Report No. 2b **STANDARDS**

Data Management

b. Metadata

Arthur D. Chapman – 11 Jan 2004.

See also Guidelines 2: Guidelines to Documentation of Species and Vegetation Data from Environment Australia (**http://www.deh.gov.au/erin/documentation/metadata.html**) at

Background:

Metadata is information that describes datasets. If data is documented following good metadata standards it provides a consistent approach to the storage and retrieval of information. Spatial metadata standards have been developed for a number of countries and are used for documenting data storage and for use in accessing data through automated technologies on the internet using distributed search and retrieval.

The United States developed the Spatial Data Transfer Standard (SDTS) in 1994, with the current version released in 1998 (USGS 1998). This was a major breakthrough in metadata standards and data documentation. This standard is extensive, and is quite detailed, indeed, too detailed for many applications. The purpose of the SDTS was "to promote and facilitate the transfer of digital spatial data between dissimilar computer systems, while preserving information meaning and minimizing the need for information external to the transfer" (USGS 2004).

In 1996, the Australian and New Zealand Land Information Council (ANZLIC) – an inter-governmental organization developed an offshoot standard for use in Australia and New Zealand. This was somewhat simplified standard and one that has now been universally adopted throughout Australia and New Zealand as the ANZLIC Metadata Guidelines (ANZLIC 2003).

More recently, the ISO Technical Committee for Geographic Information/Geomatics (ISO/TC 211) is working on an International Standard for Metadata (ISO 19115). This standard is now published and available for a cost of 208 Swiss Francs (ISO 2003). The Australian standard (ANZLIC 2003) is currently being modified to comply with the standard data elements listed under the ISO 19115, however, this is likely to cause little change the basic structure of ANZLIC Metadata elements.

Geospatial Metadata Standards in Brazil

In 1996, a Case was made for Digital Spatial Metadata for Federal databases in Brazil (Ribeiro *et al.* 1996).

In 1998, a number of institutions in Rio Grande do Sul examined existing metadata standards around the world and made a comparison with the idea of making some recommendations for Brazil (Weber *et al.* 1998). They made a comparison of the SDTS (Spatial Data Transfer Standard – USA) (An early version of USGS 1998), the CSDGM (Content Standards for Digital Geospatial Metadata – USA) (FGDC 1997), the SAIF (Spatial Archive and Interchange Format – Brazil) (Bins *et al.* 1998) and CEN (Europe) (European Committee for Standardization 1996). They also looked at software available for use with the standards.
This was early in the development of many of the standards, and since that study the standards have moved on considerably, with the CSDGM and the STDS being combined into the one standard (USGS 1998), and a number of new standards developed.

One conclusion of the study was that for adoption in Brazil " a metadata standard must be at the same time sufficiently inclusive, flexible to allow for adequate description of a data set and as simple as possible to allow its easy handling by producers and users of data" (Weber *et al.* 1998).

In 1998, the USGS developed some services in Brazil and Costa Rica at the same time as ERIN developed a searchable Clearinghouse for federal holdings in Australia (UKOLN 1998).

Since 1998, however, although discussed at a number of meetings and conferences, from what I can find on the Internet, little seems to have been done in Brazil toward developing either a geospatial metadata standard for Brazil or a data clearing-house for the country.

ANZLIC Metadata Guidelines

There are a number of ways CRIA, the State of São Paulo, or even Brazil could go with the development of Spatial Metadata Standards and Guidelines and/or a spatial data clearinghouse. I do believe that something does need to be done, and the Australian system is an excellent example of the benefits that can be achieved by having good metadata guidelines linked to an online searchable data clearinghouse. It may provide a simple solution for CRIA, if not for the whole of Brazil.

I cannot comment on other options that may be available, as I have little knowledge of many of them. I was, however, involved in the development of the Australian metadata standard as well as with the development and implementation of the first on-line distributed data clearing house based on that standard. Indeed, the Federal Geographic Data Committee used the Australian Clearing House as an example for their future developments (UKOLN 1998).

The ANZLIC Metadata Guidelines (Version 2 - 2001) (ANZLIC 2001a) are available as a PDF document downloadable from the Internet. They were developed from the core elements of the US SDTS, by simplifying it for use in Australia and New Zealand, and altering a number of terms for Australian usage.

The standard is simple, easy to use, and has been used successfully to document all Government spatial data in Australia since 1998. A Clearinghouse has been developed based on the Standard (see ASDD below) to allow on-line searching and download of data, and a simple data entry tool (see ANZLIC Metadata Entry Tool, below) to aid in the entry of data in the ANZLIC format.

Documents available for public use include the ANZMETA XML Document Type Definition (DTD) for geospatial metadata (http://www.ga.gov.au/anzmeta/). Additional guidelines have also been developed by the Environmental Resources Information Network (ERIN) to aid users to enter species and ecological data. It includes explanations and examples of data of the type used in CRIA (Chapman 1998).

Australian Spatial Data Directory (ASDD)

The Australian Spatial Data Directory (ASDI 2004), provides search interfaces for the discover of geospatial dataset descriptions (metadata) across Australia. It was originally developed at ERIN in 1998, and is now managed through the Australian Spatial Data Infrastructure. It has live links to 35 different metadata sites with a total of 33 561 dataset descriptions. The average time for a search is around 1 second (see Server Status http://www.ga.gov.au/asdd/tech/serverstatus/), with institutions not returning information within 60 second being times out. The site is getting around 80 000 to 100 000 Total Page Views per Quarter based on around 5,000 visitors each quarter.

It has two search interfaces – a **Basic Search interface** (http://asdd.ga.gov.au/servlet/asdd_basic), which allows simple boolean textual searches in a number of fields, and a **Power Search interface** (http://asdd.ga.gov.au/servlet/asdd_power) which allows geographic searches (by either typing in bounding coordinates, or through the use of a map interface), with searching of overlapping or included datasets and temporal searches in addition to the simple textual searches mentioned previously.

Technical details are available at http://www.ga.gov.au/asdd/tech/. The system is based on XML and Z39.50 and technical information is freely available for use in setting up nodes etc.

> "This WWW-Z39.50 gateway interface to the ASDD is based on a java servlet called the MetaStar Gateway which is developed by Blue Angel Technologies. Anyone can also build their own independent gateway to the ASDD. The underlying Z39.50 servers can also be accessed by true Z39.50 client software (see lists at Z39.50 Maintenance Agency and DSTC)."

Further information is available from the site ASDI site http://www.ga.gov.au/asdd/.

Environment Data Directory (EDD) – Downloadable data

The Department of the Environment have developed a subset of data in the ASDD that specifically related to the Environment (Environment Australia 2003). This is developed using similar architecture to the ASDD, and indeed began life as the prototype for the ASDD.

In addition to the searches available through the ASDD, one is able to search for Downloadable data (Freeman *et al.* 2000). At present 81 downloadable data sets are available http://www.deh.gov.au/erin/edd/index.html. When one clicks on the downloadable data, a form pops up that asks for some basic information such as name, contact information and questions as to how the data will be used. Once this information is filled in, a Licence Agreement appears with the details of the previous page included. Once you agree to the conditions, the data is made available for download in one of several formats. Along with the packaged data in a .zip file, one will receive a copy of the Licence Agreement and a copy of the Metadata description of the data (although a recent test didn't provide those in the downloaded file, only the data).

ANZLIC Metadata Entry Tool (MET)

The ANZLIC Metadata Entry Tool (MET) (ANZLIC 2001b) is a freeware tool designed for document entry and maintenance in a form compliant with the ANZLIC Metadata Guidelines (ANZLIC 2001a). It has been developed in Microsoft Access 97. It apparently it has not yet been developed to operate in the Windows 2000/Office 2000 environment. I was able to install, but not open the database.

Other Metadata Entry Tools

A number of other Metadata entry tools and other tools have been developed independently, but I have not attempted to use any of them. A list and links to some of these can be found on the ASDD Web site at http://asdd.ga.gov.au/asdd/tech/tools.html.

Other Spatial Data Clearinghouse examples

A number of other Spatial Data Clearinghouses and Data Directories exist around the world. Some include distributed searches, while others appear to be largely internal. They are listed below without comment.

- National Geospatial Data Clearinghouse
  - 250 spatial data servers
  - based on SDTS
  - http://clearinghouse1.fgdc.gov/
- USGS Node of the National Spatial Data Clearinghouse
  - One of many nodes of the above
  - http://nsdi.usgs.gov/
- New Jersey Geographic Information Network
  - https://njgin.state.nj.us/NJ_NJGINExplorer/index.jsp

- Australian Antarctic Data Centre
  - http://www.aad.gov.au/default.asp?casid=3802
- European Environment Agency – Data Service
  - http://dataservice.eea.eu.int/dataservice/
- Global Change Master Directory
  - http://gcmd.gsfc.nasa.gov/
- NOAA Environmental Services Data Directory
  - http://www.esdim.noaa.gov/NOAA-Catalog/NOAA-Catalog.html
- International Arctic Environmnet Data Directory
  - http://arctic.unep.net/
- Australian Oceanographic Data Centre
  - http://www.aodc.gov.au/
- The Global Information Locator Service
  - I was involved on a committee that oversaw the development of this, I wasn't aware it was still operating.
  - http://www.gils.net/
- UNEP GRID
  - http://www.grida.no/
- South African Marine Fisheries and Environmental Metadata Directory
  - http://sea.uct.ac.za/idyle/gis/matt/

In addition, the Australian Antarctic Division (http://www.antdiv.gov.au/default.asp?casid=4079) provides links to a number of directories as well as links to other metadata standards.

Conclusion

It would appear that Metadata Guidelines have not yet been developed for Brazil. Australia, has many similarities to Brazil with similar environments, and with probably a similar number of datasets.

One of the deficiencies in CRIA and Brazil as a whole at the moment is the availability of a simple, efficient and user-friendly on-line environmental data discovery system. Data is generally not documented in a consistent and standard format that makes it easy for use in data discovery tools. Good metadata documentation is an essential first step.

The Environmental Resources Information Network (ERIN) in Australia, and CRIA in Brazil have many similar requirements as far as metadata and access to environmental datasets are concerned. The development in Australia of standards and guidelines such as the ANZLIC Metadata Guidelines and data discovery tools such as the Environmental Data Directory and the Australian Spatial Data Directory may provide the basis for the development of a robust and user-friendly metadata system for use in CRIA, and perhaps elsewhere in Brazil. I believe that this may be the simplest way for CRIA to go.

It is essential, that whatever is done at CRIA, is consistent with the new ISO Standard 19115.

References

ANZLIC (2001a). *ANZLIC Metadata Guidelines: Core metadata elements for geographic data in Australia and New Zealand.* Version 2. (February 2001). Canberra: ANZLIC http://www.anzlic.org.au/download.html? oid=2358011755 [Accessed 11 Feb 2004].

ANZLIC (2001b). Metadata Entry Tool. Canberra: ANLIC. http://www.anzlic.org.au/infrastructure_MET.html [Accessed 11 Feb 2004].

ANZLIC (2003). *Spatial Metadata*. Canberra: ANZLIC. http://www.anzlic.org.au/infrastructure_metadata.html [Accessed 11 Jan. 2004].

Australian Spatial Data Infrastructure (ASDI) (2004). *Australian Spatial Data Directory*. http://www.ga.gov.au/asdd/ [Accessed 11 Jan 2004].

Bins, C.A., Ruschel, C. and Moraes, E.A.S. (1998). *Pró-Guaíba: sistema de informações geográficas*, in Proc. GIS Brasil98. Brazil: Curitiba.

Chapman, A.D. (1998). *Guidelines for documentation of species and vegetation data*. Canberra: Environment Australia. http://www.deh.gov.au/erin/documentation/pubs/metadata.doc [Accessed 11 Feb 2004].

Environment Australia (2003). Welcome to the Environmental Data Directory. Canberra: Department of the Environment. http://www.deh.gov.au/erin/edd/index.html [Accessed 11 Feb 2004].

European Committee for Standardization (1996). *CEN/TC-287*. Bruxelles: ECS.

FGDC (1997). *Content standards for digital geospatial metadata*. Washington: Federal Geographic Data Committee.

Freeman, N., Boston, T. and Chapman, A.D. (2000) Integrating Internal, Intranet and Internet Access to Spatial Datasets via ERIN's Environmental Data Directory **in** *Proceedings of the 26th Annual Conference of AURISA, Perth, Western Australia, 23-27 November 1998*. AURISA.

ISO (2003). *ISO 19115*. Geneva, Switzerland; International Organization for Standardization. http://www.anzlic.org.au/infrastructure_metadata.html [Accessed 11 Jan. 2004].

Ribeiro, G.P., de Souza, J.M. and de Freitas, A.L.B. (1996). Digital Geospatial Metadata: a Brazilian Case of Federal Databases. IEEE. http://www.computer.org/conferences/meta96/ribeiro/ribeiro.html [Accessed 11 Jan. 2004].

UKOLN Metadata Group (1998). FGDC – Content Standards for Digital Geospatial Metadata. Environment of Use. http://www.ukoln.ac.uk/metadata/desire/overview/rev_09.htm [Accessed 11 Jan. 2004].

USGS (1998). *View the SDTS Document*. Washington: USGS. http://mcmcweb.er.usgs.gov/sdts/standard.html [Accessed 11 Jan. 2004].

USGS (2004). *What is SDTS?* Washington: USGS. http://mcmcweb.er.usgs.gov/sdts/whatsdts.html [Accessed 11 Jan. 2004].

Weber, E.J., Lisboa F.J., Iochpe, C. and Hasenack, H. (1998). Geospatial metadata in Brazil: an experience in data documentation of an environmental GIS application. Proceedings of GIS Planet 98 – International Conference & Exhibition on Geographic Information. Lisbon, Portugal. http://www.dpi.ufv.br/~jugurta/papers/gisplanet98.PDF [Accessed 11 Jan. 200