World Spatial Metadata Standards 3
H. Moellering, H.J.G.L. Aalders & A. Crane (Editors)

An Introduction to Metadata for Geographic Information

Henri J.G.L. Aalders

Delft University of Technology, Research Institute for Housing, Urban and Mobility Studies, The Netherlands
Katholieke Universiteit Leuven Faculty of Engineering, Belgium
e-mail: h.aalders@otb.tudelft.nl

1. Introduction

Traditionally, geographic information was produced and used by the geo-spatial community, such as experts in specialised fields as geography, cartography, geodesy, photogrammetry, remote sensing, bathymetry, hydrography, geology, soils, physical planning, architecture, etc.

The work of these experts resulted in the cartographic representation of the real world originally in maps and later as a digital computer model. Combinations of different geo-spatial datasets in various scales, projections and co-ordinate systems, content and appearance create a united computer model of the earth as a whole or in parts. This includes thousands of aggregates of digital geographic, topographic and thematic maps (such as ortho-photomaps, city maps, air- and space-photo images, (3D-) terrain models with gravimetric, photogrammetric and cartographic information) integrated under common idea and ordered by geodetic base, content, design, and reference information stored digital form.

To make the geographic information available to users, the geo-spatial community prepares and fulfils measures to develop and implement conceptual and methodological bases, normative and legislative regulations and dataset descriptive standards including formats for spatial data interchange. It also develops implementation methods, hard- and software tools and technologies for acquisition and the transfer of datasets to users [Moellering 1991], [Moellering 1994], [Moellering 1997], storage, analysis and processing digital cartographic data and the creation of analogue and digital representations of these datasets.

In daily life, the public is used to product labels to receive information about the products they acquire. These labels have several functions: they identify the product as well as promote and describe it. For instance in many food industries, legislation clearly states what information the labels need to contain and how it should be presented, while for technical equipment the product label summarises the content of a package and a guide describes its use. In addition, this information is widely controlled by international regulations. For geographic data, no such legislation or internationally controlled system exists. However, now that the geographic community enters into the age of global spatial data infrastructures to make these datasets available to non-specialised users, knowledge about widely distributed geographic datasets is essential to universally allow users to locate, evaluate, extract, and employ the data. As expert personnel changes and time passes, new workers may have little understanding of the datasets' content and consequently the information about the organisation's data will be lost and the data may loose its value. Therefore standardised descriptions of datasets are required following the rules set in metadata standards – both official and defacto.

Descriptions of geographic datasets have been in existence for some time now. In many cases these geographic dataset descriptions for different information communities on regional and national levels evolved in different ways and are incompatible. As the existing general dataset descriptions, that have become national or regional standards, provide hardly global interoperability, they are at least insufficient to support geographic information from different areas in the world. This incompatibility and insufficiency was the motivation for the development of a global standard for data descriptions: ISO 19115, Geographic Information – Metadata. The ISO (International Organisation for Standardisation) Metadata Standard was designed to:

- support geographic information;
- work with wider information technology standards and practices;
- serve the global community, in a multi-national, multi-language environment;

based on a foundation of national, regional, and special information community standards and experiences and a thorough requirements analysis, and implementation testing.

The ISO 19115 Metadata standard defines and standardises a comprehensive set of metadata elements and their characteristics, along with the schema necessary to fully, and extensively, document geographic data. The standard applies to all geographic data – it is applicable to datasets in series, datasets, individual geographic features, and their attributes. The standard defines the minimum set of metadata required to serve the wide range of metadata applications, as well as optional metadata elements to support a more extensive description of geographic data. It also provides a standardised manner for users to extent their metadata and still ensure interoperability allowing other users to comprehend and exploit this extended metadata when browsing the web for certain type of geographic data.

In the field of Information Technology (IT), there are two other prominent metadata initiatives: the Dublin Core Metadata Initiative (DCMI, Dublin Ohio, U.S.A., see http://www.xml.com/pub/2000/10/25/dublincore/) and the Institute for Electrical and Electronic Engineers (IEEE) Learning Object Metadata (LOM) Working Group, see http://grouper.ieee.org/p1484/doc/wg12/LOM-WD3.htm).

The Dublin Core metadata element set (or the basic interlocking brick) is intended to support cross-subject search and retrieval. It can be thought of as a simplistic or pidgin metadata language that helps the user navigate through disparate subjects, languages, and cultures. Adoption of the Dublin Core by governments, libraries, museums, archives, publishers, environmental science repositories, prints and e-print archives, to name a few, testifies to its success in this role. There are emerging applications in the commercial sector, as well, with health care organisations and financial industries using the Dublin Core as the basis for organising and exchanging information.

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 3

2. History of metadata

To many in the hard-copy map world, data description, metadata and marginalia are often considered to be the same. They help to describe the properties and characteristics of the dataset's content (in the hard-copy map world called the 'map').

The Greek geographers applied descriptions of map-objects as early as the beginning of Christian era. Many of the Middle Age geographers applied beautiful drawings to accompany their vision of the world in maps [Bertius 1630]. Often in the Middle Ages the symbolisation on maps were very religious and descriptions referred to religious descriptions but usually carried dates and author names.

From the early to mid 18th century, topographic maps and sea navigation charts often carried descriptions and explanations in order to enable the accurate reading of the map information, called marginalia. The marginalia captured dates, bounding co-ordinates, grids, scales, accuracy, author, etc.

The term **metadata** – or used in other forms as *meta data* [McIntosh 1968], *meta-data* [Homer 1978], Ziegler 1978], [Weber 197], [Schelling 1978] and *metadata* [DLC 1979] – first was used in Computer science/Information sciences literature in 1968. Since, it has been used in any of the three forms: firstly, metadata was very popular in the European community in the beginning of the seventies and later started to show up in USA literature (as metadata – all one word) in military documents.

The word 'metadata' stems from Greek and can be described as 'data about data'. The ISO/TC 211 defined metadata by data about the content, quality, condition and other characteristics of data. Although the ICA Commission on Spatial Data Standards adopts the ISO definition on metadata, ariginally it used a description as 'information given along with geographical information and which allows a better understanding of geographical data'. According to this definition, metadata information encompasses transfer format templates, counts of data items in the transfer, conceptual data model and catalogues, positioning references of geographical data, quality reports and the logical description of the transfer metafile.

The term metadata took hold in the USA in the geo-spatial data community in the early 1990's with the development of the concept of the National Spatial Data Infrastructure and the signing of Executive Order 12906 by President Clinton in April 1994 (see: http://www.fgdc.gov).[1]

3. Fundamentals of metadata

Following the concept of set theory, a geographic dataset can contain either several homogeneous geographic datasets or one or more occurrences of geographic entities or any set of attributes or occurrences of geographic entities and/or relationships. As such, one could also

[1] The Metadata Company trademarked the term 'Metadata' in 1986. The Metadata Company sells goods and services related to metadata. Since the FGDC does not sell any goods or services, the FGDC does not feel it is in violation of the trademark as determined by the Department of the Interior, Office of the Solicitor (see FGDC Official statement at http://www.fgdc.gov/metadata/meta_trademark.htm). Since the trademark issue involves goods and services, several companies that make metadata products have tended to avoid the term metadata and have opted to use the term meta-data or even different terms such as data properties and data cataloguing.

speak of levels of datasets. Therefore, data shall be given to describe the data, which occur in a dataset.

Metadata serve a wide range of applications in geographic information as, apart from data description also the organisation and maintenance of data. Complete metadata descriptions of the structure, contents and accuracy of all datasets is an important requirement for database design. Such descriptions also may provide protection for the misuse of the data in a dataset.

A model for a metadata includes the definition of mandatory or optional metadata and which metadata is mandatory under special conditions (conditional metadata). In addition, the minimal metadataset should be clarified. Metadata can be organised in the computer in many ways: as an incorporated part of data within the dataset, in a separate database or as a text file. One may choose the manner of storage of metadata according to a management strategy, budget or other institutional or technical factors.

The best time to define and collect metadata is while the data are being developed, i.e., when the information needed for metadata is known. Probably waiting until after data collection is finished results in less accurate information lacking the description for the surveyors of what ought to be collected in the database. Besides, searching information, when lacking metadata may become cumbersome and increase the cost. However, documenting existing data can be daunting. Details may be long forgotten and costs can be high. Though these concerns are valid in well-considered decisions, they should be taken allowing documenting metadata for existing databases.

Using and combining data from different sources (and maybe also applying processes to them) that are available at different places requires the GRID approach of distributed computing and resources to conduct new and innovative specific applications. This may result in a Virtual GIS (V-GIS) environment in which the Digital Universe resident in the new databases can be seamlessly explored across the entire spectrum carrying the characteristics and possibilities of each of the single databases and spatial processes. This ‘digital era’ has been a topic for international discussion in the past year by the G8 countries, United Nations, heads of Governments, the non-profit community and the international private sector.

3.1. *Hierarchy in standardisation*

The application of standards is different in different countries. In some countries, law or governmental (executive) order, regulates the application of standards, while in other countries the use is left in free will to potential applications between parties. Countries where standards are applied according to state regulations make distinction between:

- state standard adopted by the State Committee on Standardisation, a standardisation law or a governmental (executive) order;
- domain standards that have been adopted by some ministries (departments) or by a group of organisations in the same field;
- application standard used only within certain – private sector – enterprise(s).

An International Standard usually is the standard adopted by the ISO, while a Regional Standard is adopted by a regional international organisation, for example, interstate standards organisations such as the Comité Européen de Normalisation (CEN). Governments or national standardisation institutes joining the agreement on co-ordinated policy of standardisation and

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 5

certification adopt such standards. A National Standard is the standard adopted by national standardisation authority of a certain country. Many standardisation processes result in a set of interrelated standards in a certain field (sometimes also called ‘a complex of standards’) that have the same purpose and establish co-ordinated requirements to the interrelated subjects of standardisation; e.g., the set of standards of ISO 191XX (see also the chapter on ISO standardisation in part 2 in this book).

Since metadata standards include the object definition, in many cases it will be impossible to create a definition for so many different types of users. For example, the definition of a road for a car driver is completely different from the definition for a road used by a road manager, while different road managers use specific definitions in their own application. Therefore, data definition (in the geographic community also indicated by ontology) should follow a standardised way for data definition but may occur at different levels:

1. international level;
2. national level;
3. domain level;
4. application level.

In this situation, the definitions at a lower level follow the definitions at the higher level and the transfer of the meaning of object definitions to be applied in different geographic datasets.

Standards may contain different aspects such as prescriptions for the transfer, ontology, quality, conceptual models, etc. A standard explaining technical terms is a kind of a normative document on certain production or service adopted by the developing organisation, as a rule, in co-operation with potential user organisation(s) and customer(s). Metadata standards follow this terminology.

All standards are based on generalised results obtained from scientific, technical and practical investigations, and they are intended to reach the optimal public benefit. In geographic information, the development of standards involves the methodology, theoretical basis (concepts, methods), technologies of creating and using digital datasets in GIS. In addition, metadata standards follow these general organisational, scientific, technical and application aspects.

3.2. *Metadata modularity and extensibility*

Another concept in metadata applications is modularity [Duval 2002]: metadata modularity is a fundamental principle where vastly diverse sources, styles and content management and approaches to resource description exist. Modularity in metadata allows designers to create new assemblies based on existing metadata schemas: in a modular metadata environment, data elements from different schemas as well as vocabularies can be combined in a syntactically and semantically interoperable way. This allows designers to benefit from re-usability while gathering exiting datasets of metadata and ‘snap’ them together into larges structures, even though the semantics do not appear as obvious to be combined as long as there exist a common syntactic foundation.

Metadata modularity includes also the possibility for extension in order to accommodate particular needs in a specific application. For example, the *identification* of the dataset will be described in any metadataset, while the *degree of cloud cover* will only appear in remote sensing data of aerial photography.

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 6

Furthermore, the metadata modularity requires sufficient refinement to make a more specific meaning of a metadata element, e.g.: *illustrator, author, composer or sculptor* are all particular types of the more general term *creator*, or *date of creation, date of modification, date of acceptance* all have a narrower meaning as the *date* attribute. In addition, the range of values for a given element is a refinement of the metadata element.

In addition, multilingualism becomes a very important aspect since the Web connects a diversity of linguistic and cultural aspects. The Web will fail to achieve its potential as a global information system, unless resources can be made available to users in their native language, in appropriate character sets and with metadata appropriate to resource management.

Here *internationalisation* and *localisation* become apparent, though they may be contradictory: while global resource discovery is best served by internationalisation (using common conventions of practice, languages and character sets throughout the world), the needs of any given community may be better served by supporting local conventions. Basic to the promotion of a global metadata architecture is to translate relevant specification and standard documents into a variety of languages. DMCI maintains a list of translations of its basic documents, as the European workshop on Learning technologies is maintaining translations of the LOM specifications.

At last, namespaces is a fundamental part of the infrastructure of the Web, and which are used in metadata queries through the Web. They are a part of the modular metadata systems. For example, the declaration of namespaces in the Web in a URL may be specified by a prefix:

- *dc* to indicate a search only in the Dublin Core metadata element structured sets, while
- *lom* recognises the IEEE-LOM metadata element set, searching only for data, listed in IEEE-LOM metadatasets.

3.3. *The Universe of Discourse (UoD)*

All human beings observe the reality with their senses and create a model of the reality in their mind. In creating digital geographic data, it is attempted to model and describe objectively the real world phenomena that have a relation to the earth' surface, using IT concepts in geographic and spatial computer analysis and graphic display. Any description of the reality is always an abstraction of the real world objects, always partial and always just one of the possible real world 'interpretations'. This are not an exact duplication of the real world: some things are approximated, others simplified, and some are ignored. There is no such dataset as perfect, complete and correct, or it would be the reality itself. To ensure that the dataset is not misused the assumptions and limitations affecting the dataset must be fully documented in the metadataset.

The Universe of Discourse (UoD) defines phenomena in the reality into an abstract form, by theme, position and time in order to make these objects intelligible to represent them in a database. The UoD applies an ontology – in GIS usually understood as 'the definition, classification and structure of objects considered for a specific database application' – as a limitative collection of unambiguously defined concepts [Uitermark 2001], following the Greek philosopher's Parmenides (500–450 BC) doctrine on ontology and the Indian Nyāya–Vaiśesika school of philosophy on the classification and structuring of objects. In this way, the UoD contains the limitative set of characteristics of selected objects in the real world and their descriptions. However, real world objects are modelled into an object-based database, in

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 7

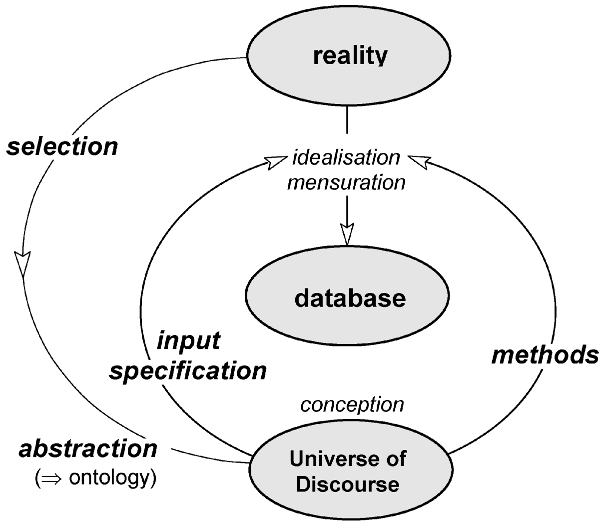

Fig. 1. From Universe of Discourse to Database. From reality the intended objects are selected and abstracted resulting in an object definition (database concept); with the input specification and the methods to be used (resulting in the surveying rules), the surveyed data are stored in the database.

which for each real world object a one-to-one representation. So, the process for defining the UoD is decomposed into two steps:

1. conceptualisation including the selection of what should be considered in the real world and the abstraction of the selected objects, resulting in a dataset content specification;
2. mensuration, specifying the measuring methods and equipment, the measurement requirements and the capturing and storage of the data itself.

Since the UoD (see Figure 1) includes any selection of objects from the reality as seen by the-potential-user, his (the customer's) satisfaction is determined by the perceived performance and expectation of the data for the user's application. This gives a product image, both emotional and cognitive factors and can be recognised in the process of the customer's satisfaction contributing to the customer's profitability of the data.

In America, often the term 'abstract view of the universe' is used for the concept of the Universe of Discourse. However, the 'abstract view of the universe' does not reflect the selection process, which is an integral part of the definition of the UoD. ISO/TR 9007, 1987 [ISO 1946] defines the UoD as 'all those objects or phenomena of interest that have been, are or ever might be'.

3.4. *Application of metadata*

No single metadata element set will accommodate the functional requirements of all applications, and as the Web dissolves boundary access, it becomes increasingly important to be able to cross discovery boundaries. This is served by:

- *conceptual models* – defining the UoD – needed to structure the representation of the semantics and syntax, and to replace the textual descriptions or lists of feature and attribute types, which are often found as metadatasets. Though lists of feature and attribute types give easily the impression of exact information, they often lead to misinterpretation when not applying the definitions of used terms. So, conceptual models limits to those parts and characteristics of the reality, which are relevant for the particular information system. A conceptual model names the object types in the UoD, the attributes of the object types and the relationships between the object types and the attribute domains (and the behaviour of the object types in an object-oriented model);
- *syntax and semantics*. Syntax is about form; semantics is about meaning. Agreements about both are necessary for sharing metadata; however, it is important to make a clear distinction between syntax and semantics. Where the syntax defines the manner of the transfer of data, the semantics will tell what is transferred. In addition, there is a third component for interoperability, that is beyond syntax and semantics: *content vocabularies*; this can be in an open and unconstrained natural language as Chinese, English, Japanese, Russian, Spanish, etc., but may also be a specific controlled vocabulary further narrowing down the scope and increase the precision of a description;
- *application profiles*. An application profile is a combination of metadata elements selected form one or ore metadata schemas, expressing the principles of modularity and extensibility, to tailor to the functional requirements of a particular application. Application profiles achieve this modularity through the following mechanism:
  - *cardinality*, referring to the constraints on the appearance of a metadata element (being mandatory, optional or conditional);
  - *value space restriction*, for specific applications differing from the general specification;
  - *relationship and dependency specifications*, by defining interrelations between data elements and their value spaces for a specific application;
  - *declaration of namespaces*, when supporting multiple namespaces.

One of the first steps in designing a database is the data analysis based on *user requirements*, based on how users perceive their application. The result is a set of external views representing exactly the data needed. External views represent each separate application of the information system; the conceptual schema is the combination of all external views of the information system. However, databases and their applications have life cycles and it is most probably that new ideas for data products appear later. Such new external views shall be derived to include in the conceptual model.

3.5. *Metadata element sets*

Metadata standards provide a mean to describe the content of a dataset at different levels of detail. The most common, and most vaguely defined metadata element is an *abstract* or an *overview*, providing a brief summary of the dataset's content, usually expressed in natural language without any structure. Although it gives the supplier the maximum freedom conveying the dataset content, its usefulness is fully depending on the evaluation by the data users. Because the natural language allows ambiguous and inconsistent expressions – despite of a defined grammar – it creates fuzziness and confusion requiring formal models for computer interpretation of metadata.

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 9

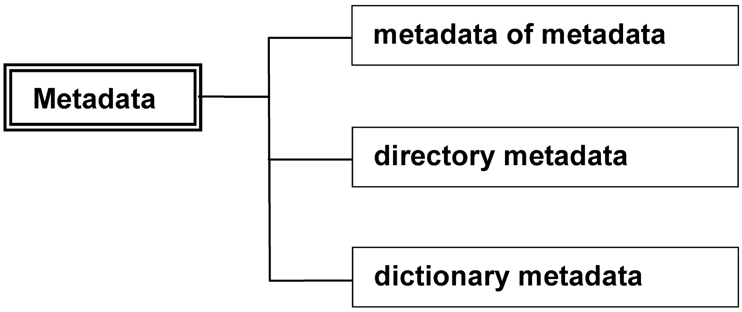

Fig. 2. Types of metadata.

More structured as free text is a list of objects types and their attributes, including hierarchies and object and attribute type definitions. A further approach is to provide a conceptual schema in a formal way. This leads to the necessity to list and define metadata elements. Metadata elements can be distinguished in three types, together forming the set of metadata elements (see Figure 2):

1. *content* of metadata (metadata of metadata). In order to understand the metadata of a dataset it is required to state the characteristics of the metadata such as the language and coding system of the metadata itself as well as the reference and coding systems that the metadata uses to define the thematic, geometrical and temporal extent;
2. *directory* metadata including the identification of the database, as name and origin as well as locating specific type of data as names and addresses of owner, distributor and manager and the description of the extents of the database in semantic, geometric and temporal sense, and quality and security measures, etc.;
3. *dictionary* metadata including the data definition, referring to semantic, geometric and temporal definitions of the data, data organisation by conceptual schemes for homogeneous datasets and quality information including the quality conceptual schema and the description of the parameters for the data as idealisation, accuracy, reliability, currency, completeness, consistency and currency.

3.5.1. *Metadata of metadata*

In order to understand the metadata of any dataset, the organisation of the metadata, the language used and the reference system for semantic, geometric and temporal extent should be given. (Metadata may refer to a homogeneous dataset series, a dataset or a set of occurrences of attributes of entities and/or relationships.) These characteristics can be given referring to other standards or by a description:

- *character set* used for encoding the data in the metadataset (using, e.g., ISO 8859-10);
- *language* for textual statements in the description of geographic datasets may follow a coding system as defined by ISO 639 or can be given by a text like ‘Chinese (Hong Kong)’, ‘Chinese (Singapore)’, ‘Chinese (Taiwan)’, ‘English (Australian)’, ‘English (UK)’, ‘English (US)’, ‘French (Standard)’, ‘French (Belgium)’, ‘French (Canadian)’, ‘German

(Standard)', 'German (Swiss)', 'Portuguese (Standard)', 'Portuguese (Brazilian)', 'Russian (Moldova)', 'Russian (Russia)', etc.;
- *date description*, e.g.: by EN 28601, ISO 10303, part 41. The advantage of using ISO 10303 part 41 is that it also standardises the description of persons, organisations and addresses;
- *geometric reference system* applying geodetic standards as well for planimetric geometry, as for the vertical component;
- *quality element definitions* (Morrison, 1995); e.g., how to describe the original intended purpose of production by the producer, the usage prior to the present intended use, the lineage of the dataset describing the process history, the thematic, positional and temporal accuracy, the completeness and consistency of object types and the homogeneity of the given quality elements;
- *administrative data* about metadata, being:
  - *point of contact* in the organisation that created or manages the metadata;
  - *dates of creation*, last check and update future updates of the metadata;
  - *constraints and security* measures of the metadata.

The metadata language should be a mandatory element. The positional reference system whether it be geodetic (using co-ordinates) or non-geodetic (as addresses including postal codes, parcel identification, road networks identification, administrative subdivision of a country, etc.) are required when geometric data is given in the dataset. Here, conditional metadata may also exist: a dataset with geographic information may contain semantic geographic data only where no positional data is available (then no positional reference systems need to be given in the metadata).

3.5.2. *Directory metadata*

The elements of metadata to find the location of the dataset at least comprise:

- *identification* of the dataset by a code or a name, that defines the dataset uniquely and clearly amongst other datasets including the version number of the dataset when appropriate. Besides, other names can be given as alternative title(s). This may be the dataset's name in other languages or an abbreviation of the dataset name. Also the dataset overview, giving an overall description of the dataset including a summary (in text) and describing the content of the dataset can be given;
- *administrative* metadata to acquire a dataset information, regarding where and how the dataset is held as well as its procurement is given including:
  - *organisation*: name and abbreviations (if available) of the organisation, the address (postal address, visiting address, telephone number, telefax number and/or electronic-mail address, home page address on the web;
  - *organisation role*: as responsible authority for the dataset, producer organisation, etc.;
  - *point of contact*: in many cases it is evident that personal contacts lead to the best way of transferring data from provider to user. In such cases, direct contact between persons is necessary. Though personnel basis is an extreme aid for mutual use of data, the organisations should still appoint officials for this purpose. Here the name, function and addresses of the point of contact can be given;
  - *distribution* includes descriptions of restrictions on use, copyrights, units of distribution (e.g., per tile, per square kilometre, per administrative unit, etc.), pricing and discounts of

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 11

the data (types) per unit, data media on which the dataset can be recorded and retrieved or on-line accessed, delivering data formats, procurement, giving instructions for ordering the data and the delivery service and services for processing the data;

- *origin* of the data:
  - *producer's purpose* of the dataset, describing the original purpose by the producer for which the data in the dataset was captured. It may include the original intended application scale. Much of the digital geographic data is collected for direct data presentation on maps. The content definition of the dataset (selection and abstraction in the definition of the UoD) is usually based on what should and can be represented on those maps and is very dependent on the original purpose of the dataset;
  - *capturing method* and type of semantics describing the way of original collection and the type of data that can be found in the dataset, e.g., data from space, aerial data, landsurvey data, etc.;
  - *potential usage* give the provider's view on the potentials of the dataset, i.e., the possibilities of the data for different applications. In addition, the usage can be described to give the future users an idea for which applications the dataset was already used. Important in these cases is also to give an impression of the successes and failures that are experienced in the specific applications of the dataset. References to other relevant published documents or public available additional documentation may be provided;
  - *type of spatial schema* used in the dataset. This may be a standard or a user defined spatial schema, describing the main characteristics and components of the schema. On top, the definition of the spatial reference system should be given in the case of use of a spatial schema. The metadata may also describe geographical data that does not contain any spatial reference. In this case no spatial schema has to be defined in the metadata;
  - *examples* taken from the dataset and being representative for the whole dataset can be provided as samples, e.g., as a browse graphics either in raster or in vector format;
  - title and/or code and provider of *related datasets* of possible interest to a potential users can be enlisted for further information about the possible uses of the dataset;
- *dataset quality* elements. The dataset quality describes the difference between the actual dataset and the user's and producer's UoD. By describing the quality of the dataset, the user can determine whether the data has potential for the intended application before the transfer of the data. The dataset quality elements follow the designed quality conceptual model and comprises:
  - *source* describing the producers'-, owners'-, managers'- and providers' organisation as well as the point of contact within the organisation;
  - *homogeneity,* giving a description of how well the data in the dataset follow the overall uniformity of the data;
  - *usage* gives an overview of the applications for which the information in the dataset has been used previously and how well the data fitted in these applications. Also the *potential use*, is an important aspect in these which gives an indication of the possibilities of the data seen from the providers' perspective;
  - *lineage* giving a description of the origin of the information contents in the dataset and everything that has happened to the data since, until the moment of transfer;
  - *quality parameters* as positional accuracy, thematic accuracy, temporal accuracy, logical consistency, completeness and their values;

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 12

  - *currency*, dating the information in the dataset as well as indicating how well the data is kept up to date;
- *extents*. Data about geographic objects are positioned in space and time. In order to judge for the user whether a dataset is suitable for the intended application the planar, vertical and temporal extent of the data in the dataset should be available. This can be done by one or more 'bounding range(s)' of the dataset giving the maximum and minimum coordinates appearing in the dataset or by one or more 'area boundary(ies)' delimiting the area(s) covered by the dataset. In addition, the currency of the extent, indicating the status, completeness and validity should be given. For the temporal extent an indication of the range (from start to expiring date) should be given (this may still be 'on-going', i.e., having no expiring date yet), but also a descriptive text may indicate such as 'medieval period', '20th century', 'annually update', 'continuously update', 'last update', etc.

3.5.3. *Dictionary metadata*

The dictionary data describe the semantics of the data in the dataset as well as the conceptual schema that has been used for data modelling. This allows search engines to not only access the data in the dataset but also interpret the semantics for determining the requested information and gives the user access both to the data and its semantics. Amongst others, dictionary metadata consists of the following elements:

- *data definition*. Objects may be defined in different ways in different datasets. To enable comparison between them, data describing the characteristics of objects are required. Data definition and classification, describing the differences between classes in order to distinguish between objects in different classes and to define the relationships between the classes do this. For this purpose, data is provided giving the definition of object types, attribute types and relationship types, wherever they exist in the dataset. Together with the application schema, all objects that are represented by data in the dataset should be defined and become understandable for the receiver of the data. Sufficient description of the data definition is given by:
  - *object type*, attribute type and relationship type name and definition should be given completely. If object types, attribute types or relationship types are indicated in the dataset by codes, also the coding should be part of the definition. Relationship types should also indicate the object types they relate (a 'from' and 'to' object type);
  - *classification*. The object, attribute and relationship types of a dataset can be described in a classification system, in which the object, attributes and relationship types may take part in a given hierarchical organisation. These hierarchies should be given in the form of a thesaurus according to the standard ISO 2788: 1988. In the case of using a thesaurus the name, version number or version date, the thesaurus administrator's point of contact and references to other publications about (use of) the thesaurus should be available to the user. If not all the thesaurus elements are used in the classification system, then those thesaurus elements that are used shall be listed in the classification part of the directory data within the metadataset. When a thesaurus is provided in the metadataset, besides the definition (if it does not belong to a 'de jure' standard) of each thesaurus element, also the code and synonym should be given; related, narrower and broader terms as well as pictures can accompany this;

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 13

– *spatial reference system*. The spatial reference system may be a direct, geodetic reference system based on co-ordinates, or a non-geodetic reference system, e.g., based on addresses or others. A dataset can contain objects referenced by a multiple of geodetic and/or non-geodetic spatial reference systems. Elements for the definition of the non-geodetic reference systems are: type (e.g., the system in which the references are given as country, county, municipality, etc.), title and owner and version date of the non-geodetic reference system. For geodetic reference systems the name of the reference system, the co-ordinate representation and its units (for both planimetry and height), the name of the datum, ellipsoid, map projection, etc.

4. Minimal metadataset

Minimally a dataset should have some metadata to stay understandable for any enquiry by users. It is noted that elements have descriptive text to convey a common semantic understanding of the elements. However, controlled vocabulary may be required to promote global understanding of the element values.

The development of the ISO 19115 Metadata standard has including a number of studies on different metadatasets available in different regions and organisations such as the ANZLIC Metadata: Core Metadata Elements (1995), Canadian Directory Information Describing Digital Geo-referenced Data Sets (1994), CEN ENv 12657, Standard for Geographic Information – Metadata (1996), FGDC Content Standard for Geo-spatial Metadata (1994), DIGEST (1994), IHO S 57 (1995). The system enables international users to locate datasets of interest, and contact the appropriate National Mapping Agency (NMA). Comparison to the from origin bibliographic Dublin Core Metadataset (that has much experience with metadata to structure the vast amount of publications for consulting), one can conclude for geographic dataset to the following minimal metadataset:

- *identification* of the dataset by a name given to the dataset by the original producer or publisher, being unique to distinguish the dataset from others. This might also be done by an unambiguous code for the dataset;
- *provider*, the organisation responsible for making the dataset available in its present form, such as a publishing house, topographic service, private company, municipality, etc.;
- *original producer*, the organisation responsible for the original capture of the data in the dataset;
- *reference system* for spatial reference, semantic definition by thesaurus and time definition;
- *extent*, describing the geographical, semantic and temporal coverage of the valid data in the dataset and whether or not spatial attributes are carried in the dataset. The geographical coverage may be a description by a set of quadrangles or geographically defined circumferencing boundaries. The semantic coverage may be a list of the type of objects by semantic attributes and the temporal coverage indicates the time period of the valid data in the dataset;
- *date*, indicating the date of validity of the data in the dataset. Many ways of writing dates are possible but if used they should be written in a clear and unambiguous manner;
- *language* of the metadata;

- *syntax* of the dataset transfer, to identify the software that will be required for reading the data;
- *quality*, describing the spatial, semantic and temporal quality parameters for the dataset. In the quality definition of larger datasets information should be available of the meta-quality indicating the quality indicators, e.g., for spatial quality a relative standard error should known as 'average', 'maximum', 'minimum', 'expected', 'required', etc.;
- *rights and management* indicating the copyrights, constraints on use, way of distribution (e.g., by tiles, by square kilometres, etc.).

5. Metadata query and distribution services

There are also many ways to transmit, communicate and present metadata. Metadata elements will be valued in different ways by different users or by one user for different tasks. The Internet and other technologies are causing rapid change in means to provide information although the need for metadata on physical means – including paper – will continue to exist.

The web has become an important universal information tool, embracing vast stores of information with many purposes, multiple disparate sources, and quite a few unpredictable users. There is a clear need to improve access to this mass of information and for the development of better search, retrieval, and organisational tools. Metadata is a fundamental part of the solution to these challenges. Browsing the web to find data resources three levels can be distinguished (see Figure 3):

- discovery metadata to find data resources (a front office application);
- inventory metadata to select the datasets, which will be acquired (also a front office application);
- data model access to enable the interface to load the geographical data into the application using the metadata that accompanies the dataset received from the data provider (a back office application).

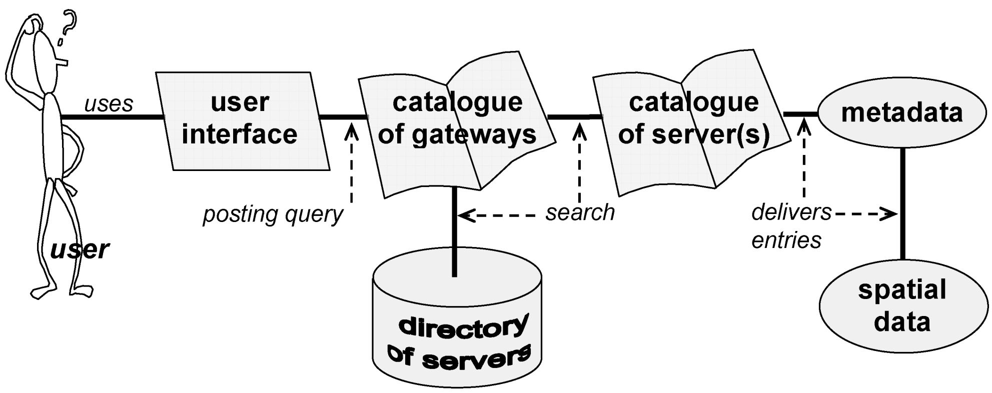

Fig. 3. Finding spatial metadata through a catalogue of gateways and servers to find the spatial data itself.

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 15

5.1. *Discovery Metadata – Metadata to find data resources*

Effective use of metadata in finding data on the web, requires three things:

1. set of commonly-understood terms to describe the content of information resources (semantics);
2. standard grammar for connecting those terms in meaningful metadata sentences (syntax);
3. framework that allows us to transfer and recombine those metadata sentences across different applications and subjects.

These three elements – standardised semantics, a definitive syntax, and a framework for transfer – provide architecture for resource description that can work across all subject areas on the web.

5.2. *Inventory Metadata – Descriptions of datasets that could be ordered*

In its simplest form, a central database is maintained, with data providers updating dataset descriptions (metadatasets) as necessary. Data providers are provided with a software update tool, which includes the current version of the central database. They can update information about existing datasets or add new datasets before returning the media to the central, when the new data are then transferred to the main central database. On the other hand, a Spatial Data Infrastructure (SDI, see Figure 4) may link the centre to all metadata providers [Williamson 2003]. Locally, the providers will maintain their datasets (both the metadata and the data itself). Users may consult the metadatasets for inventory by using the appropriate web browser,

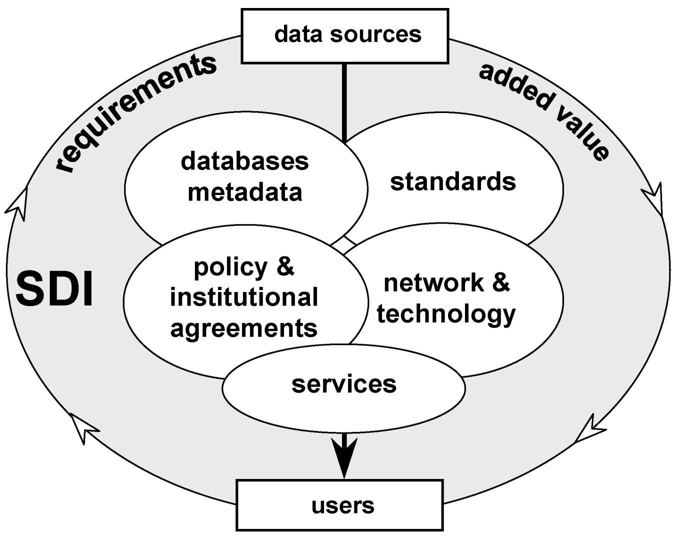

Fig. 4. Interactions between SDI components for the transfer of data to users: databases with metadata, standards, policies and agreements, network and technology and services. By using the data more experience is gained, allowing providers to improve the data quality.

protocols and standardised catalogue services to support technologically the interpretation of the (meta-)data.

Web pages are derived automatically from the database, and the information is made available free of charge to all Internet users. In a way, it can be described as an Internet gateway dedicated to Geographic Information. Regional and national examples of these are:

- MEGRIN Geographical Data Description Directory service in Europe, as well as the European Spatial Metadata Infrastructure (ESMI);
- Asia-Pacific Spatial Data Infrastructure (APSDI) as proposed by PC GIAP;
- Inter-American Geo-spatial Data Network (IGDN) in Latin America and the Caribbean;
- National Geospatial Data Clearinghouse in the U.S.A., following the EO 12906;
- Canadian Geospatial Data Infrastructure GEONet, developed by GeoConnections;
- Australian Spatial Data Infrastructure Distribution Network;
- the Japanese GSI clearinghouse gateway system with multi-byte character code sets using the GEO profile based on ISO 23950.

5.3. *Model metadata – Metadata for data transfer*

Most of the present transfer standards contain metadata on top of the geographical data that are transferred. Examples are the CEN ENv 12657: Geographic Information – Data description – Metadata, issued 1998 and the ISO 19115 Geographic information – Metadata issued in 2001. Apart from the used terminology in the standards, they also give the metadata elements and a schema, which, when properly implemented by a data producer, will enable users to locate, access, evaluate and employ geographic data. Users will be able to discover, select, and purchase geographic data; determine whether data in a holding will be of use to them, and be able to apply geographic data in the most efficient way by establishing a common set of metadata terminology, definitions, and extension procedures. Supplementary benefits of these standards for metadata are to facilitate the organisation and management of geographic data: it furnishes data producers with appropriate information so they can properly characterise their geographic data. It facilitates dataset cataloguing enabling data discovery, retrieval and reuse.

The standards models reflect their period in time for development. The ENv 12657 uses EXPRESS and EXPRESS-G as a modelling tool, while the ISO 19115 applies UML (Universal Modelling Language). These models are an integral part of the entire respective abstract models for geographic information. The metadata model diagrams are presented in computer readable text and their graphic representations in view. The model diagrams provide a good illustration of the classes and their attributes, and the type and cardinality of the relationships between classes for geographic metadata.

6. Experiences in the use of metadata

In the use of metadata some striking results were obtained; some examples given by [Ahonen-Raino 2001], [McCelland 2002] learn the following:

1. metadata are incomplete and contain errors. As users of data rely on descriptions given by suppliers, users with unexpected applications look for information in the metadata that is

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 17

not available because the UoD did not foresee in the apparent application. For the user it appears as if the metadata is incomplete or contains errors, while the supplier may have listed the metadata correctly according to his UoD;

2. specialised vocabularies complicate data import and understanding. Localisation of metadatasets often implies specialised vocabularies, which are not understandable for many users outside the geographical area or outside the field of application. Suppliers do not often realise that users may use data for unexpected applications. The use of metadata for non-experts leads often to the Lego™method of building and children are not biased by the suppliers intention and may apply their play constructions in an unexpected water or desert type imagination. Non-experts users may apply the same philosophy which in their situation leads to inaccurate metadata descriptions;
3. in general standards are evolving together with the development of applications. Metadata specifications are not different. Therefore, conceptual models should follow these new applications.

In the past, many of the developments of metadata standards were focussed on the coordinate based geometry and raster geometry played a minor role, as is also correct for the ISO 19115 Metada standard. Therefore, in May 2003 ISO/TC 211 has decided to set up a project to develop ISO 19115-2 *Geographic information – Metadata – Part* 2: *Extensions for imagery and gridded data* . At the time of the press of this book the project was at WD stage (Working Group draft).

7. Regional and national efforts developing metadata

In some regions, specific developments have led to the use of metadata standards. Some of these were national oriented, while other were developed by some regional countries or regional bodies and accepted by its members. Part 2 of this book will describe these in detail.

7.1. *Europe*

The first attempt for the development of standards in the field of Geographic Information took place in Europe by the Technical Commission 287 on Geographic Information of the European Committee on Standardisation (CEN/TC 287) from 19991 until 1999. Its basic objective was to enable geographic information to be accessed by different users, applications and systems, and from different locations. This requires a standard way of defining and describing this information, a standard method for structuring and encoding it, and a standard way of accessing, transferring and updating via geographic information processing and communication functions, independent of any particular computer system. Suppliers and users of data and developers of GIS's and GI-applications may use the CEN family of standards to enable databases and applications that are different in structure form and content to interconnect and inter-operate.

The CEN ENv 12657, Geographic Information – Data description – Metadata is widely accepted in Europe and many countries use the standard for the creation of metadata in developing a National Spatial Data Infrastructure. Also pan-European projects as the Geographical Data Description Directory (GDDD), Geo-Scientific Electronic Information Exchange System (GEIXS), Added Value Information Dissemination for hydrographic Datasets (AVID),

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 18

European Spatial Metadata Infrastructure (ESMI), etc. apply the ENv 12657 for the dissemination of data (see the European chapter in part 2 of this book).

Especially in Europe the multi-lingual situation emerged in projects dedicated towards dealing with this issue like the Electronic Trade for Geographic Information (GISEDI), Methods for Access to Data and Metadata in Europe (Madame) and CLEF (Cross-Language Evaluation Forum); the last one attempting to develop information retrieval system that focuses on non-English approaches including Dutch, Finnish, French, German, Italian and Spanish. The Open Archive Forum supports projects and national initiatives using open archiving approach as an interoperability framework implementing Open Archives Initiative (OAI) metadata protocol in Europe.

Many countries in Europe have implemented the CEN ENv 12657 metadata standard; with the approval of the ISO 19115 Metadata standard, they have to decide to transform the metadata information into this new standard and introduce this into the geo-spatial data infrastructures.

Because of the developments of standards for Geographic Information within ISO/TC 211, CEN/TC 287 has ceased its activities in 1999. Facing the regulations of CEN, making ENv's mandatory for CEN-members to use, CEN/TC 287 has been revived in November 2003 to harmonise the existing ENv's with the new ISO 191XX suite of standards in the field of Geographic Information. Meanwhile CEN ENv's are abandoned and the acceptance of the ISO ISO 191XX standards within CEN is under a voting process. Due to this voting the following ISO standards have been accepted as European standards:

- EN-ISO 19101: 2002 Geographic Information – Reference Model;
- EN-ISO 19105: 2000 Geographic Information – Conformance and Testing;
- EN-ISO 19107 Geographic Information – Spatial Schema;
- EN-ISO 19108: 2002 Geographic Information – Temporal Schema;
- EN-ISO 19111: 2003 Geographic Information – Spatial Reference by Coordinates;
- EN-ISO 19112: 2003 Geographic Information – Spatial Reference by Geographic Identifiers;
- EN-ISO 19113: 2002 Geographic Information – Quality Principles;
- EN-ISO 19114: 2003 Geographic Information – Quality evaluation procedures;
- EN-ISO 19115: 2003 Geographic Information – Metadata;

while the following are still in process:

- prEN-ISO 19116 Geographic Information – Positioning services;
- prEN-ISO 19125-1 Geographic Information – Simple feature access, Part 1: Common architecture;
- prEN-ISO 19125-2 Geographic Information – Simple feature access, Part 2: SQL;
- prEN-ISO-TR 19120 Geographic Information – Functional standards;
- prEN-ISO TR 19121 Geographic Information – Imagery and gridded data.

Based on environmental issues in the EC FP-6 programme, the INSPIRE project was initiated in 2001–2003. The project aims to enable communication between users in a meaningful manner, despite the existing differences in language, context and content of spatial datasets in Europe. The project has two phases: a design phase (2004–2007) and an implementation phase (2007–2015) (see also chapter 1 of Part 2 in this book). INSPIRE intends to apply the ISO 191XX suite of standards.

7.2. *North America*

Between Canada and the U.S.A. a joint committee provided a forum to develop consensual positions, which then can be ratified by the main groups of users and providers. This process is being followed in both countries and has resulted in an identical Canadian and U.S.A. national metadata standard, becoming the North American profile of ISO 19115. In addition, there have been informal discussions between the United States and Canada on a common North American strategy to adopt and implement other ISO/TC 211 standards. Now, the United States has developed a Memorandum of Understanding (MoU) with Canada that will outline procedures for ANSI/INCITS L1 and SCC to develop common profiles of ISO/TC 211 standards.

Both, Canada and the U.S.A. are member of and actively participating in the Permanent Committee on Spatial Data Infrastructure for the Americas, which also contributes to the development of the Global Spatial Data Infrastructure.

7.2.1. *United States of America*

The Federal Geographic data Committee (FGDC), founded in 1990, has the task to develop geo-spatial data standards that enable the sharing of spatial data among producers and users and support the National Spatial Data Infrastructure (NSDI). FGDC working groups and sub-committees develop standards in the field of data transfer, metadata, cadastral data, ortho-imagery, and positional accuracy in consultation and co-operation with state, local, tribal, private academic and international communities through a structured open consensus. The FGDC sub-committees and working groups establish and implement strategic guidance and specific actions that support improved collection sharing, dissemination and use of geo-spatial data, and by doing so, contributing to the development of the NSDI.

The Metadata working group is established to promote an awareness of the metadata dimension to geo-spatial data, among FGDC members and geo-spatial data users. Presently most work of the working group deals with the harmonisation of the FGDC metadata standard with the ISO standard 19115 for Geographic Information – Metadata on order to adopt the ISO standard as an American National Standard. Proposals have been discussed for changes in the FGDC metadata standard, which may not be contained in the SDTS transfer set. Work is done or underway for profiles of the FGDC metadata standard for biological data, shoreline data, remote sensing data.

The U.S.A. is also member of Pan American Institute of Geography and History (PAIGH) and provides technical assistance, conducts training at research centres, distributes publications and organises technical meetings in the areas of cartography, geography, history and geophysics.

As consequence of the Executive Order 12906 of April 1994, to establish a National Spatial Data Infrastructure, also state, local and tribal governments are involved in the implementations of NSDI. In addition, in the U.S.A. three cross-sector stakeholders play also an important role in the development of NSDI and the use of metadata:

1. OpenGIS® Consortium, an international non-profit member consortium developing interface specifications for geo-spatial services;
2. I-team Geo-spatial Information Initiative, an interdependent partnership of federal, state, local and tribal authorities, academia and the private sector for interoperability implementations;

3. GeoData Alliance, a non-profit organisation promoting the creation and effective use of geo-spatial data.

The Geo data Alliance and the I-teams Initiatives will bring all geo-spatial actors in the U.S.A. together to improve the national co-operation in the field of GIS as part of the Geo-spatial One Stop electronic government.

The FGDC Metadata Content Standard for Digital Geospatial Metadata, STD-001-1998 Version 2.0, is well distributed and applied throughout the America's and the Caribbean by workshops to disseminate the information about metadata and clearinghouse implementation.

7.2.2. *Canada*

Since 1997, the Canadian GeoConnections develops the Canadian Geo-spatial Data Infrastructure (CGDI) co-ordinating numerous Canadian databases with geographic information and make them accessible through a common window on the Internet in order to promote partnership among federal, provincial and territorial governments, the private sector and the academic community. Geographic or geo-spatial information is seen as a key driver of the transition to a knowledge-based economy in Canada. In this case, everyone will benefit from having access to (geo-spatial) data and information.

The CGDI has seven GeoConnections programmes:

1. through cost-shared development, GeoExpress-Access provides the private sector, Canadians and government agencies *access* to government geo-information via a common Internet system;
2. national geo-information *framework* (foundation) with diverse data from various organisations, upon which other databases will be integrated or constructed;
3. *federal-provincial/territorial partnership* program to facilitate data sharing, and guide the vertical harmonisation of information sharing and access to geo-information holdings. The GeoConnections Secretariat co-ordinates federal/provincial/territorial ideas, programs and project activities with private initiatives;
4. *industry-partnership* (GeoInnovations) to accelerate the development of promising and advanced geo-information technologies by investments in the private sector to develop new tools and innovative technologies;
5. strengthen Canadian *communities* capacity of to effectively plan and manage their economic, environmental and social development using geo-information and services through selected pilot projects;
6. *National Atlas of Canada* to provide national perspectives on Canada's physical, environmental, economic, social and culture issues for students at all levels and the public in general. Examples of issues include climate change, industrial development and population;
7. *geomatics skills-matching network* fostering knowledge-based industry growth, with an Internet-based information clearinghouse to disseminate cross-jurisdictional information and facilitate employment.

7.3. *Asia/Pacific*

The Permanent Committee on GIS Infrastructure for Asia and the Pacific (PC GIAP) has proposed the Asia-Pacific Spatial Data Infrastructure (APSDI), being a network of databases, located throughout the Asia-Pacific region, that together provide the fundamental data needed to achieve the region's economic, social, human resources development and environment objectives. APSDI Data Node is a network of nodes established to distribute, manage, and maintain fundamental data in each country connected to APSDI. According to the project plan, Australia, China, Japan and Iran will act as demonstration data nodes in the first stage. In each country, a sub-node of the APSDI is planned and they will be linked with each other.

7.3.1. *Pacific*

The co-ordinating Spatial Information Council for Australia and New Zealand (ANZLIC), originally established in January 1986 as the Australian Land Information Council, aims to co-ordinate a widening range of activities and objectives to make land and geographic information readily available to different users. In particular, ANZLIC promotes the availability of land and geographic information for land resource management, environmental monitoring and economic development for both the public and private sector. A key concept of the Council is that each delegate represents a co-ordinating structure within his or her own jurisdiction and that close links are maintained with other co-ordinating bodies such as Intergovernmental Committee on Surveying and Mapping, the Public Sector Mapping Agencies, Registrars-General and Valuers-General.

Version 1 of the ANZLIC Metadata Guidelines was published in 1996 followed by an associated paper 'Recommended Guidelines for the Transfer of ANZLIC Metadata Core Elements'. These reference documents have been well received by the geo-spatial industry. However, with use, some inconsistencies and deficiencies have been identified, and some additional metadata elements have been defined. Consequently, in November 1999 the ANZLIC Metadata Working Group initiated a minor review of the ANZLIC guidelines. The review was seen as an interim measure pending the development of the international standard 'ISO 19115, Geographic information – Metadata'. A further revision of the ANZLIC guidelines is likely to be undertaken once the international standard becomes stable.

7.3.2. *Asia*

The Asian region is geographically and culturally diverse. Because of logistical reasons not many common project in the field of geo-spatial metadata have developed; most developments are country wise undertaken.

However, in the Asian region there are two organisations encouraging metadata initiatives, which are the Permanent Committee on GIS Infrastructure for Asia and the Pacific (PC-GIAP) and the Technical Committee 211 on Geographic Information/Geomatics of the International Organisation for Standardisation (ISO/TC 211). PC-GIAP, has 55 country members extending an region from Armenia, Iran, Azerbaijan and Kazakhstan in the north-west to New Zealand, French Polynesia, American Samoa and Cook Islands in the South East, and has the focus on SDI development in the region, including the Pacific area. The metadata work is part of the Date Node Project that started in 2000.

Although it was not in the scope of ISO/TC 211, developing ISO geo-spatial standards in the commission has brought together many experts and so created a forum in the Asian region for information exchange about metadata projects stimulating national developments based on ISO 19115.

moellering v.2005/06/23 Prn:2/08/2005; 15:09 F:moellering01.tex; VTEX/J p. 22

7.4. *Central and South America and the Caribbean*

The region is very active in the developments of metadata as part of SDI developments. Four regional organisations stimulate these developments: PCIDEA, PROCIG, PAIGH and DIGSA.

In developing SDI's, the Permanent Committee on Spatial Data Infrastructure for the Americas (PC IDEA) is taking advantage from the lessons learned elsewhere in the world. Still, it has to face the differences in culture and politics, resulting in specific conditions for each country in the region. Therefore, two sub-regional committees are founded: for the Caribbean and for Central America. In addition, two surveys have been carried out identifying the key actors in the region. Although many countries in the region apply the FGDC metadata standard, the ISO19115 metadata standards is considered more and more.

The Central American Development Project for GIS (PROCIG) – the project ended June 2001 – promoted the integration of statistical and census data with other geographic information for public dissemination. Additionally, PROCIG supported national SDI activities, implementing metadatasets and clearinghouse nodes as well as training and consultancy to strengthen the national knowledge about SDI.

Pan American Institute of Geography and History (PAIGH) is an old institution in the Americas, dating back from 1928. Nevertheless, it encouraged the Inter-American Geo-spatial Data Network for the developments of a regional SDI for the Americas by USGS/Eros Data Center and USAID. Besides, PAIGH also co-financed a project called Atlas of the Americas to develop tools for sustainable decision-making. As part of this project in Argentina, Costa Rica, El Salvador, Mexico and Peru a SDI with metadata of existing datasets in the SDI was created.

The Directory of Geographic Institutes from South America, Spain and Portugal (DIGSA) is a policy-makers summit of national mapping institutions in South America, Spain and Portugal. The summits eases information exchange between members including information about SDI and metadata, although the focus is on licensing, copyrights, training, data production, updating policy and product quality.

Many of the application scenarios in the region are oriented towards the Environmental Decision-Making and Disaster Management due to natural catastrophes regularly returning in the area, as the hurricane Mitch in 1999.

7.5. *Africa and the Middle East*

Africa and the Middle East are entering a period that will show a rapid growth in capturing and publishing spatial metadata as a part in the developments of spatial data infra structures, although at present many of the NMA's stay with the traditional production of paper-based maps. However, the initial efforts to develop SDI, coinciding with the completion of the ISO 19115 Metadata standard, have resulted in an ISO equivalent development and working groups have been established for the standardisation of metadata and for creating a sound geodetic network for Africa.

The United Nations Environmental Programme (UNEP) has, in collaboration with the Regional Remote Sensing unit of the Southern African Development Community in Harare (SADC), drafted a strategy to develop the State of Environment metadataset for South Africa and will soon develop the same for West Africa.

8. Summary and conclusions

The applications of GIS has been globalised in the past and its cross-boundary use is demonstrated by applications as:

- environmental, forestry, marine and cultural protection analysis, as well as disaster regulations programmes, etc., applying GIS analyses in the future Virtual GIS environment using distributed databases and software programmes;
- decision support systems for policymakers in land, and socio-economic planning;
- location based services using GIS for route finding, building constructions, precision agriculture, etc.;
- and many more.

In order to make use of the many spatial datasets in such situations, the geographic information scene has realised to make well-equipped standard metadatasets enabling access to these datasets to the public. Even for the experts in the field, it was recognised that modular standards need to be developed for data dissemination and use. This has lead to different developments of metadata standards in several regions, as:

- in the USA, the FGDC standard Version 2.0 is developed in 1998 and applied in many American countries;
- in Europe, by the CEN/TC 278 designed the ENv 12657 Metadata standard in 1999, which is wide spread applied for geographic data discovery in national and regional SDI's and many project funded by the EU;
- in the Pacific region, ANZLIC published in 1996, a set of core metadata elements contained in ANZLIC's Metadata Guidelines, applied in Australia and New Zealand.

In the 1990's, the need for global standardisation of metadatasets for data discovery and transfer became apparent, which has lead to the development of the IS 19115 – Metadata standard by ISO/TC 211 – Geographic Information/Geomatics. Many countries in the world have stated they will harmonise the locally used metadata standard with the IS 19115.

As can been read in this book many countries are progressing in setting up the metadatasets. In Western Europe, North America, Australia, New Zealand the development of SDI's is well underway, and many metadatasets are already available. In Southern America and Asia, many countries have learned form the lessons from other regions and are creating SDI's and metadatasets. In Africa, the awareness of the need to describe datasets is still evolving.

Acknowledgements

We would like to acknowledge Olaf Østensen, Alexander Martynenko, Harold Moellering, Richard Pearsall and François Salgé for their contributions to this introduction.

We are also grateful to the Department of Geodesy and the Research Institute of Housing, Urban and Mobility Studies of the Delft University of Technology, as well as the Department of Geomatics of the Melbourne University to allow time to make these contributions to this book possible.

References

[Ahonen-Raino 2001]
Description of the content of geographic datasets
P. Ahonen-Raino (National Land Survey, Finland), in: J.T. Bjørke, H. Tveite (eds), Proceedings of the 8th Scandinavian Research Conference on Geographical Information Science, 25–27, June 2001, Ås, Norway

[Bertius 1630]
Geografie der Oudheid
P. Bertius, in: Antiquity Geography, Flanders, 1630

[DLC 1979]
Technical (Unclassified) Report UG478.A88, 1979
Defense Technical Information Center, Defense Logistics Center, U.S.A.

[Duval 2002]
Metadata principles and practicalities
E. Duval (KU Leuven, Belgium), W. Hodgins (Autodesk), S. Sutton (Univeristy of Washington), S.L.Weibel (DCMI), D-Llib Magazine April 2002, Vol. 8, no. 4, ISBN 1082-9873.

[Homer 1978]
A mathematical model of the flow of data in a management information system
E.D. Homer, 11th American Meeting of the Institute of Management Sciences, 1978.

[McCelland 2002]
Challenges for service providers when importing metadata in digital libraries
M. McCelland, D. McArthur, S. Giersch (Colleges Eduprise), G. Geisler (University of North Carolina), D-Llib Magazine April 2002, vol. 8 (4), ISBN 1082-9873.

[McIntosh 1968]
Information Processing 68
S. McIntosh, D. Griffel, Proceedings of International Federation for Information Processing (IFIP) Congress, August 1968.

[Moellering 1991]
Spatial database transfer standards: current international status
H. Moellering (ed.), Published on behalf of the International Cartographic Association by Elsevier Applied Sciences, London 1991. ISBN 185166677X.

[Moellering 1994]
Technical characteristics for assessing standards and for the transfer of spatial data and brief international descriptions
H. Moellering, C. Clement (eds), International Cartographic Association Standards Commission for the transfer of spatial data, Columbus, Ohio, U.S.A., 1994, 104 pp.

[Moellering 1997]
Spatial database transfer standards 2: characteristics for assessing standards and full description of the national and international standards in the world
H. Moellering (ed.), Richard L. Hogan associate editor. Published on behalf of the International Cartographic Association by Elsevier Applied Sciences, Oxford, UK and Tarrytown, N.Y. 1997. ISBN 0 08 042433 3.

[Morrison 1995]
Elements of Spatial Data Quality
J. Morrison, S. Guptill (ed.), International Cartographic Association, November 1995.

[Schelling 1978]
The use of IBM's data dictionary
G. Schelling, Proceedings of the Conference on Data dictionary Systems, London, November 1978

[Weber 1978]
Data Base theory and practice
H. Weber, A.I. Wasserman, Conference in Berlin, 13–15 September 1978.

[Williamson 2003]
Spatial DataInfrastructures – From Concept to Reality
I.P. Williamson, A. Rajabifard, M.-E.F. Feeney (eds), Taylor and Francis 2003. ISBN 0-415-30265-X.

[Uitermark 2001]
Ontology-Based Geographic Data Set Integration
H. Uitermark, PrintPartners Ipskamp Enschede, The Netherlands, ISBN 90-365-1617-X, 2001.

[Ziegler 1978] *Distribution: A New Impetus Toward Understanding Data*
K. Ziegler (International Business Machines, Inc.), August 1978.

Organisation	For on-line contact see URL:
Dublin Core Metadata Initiative	http://www.xml.com/pub/2000/10/25/dublincore/
INSPIRE	http://www.ec-gis.org/inspire
Institute for Electrical and Electronic Engineers, Learning Object Metadata Working Group	http://grouper.ieee.org/p1484/doc/wg12/LOM-WD3.htm
U.S.A. National Spatial Data Infrastructure	http://www.fgdc.gov
FGDC Official statement, Department of the Interior, Office of the Solicitor	http://www.fgdc.gov/metadata/meta_trademark.htm
International Organisation for Standardisation	http://www.iso.ch
ISO/TC 211 Geographic information/Geomatics	http://www.isotc 211.org/

URL's mentioned in this chapter were accessed on May, 2005, unless otherwise stated.