applying the LCIM model in the design of SoS supporting integration, interoperability and orchestration or chaining of systems.

Metadata are key elements for Information Infrastructures (II), especially for SDIs. In addition to carrying out discovery, evaluation, access and use functions (Gayatri & Ramachandran 2007, Johnston 2005, GSDI 2004, Gilliland-Sweetland 2000, Beard, 1996), they may support interoperability. In this sense Tolk, Diallo y Turnitsa (2007) state that ‘we are moving towards a “Dynamic Web”, supporting the orchestration and alignment of agile components at least up to the dynamic layer with standardized metadata’. In our IMBM-SDI model we have also verified that the main interoperability levels supported by ISO 19115 metadata are the semantic, dynamic and organizational levels.

West and Hess (2002) and Guptill (1999) state that the manual creation of contents adequately describing a geodata set is a boring, tedious work consuming a great deal of resources, besides being prone to errors. Anderson and Pérez-Carballo (2001) maintain that the metadata created by automatic procedures tend to be more efficient, more consistent and cost-effective than the manually created metadata.

The purpose of this paper is to show that ISO19115-conformant metadata automatically produced for a dataset, support the interoperability of systems in an SDI. Given the importance of the dynamic interoperability and the complexity of the metadata standard in the composition of services, the scope of this demonstration is limited to metadata items capable of being automatically produced and supporting such interoperability level.

The remainder of this paper is organized as follows. Section 2 reviews the interoperability levels described in the literature and related to SDI; the most prominent authors are indicated. Section 3 describes the meaning and objective of the levels in the “Integrated Model of Interoperability for SDI”. Section 4 describes the role of metadata at the interoperability levels of the integrated model. Section 5 enumerates the metadata items supporting dynamic interoperability and able to be automatically generated by extraction, computing or inference; a description about how to generate the metadata is included. In Section 6 the conclusions and future research lines are presented.

2 UNDERLYING INTEROPERABILITY LEVELS

In SDI different hardware and software components, supplies, policies, procedures and people must interoperate, for storage, processing and enabling access to spatial data. Interoperability has many meanings, including the facets of communication, exchange, cooperation, and sharing of information between systems. In fact, the
The essence of interoperability is being relationships between systems, where each relationship is a form of communication, exchange, cooperation and sharing (Carney et al. 2005). Usually different definitions of interoperability differ in terms of the relationship description and the system components. Some focus on system and hardware components, others on services in order to provide information and components sharing information, and still others on how to use the exchanged information in a meaningful manner without any special manipulation (Ford et al. 2007, Flater 2002, Buehler and McKee 1998). In contrast, interoperability definitions proposed by policy organizations have stressed the engagement process necessary to exchange and re-use information. Some authors point out that such a process should ensure that systems, procedures and the culture of an organization are managed in such a way as to maximize opportunities for exchange and re-use of information, whether internally or externally (Dekkers 2007, Nedovic-Budic and Pinto 2001, Miller 2000).

The levels of interoperability are a set of criteria and associated processes for assessing system capabilities in the context of the degree of interoperability required. Several levels have been proposed in the literature according to the interoperability required: semantic, syntactic, technical, organizational, schematic or structural, pragmatic, dynamic, social, legal, and others.

**Semantic** Interoperability is related to the meaning of the information. Some authors have emphasized the need for a common reference model for information exchange and interpretation of concepts (Turnitsa and Tolk 2006, Kalantari et al. 2006, Antonovic and Novak 2006). The fact that the geographic space may have more than one description is handled as semantic heterogeneity (Kuhn and Raubal 2003). Standardized taxonomies in domains related to geography as well as new approaches for the description of semantic proximity between objects are objectives of this level (Probst 2006, Rodríguez and Egenhofer 2003).

The **Syntactic** Interoperability level provides a common structure to exchange information (Turnitsa and Tolk 2006) or a common message format (Shekhar 2004).

Communication, transport, storage and representation standards are understood as **Technical** Interoperability aspects (Miller 2000), and system linking is the basic objective (Schekkerman 2004).

The **Pragmatic** Interoperability deals with the exchange of services and processing (Shanzhen et al. 1999). This level is reached when the interoperating systems are aware of the methods and procedures every one of them is employing (Turnitsa and Tolk, 2006).

The **Organizational** Interoperability deals with enabling a cooperation process by encouraging partnership among organizations to assist best practices in data sharing (Lance et al. 2008, Georgiadou and Harvey 2007).

Abstract model interpretation by third party is the main objective of the **Conceptual** Interoperability level (Turnitsa and Tolk 2006), in such a manner that the model might be documented by engineering methods. Some authors have studied different data types, label inconsistencies, aggregation discrepancies and generalization conflicts (Shekhar 2004, Bishr 1998, Goh 1997).

When systems are able to detect state changes and to take advantage of those changes (Turnitsa and Tolk 2006) or when systems are able to locate resources for
their use based on the existence of standardized metadata (Shanzhen et al. 1999), Dynamic Interoperability is achieved.

Intellectual Property (IP) rights (Miller 2000) and directives, rules, parameters and instructions for the management of business workflow, and considering incorporation of information and communication in the business of land administration (Kalantari et al. 2006) are all aspects dealt with by Legal Interoperability.

The Social Interoperability identifies aspects such as interests, beliefs, expectations and commitments (Assche 2006, Mohammadi et al. 2006).

Assuming that a large set of systems and SDIs must interoperate in the near future, we claim that there will be a need to study the complexity of the relationships between interoperability levels of SoS, and to develop new interoperability models. One example is given by the design of the Global Earth Observation System of Systems that aims to provide decision-support tools to a wide variety of users (GEOSS 2008).

3 INTEGRATED MODEL OF INTEROPERABILITY FOR SDI

As we have described in ‘Towards an Integrated Model of Interoperability for Spatial Data Infrastructures’ (Manso et al. 2009), before reviewing interoperability levels and models, some levels are not pertinent to this context and some are included within other levels. This integrated model of interoperability for SDI is made up of levels: Technical, Syntactical, Semantic, Dynamic, Pragmatic, Conceptual and Organizational. Our integrated model assumes the associations across all interoperability levels with different intensity. Figure 1 present an overview of the resulting pair-wise relations of our interoperability model.

**Figure 1: Integrated Model of Interoperability for SDI.**

The meaning and objectives for these levels is summarized below.

Technical interoperability is the level enabling the interconnection of systems through common communication protocols allowing information exchange at its most basic level: bits and bytes. Regarding SDIs, technical interoperability is a series of technical aspects: character sets, character encoding, file identifiers, description of processing environment, file names, service types and versions, transfer size, file formats and versions, storage means, links and protocols.
Syntactic interoperability is about the information exchange between systems by using a common data format or structure, language, logic, registers and files. Key aspects are the standards or format specifications which structure information, so that the information could be interpreted and processed. XML Schemas (XSD) defined by Open Geospatial Consortium (OGC) is a good practice enabling syntactic interoperability.

Semantic interoperability is about information exchange using a shared, common vocabulary that avoids inaccuracies or mix-ups when interpreting the meaning of terms. Web Service Description Language (WSDL), the Simple Object Access Protocol (SOAP) at the level of service interconnection or the Geographic Mark-Up Language (GML) for the transfer of vector GI, and the Style Layer Description (SLD) for definition of a visualization style are good practices enabling semantic interoperability.

The dynamic interoperability allows systems to supervise other systems and to respond to detected changes in information transfer or time delay, taking advantage thereof. In order to enable switching from the use of one service to another by supervising the functioning of the network and other services, systems need dynamic discovery capabilities of services complying with the requirements called for. Initiatives such as the spatial data themes of Directive 2007/2/EC, classified into a topic category metadata help to keep the reliability of dynamic service interchange.

Conceptual interoperability is about knowing and reproducing the functions of a system based on documentation usually stored in formats used in Engineering. Aspects of conceptual interoperability are those describing data and system model in standardized documentation format from an engineering viewpoint. OGC WFS describeFeatureType response containing a GML application schema is a good practice of conceptual interoperability.

The organizational interoperability allows knowledge of business targets, process models, regulations and policies of access and use of data and services. It has to do with aspects related to expectations, contracts and culture. Knowledge about goals, responsibilities, access and use policies are considered as a constraint or as identification information helpful to evaluate the use of metadata elements.

We advance an interoperability model which, in addition to the models proposed by the LCIM, includes the organizational interoperability level, in which the legal aspects and the relationships between data providers and users are framed, as mentioned above in the Introduction, along with the definition of SDI.

Although the scope of the integrated model of interoperability has been defined for SDIs, we believe that it may also be used within wider contexts, such as SoS.

4 METADATA ROLES IN THE INTEGRATED METADATA MODEL

As mentioned above, metadata in II, and especially in SDIs, allow the discovery, evaluation, access and use functions. The wider, boosted utilization of metadata allows discovery and access to data, as we are observing in INSPIRE Implementation Rules (IR) Metadata. However, from the perspective of system interoperability, some authors highlight the usefulness of metadata to enable different interoperability levels in the LCIM model (Tolk, Diallo and Turnitsa 2007). These authors emphasize the usefulness of metadata for the communication among intelligent software agents, ‘to communicate about situations’, ‘to enable software agents to select different components and compose them to evaluate alternative hypotheses’, ‘to support decision makers’ and
finally ‘to support the orchestration and alignment of agile components at least up to the dynamic layer’.

Under SDI context, metadata has been consider an element that support discovery, evaluation, access and use functions mainly oriented to human users. Since SoS perspective, metadata enable communication among intelligent agents, hypotheses evaluation, support decision makers and orchestration of components in dynamic contexts. In the near future intelligent software agents can run SDI components using geospatial metadata. In this scenario interoperability levels provided by metadata are important. The lack of studies regarding the analysis of interoperability levels provided by the ISO19115 metadata items bears out our consideration.

In this Section the results of such analysis are presented. Table 1 shows the count of the metadata items which support the different interoperability levels of the integrated model individually. These values can be appreciated in the main diagonal of Table 1. The remainder of the values indicates the amount of items supporting two interoperability levels simultaneously.

Table 1: Metadata elements count for each interoperability level and simultaneous relations.

<table>
<thead>
<tr>
<th>Interoperability Level</th>
<th>Organizational</th>
<th>Semantic</th>
<th>Dynamic</th>
<th>Technical</th>
<th>Pragmatic</th>
<th>Conceptual</th>
<th>Syntactic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational</td>
<td>229</td>
<td>181</td>
<td>143</td>
<td>22</td>
<td>19</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Semantic</td>
<td>181</td>
<td>196</td>
<td>127</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Dynamic</td>
<td>143</td>
<td>127</td>
<td>151</td>
<td>10</td>
<td>21</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Technical</td>
<td>22</td>
<td>11</td>
<td>10</td>
<td>32</td>
<td>8</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>19</td>
<td>4</td>
<td>21</td>
<td>8</td>
<td>24</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Conceptual</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Syntactic</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

The results of the classification study have shown that the metadata elements supporting semantic interoperability (196) also provide dynamic (127/196 – high intensity) and organizational (181/196 – high intensity) interoperability levels as well. The same results may be observed for the dynamic and organizational interoperability. The interoperability levels least favored by metadata items of the ISO19115 Standard are the syntactic, conceptual, pragmatic and technical levels.

We can say that the proposed Interoperability Model Based on Metadata for SDI (IMBM-SDI) is not a hierarchical model in which the higher levels need all the lower levels; instead each level may require the functionality of other non-subordinate interoperability levels. The proposed integrated model of interoperability might be metaphorically compared with the TCP/IP communication protocol and other models proposed in the literature such as LCIM, suggesting hierarchical structures in the style of ISO communication protocols.

The study has also allowed determining the set of metadata items providing a definite interoperability level, as well as those providing several levels simultaneously (e.g. semantic and dynamic levels).
In-depth analysis of this information may be useful to define a core set of metadata items that should maximize interoperability with a minimum of metadata or else ensure a minimum of interoperability based on the metadata. This analysis might be carried out while thinking of the future, so that expert and automatic systems will be able to exploit SDI’s functionalities and capabilities.

Such as has been comment, manual creation of geospatial metadata is tedious works that consume resources and is error-prone. This has motivated us to study automatic metadata productions to support dynamic interoperability level.

5 AUTOMATIC METADATA PRODUCTIONS TO SUPPORT INTEROPERABILITY: DYNAMIC INTEROPERABILITY CASE USE

In this section the metadata items supporting dynamic interoperability and being automatically produced by extraction, computation and inference are enumerated, as proposed by Beard (1996).

In the first stage, in order to reach this objective, the metadata items have been adopted, providing dynamic interoperability resulting from the analysis carried out to define the integrated model of interoperability. In the second stage, use has been made of the knowledge acquired by investigating the data that may be automatically extracted from the different types of Geographic Information (GI) (raster, vector and Digital Terrain Model - DTM), in order to select the metadata items that may both provide dynamic interoperability and be automatically obtained. In the third stage, the remainder of the items has been analyzed with the purpose of considering if it is possible to assign them value, by carrying out some kind of calculation or deriving information on the basis of other pieces of available information.

Table 2 shows the outcome of these three stages. After its presentation, we will reflect on the techniques used in order to be able to automatically produce these metadata items.

Table 2 contains 5 columns.
The class to which the item belongs is shown in the first column.
The item is identified in the second column.
Items classified as “P” (produced: extracted, calculated or derived) or as “M” (may appear many times according to the data type) are identified in the third column.
Items that are only applied to a certain type of GI (“R” = raster data; “D” = DTM and “V” = vector data) are identified in the fourth column.
A description of the content and an explanation appears in the fifth column.

Table 2: Automatic metadata element production enabling dynamic interoperability.

<table>
<thead>
<tr>
<th>MD_Metadata: Packed</th>
<th>Metadata element</th>
<th>Production (P), Multiple values (M)</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>distributionInfo: distributionFormat</td>
<td>dateStamp</td>
<td>P</td>
<td>Date, time of metadata generation</td>
</tr>
<tr>
<td>distributionInfo: distributionFormat</td>
<td>name</td>
<td>P</td>
<td>Implicit to GI store.</td>
</tr>
<tr>
<td>distributionInfo: distributionFormat</td>
<td>version</td>
<td>P</td>
<td>Some stores have versions</td>
</tr>
<tr>
<td>distributionInfo: distributionFormat</td>
<td>fileDecompressionTechnique</td>
<td>P</td>
<td>Some stores (mainly imagery) use compression techniques</td>
</tr>
<tr>
<td>contentInfo:MD_CoverageDescription</td>
<td>contentType</td>
<td>P</td>
<td>Raster and grid data must be distinguished</td>
</tr>
<tr>
<td>Description</td>
<td>Sequence Identifier</td>
<td>Role</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: Band</td>
<td>Band number</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_RangeDimension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: Pixel or cell maximum value</td>
<td>Pixel or cell maximum value</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: Pixel or cell minimum value</td>
<td>Pixel or cell minimum value</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: Number of bits used to encode band</td>
<td>Number of bits used to encode band</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: BitsPerValue</td>
<td>When grid data have associated categories like classified image, Name for every feature type</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>For raster and rectified grid, pixel size can be extracted, computed or inferred.</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Title inferred for dataset, based on BBOX, time information and other information that can be extracted, computed or inferred.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Date, time of metadata generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Date</td>
<td>P</td>
<td>Creation</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>DateType</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Name</td>
<td>P</td>
<td>Same as distribution Format</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Version</td>
<td>P</td>
<td>Same as distribution Format</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>FileDecompressionTechnique</td>
<td>P</td>
<td>Same as distribution Format</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>useConstraints</td>
<td>P</td>
<td>Some datasets stores use constraints. In this case, extract and include constraints.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>keyword</td>
<td>P</td>
<td>Based on raster datasets analysis, content type can be inferred. Then some keywords included in thesaurus can be inferred.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>type</td>
<td>P</td>
<td>Keyword type</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Title</td>
<td>P</td>
<td>Thesaurus Name</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Date</td>
<td>P</td>
<td>Thesaurus keywords date</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>dateType</td>
<td>P</td>
<td>Thesaurus date type</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>spatialRepresentationType</td>
<td>P</td>
<td>From dataset representation type can be extracted.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Distance</td>
<td>M</td>
<td>For raster and rectified grid, pixel size can be used to define resolution distance.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Denominator</td>
<td>M</td>
<td>From resolution distance denominator can be computed, but only one is needed.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>ExtentTypeCode</td>
<td>P</td>
<td>1 (true: inclusion)</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>WestBoundLongitude</td>
<td>P</td>
<td>West Longitude computed from dataset identifying source CRS.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>EastBoundLongitude</td>
<td>P</td>
<td>East Longitude computed from dataset identifying source CRS.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>SouthBoundLatitude</td>
<td>P</td>
<td>South Latitude computed from dataset identifying source CRS.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>NorthBoundLatitude</td>
<td>P</td>
<td>North Latitude computed from dataset identifying source CRS.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Title</td>
<td>P</td>
<td>Geographic identifier (toponym) computed by gazetteer reverse query.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>Date</td>
<td>P</td>
<td>Date time extracted from gazetteer metadata or database used to compute toponym.</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>dateType</td>
<td>P</td>
<td>Publication</td>
</tr>
<tr>
<td>Dimension: MD_CoverageDescription: GeometricObject</td>
<td>MinimumValue</td>
<td>M</td>
<td>For grid dataset containing DTM, lower value.</td>
</tr>
</tbody>
</table>
Most of the 54 items enumerated in Table 2 are produced by extraction of information from data stores, whether files, directories or databases. That is the case for the items: version of format, data decompression technique, number of bands, number and types of geometry, geometric resolution of pixels, maximum and minimum coordinates in the reference system used, information about constraints and restrictions of use and access.

Some of them, e.g. dateStamp or the name of the format may be obtained from the context (date on the clock of the computer system).

There are also a sufficient number of items (12) that may be computed when they do not appear in the spatial information store. That is the case of the maximum and minimum radiometric values of the bands in the raster data, the denominator of the scale in raster data, the maximum and minimum non-geographic coordinates or the geographic identifier from the previous coordinates.

Another set of items (>12) may be produced by inferring their values. If the type of content stored in the spatial data (data mining, image classification) is somehow determined, a set of keywords belonging to thesauri may be contributed describing that content, and a title for the dataset may also be inferred. An important item for data is also the identification of the spatial reference system for data representation. On the basis of the information stored, it is sometimes possible to deduce the identifying code.

As a first conclusion, we may state that of the 151 metadata items providing dynamic interoperability, 54 of them may be automatically produced (35%). Although this value is quite high, it should be cautiously interpreted since it is a “hopeful” value representing the ceiling of the automatic production.

Depending on the type of spatial data storage, its nature and the possibility of determining the content type, these values may diminish or increase markedly since a fair amount of items have been identified as elements susceptible of being multi-assessed (cardinality >1).

6 CONCLUSIONS AND FUTURE RESEARCH WORK

Spatial Data Infrastructures (SDIs) subsume technology, standards, networks, people, policies, organizational aspects and systems as has been suggest by different authors. SDIs are a special case of Information Infrastructures (IIs). System of Systems (SoS) is a broad term that includes IIs. From this perspective SDI can be dealt with and modeled as an SoS.

The definitions of the interoperability levels have been reviewed and synthesized in the SoS context and an Interoperability Model Based on Metadata for Spatial Data
Infrastructures (*IMBM-SDI*) has been proposed. It is based on the *LCIM* model which has been extended with the organizational interoperability.

The authors of the *LCIM* model have studied how to achieve dynamic interoperability to orchestrate and chain services in the SoS and they have committed themselves to the use of metadata to enable communication between agents. We have followed that suggestion and analyzed metadata items of the International Standard ISO19115. The interoperability levels supported by every item have been identified. Table 1 shows the aggregate results of the analysis. It has been argued that it is a non-hierarchical integrated model.

As anticipated in the conclusions for the sections, we have bet on a new interoperability model in line with the existing models in the SoS context. We have also analyzed the interoperability that may support the metadata items of the ISO19115 Standard.

It has been argued that the manual creation of metadata is a slow, costly process, prone to the introduction of errors in metadata; nevertheless we bet on metadata as an element supporting system interoperability. This study shows at the theoretical level how to automatically produce metadata items supporting dynamic interoperability, by extracting the information stored in files and databases through computations or by inference.

The results shown indicate that a high percentage of the metadata items providing dynamic interoperability – limited objective of the study – may be automatically produced. This fact strengthens the hypothesis of automatic creation of metadata useful from the point of view of interoperability.

The following lines of research are suggested: 1) Study of metadata items from the point of view of interoperability in the framework of the proposed integrated model in order to define a metadata core with this aim, just like the ISO19115 Standard proposes a core from the perspective of discovery and use of resources. 2) Study of automatic methods to produce metadata inducing interoperability at the different levels proposed in the integrated model. 3) Implementation and testing of the automatic methods of metadata production with the end of encouraging interoperability in the domain of virtual map libraries made up of geo-referenced, digitized historical cartography, accessible through standardized services.

7 ACKNOWLEDGEMENTS

This work has been partially supported by the National Geographic Institute (IGN) of Spain through the project UPM-P086005573 and for the Plan Nacional de I+D through the Projects HUM2007-31128-E and CSO2008-03248/GEOG (Prototipo de Cartoteca Histórica Virtual Distribuida – Prototype of Distributed Virtual Historical Map Library)

REFERENCES


