SAHRA Metadata Guidelines

What are Metadata?

Metadata are data or information about data which are usually stored in a structured document. Metadata describe the origin, nature, and spatial and temporal location of data. They are important in searching, sharing and using the data in scientific communities. A metadata document consists of elements which guide its creation and facilitate the search and retrieval process of data.

Metadata elements are arranged in a metadata model or schema and described using worldwide standards:

- International Standard Organization (ISO) 19115:2003 (http://www.iso.org/iso/en/ISOOnline.frontpage)
- Federal Geographic Data Committee (FGDC) STD-001-1998 (http://www.fgdc.gov/metadata/metadata.html)
- Dublin Core (http://dublincore.org/).

Scientific communities often develop more specific standards based on ISO or FGDC but add elements specific to the discipline. Some standards for hydrologic data include:

- CUAHSI Metadata Profile (http://thor.cae.drexel.edu/cuahsi/)
- Ecological Markup Language (EML) (http://knb.ecoinformatics.org/software/eml/)
- Geographical Markup Language (GML) (http://xml.coverpages.org/ni2004-03-26-a.html)
- USGS Hydrologic Markup Language (HYDROML) (http://water.usgs.gov/nwis_activities/XML/nwis_hml.htm)
- Earth Science Markup Language (ESML) (http://esml.itsc.uah.edu/index.jsp)

Geospatial Metadata Standard

Content Standard for Digital Geospatial Metadata (CSDGM) FGDC-STD-001-1998 provides a common set of terminology and definitions for digital geospatial data documentation. It is mainly used in the United States. Basic elements are organized into ten groups, seven of them major (1-7) and three minor (8-10):

1. Identification Information
2. Data Quality Information
3. Spatial Data Organization Information
4. Spatial Reference Information
5. Entity and Attribute Information
6. Distribution Information
7. Metadata Reference Information
8. Citation Information
9. Time Period Information
10. Contact Information

A detailed description of these groups is provided at www.fgdc.gov/metadata/contstan.html.

Basic elements of a metadata document include:

- Dataset Title
- Attributes
- Purpose
- Access Constraints
- Contact Information
- Citation
- Time Period of Dataset
- Status
- Spatial Domain
- Keywords
- Distribution
- Metadata Reference

The harmonization of the new ISO and FGDC geospatial metadata standards will produce a universal standard for geographical metadata ISO-19139 for digital geographical data. This standard is currently in draft stage.

CUAHSI Metadata Standard

Table 1 shows the CUAHSI metadata profile recently developed by Piasecki et al. (2005). It consists of 24 ISO-based elements which describe geo-spatial data and three additional elements from other parts of the ISO standard.

The other essential component of metadata standards is a controlled vocabulary (CV). Using CV is beneficial for consistency in metadata creation, for searching and browsing metadata, and for matching other metadata by creating links between similar terms. The main objective of CUAHSI's metadata profile was to create a CV that is interoperable with already existing CVs.

Table 1 CUAHSI metadata profile (Piasecki et al., 2005)
(M – mandatory, O – optional, and C – conditional)

Element	Requirement		
	ISO	CUAHSI	
Dataset Title	M	M	
Dataset Reference Date	M		
Dataset Language	M		
Dataset Topic Category	M		
Abstract (description of dataset)	M		
Metadata Point of Contact	M		
Metadata Date Stamp	M		
Distribution Format	O		
Dataset Responsible Party	O		
Additional Temporal Extent Information	O		
Lineage	O		
On-line Resource	O		
Metadata Standard Name	O		
Metadata Standard Version	O		
Metadata Language	C		
Geographic Location (four coordinates)	C		
Geographic Location (geographic identifier)	C	O	
Dataset Character Set	C		
Metadata Character Set	C		
Spatial Resolution of the Dataset	O		
Additional Vertical Extent Information	O		
Spatial Representation Type	O		
Reference System	O		
Metadata File Identifier	O		
CUAHSI-Thematic Keywords		M	
Dataset Legal Constraints			
Dataset Security Constraints			

More information about the CUAHSI Metadata Profile can be found at http://thor.cae.drexel.edu/cuahsi/ .

Minimum Metadata Description

ISO/CUAHSI and FGDC standards are very complex and contain over 300 elements, all of them important in the data description. Must all of these elements be described? What if sufficient information is lacking? This issue can be solved by defining a minimum searchable metadata set, which is a subset of the mandatory elements. Table 2 compares minimum requirements for description sets by CUAHSI (Piasecki et al., 2005) and by the "Denver Summit" Metadata Core (discussed at http://geology.usgs.gov/tools/metadata/tools/doc/faq.html) for geographic data.

Table 2 Minimum metadata set requirements for CUAHSI and Denver Core

CUAHSI	**Denver Core**
Publisher	Originator
Subject keyword	Theme keywords
	Place keywords
Spatial Coverage	Bounding coordinates
Temporal Coverage	Time period of content
Title of Data set	Title
Description	Abstract
	Purpose
Download link of data file	
Download of full metadata set	Resource description
File format/resource type	Geospatial data presentation form
File size	
Access control of permission	
Last successful update	
	Currentness reference
	Language

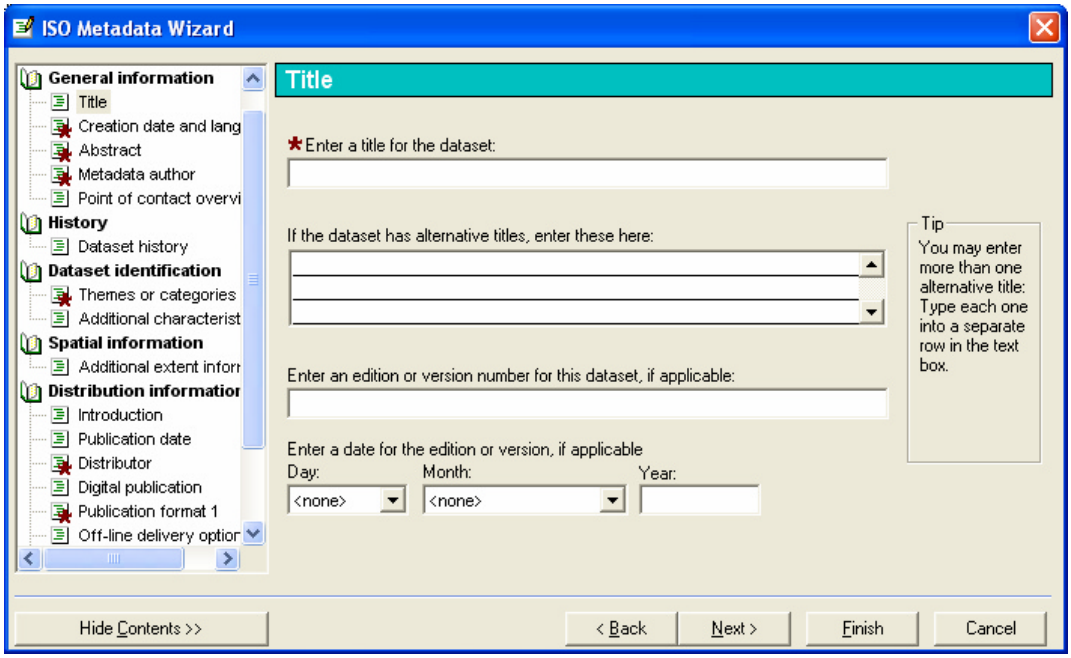
SAHRA Metadata Creation and Editing

To describe data correctly we have to understand what they represent and how they were obtained. Then we have to decide what tools should be used to create metadata, based in part on data format and software availability. The SAHRA Geo-Database uses ArcSDE as the spatial engine. The metadata and all spatial data (vector and raster), tabular data, and time series are stored in an Oracle database.

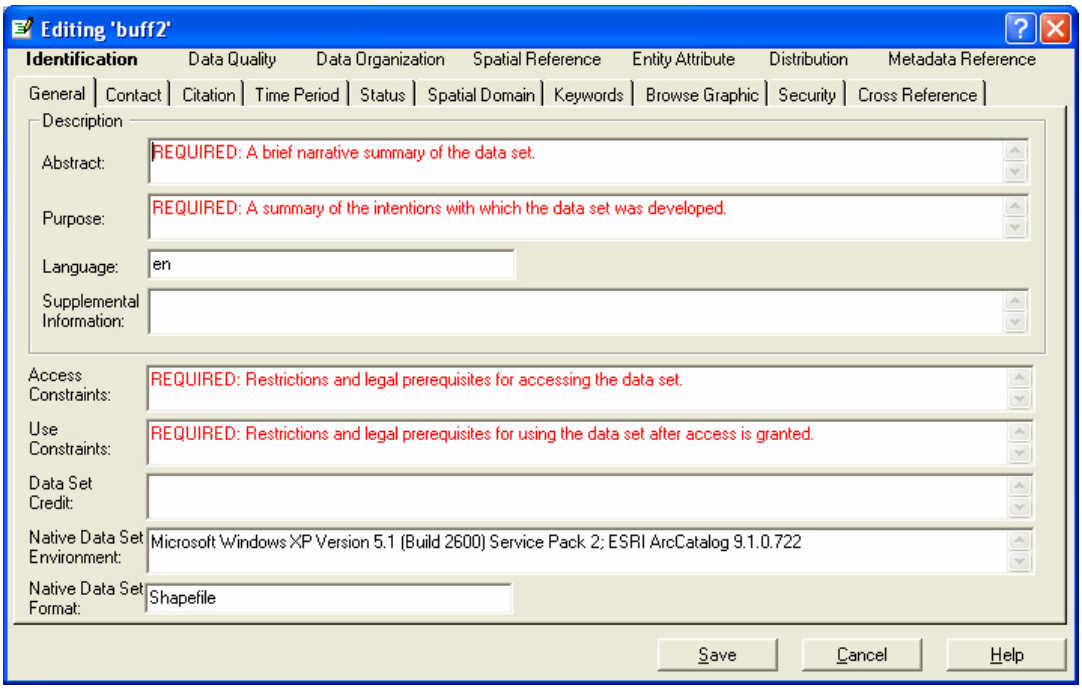
Based on the database implementation, we recommend that ESRI FGDC/ISO and CUAHSI metadata profiles be used to generate metadata files for data uploaded to SAHRA Geo-database. Under certain circumstances, metadata may be created using the minimum metadata profile.

1. GIS software - ArcGIS

- A metadata editor is a part of the ArcCatalog. This editor incorporates ISO and FGDC editors and a viewer form. It is easy to use and the required elements are marked with an asterisk. More detailed information about the ESRI Profile (extended FGDC standard) is described at http://www.esri.com/metadata/esriprof80.html. The implementation of ISO standards is described in http://www.esri.com/library/whitepapers/pdfs/supported-ogc-iso-standards.pdf .



Example of ISO Metadata Wizard implemented in ArcCatalog.
(Elements marked with * are required.)



Example of FGDC Metadata Wizard implemented in ArcCatalog.

2. Stand-alone or Internet-based metadata editors

- TKME (http://geology.usgs.gov/tools/metadata/tools/doc/tkme.html) - a Windows-based multilingual metadata entry and editing tool with the FGDC-STD-001-1998 standard implemented. It is developed by the USGS.
- XTME (http://geology.usgs.gov/tools/metadata/tools/doc/xtme.html) –a UNIX-based metadata editing tool with the FGDC-STD-001-1998 standard implemented. It is also developed by USGS.
- M3Cat (http://www.intelec.ca/anglais.html) –an Internet-based metadata tool that supports multilingual metadata entry and editing. Metadata can be created using various metadata standards (FGDC, ISO, GILS, NBII) and other metadata standards can be added.
- A list of additional metadata tools is listed at http://sco.wisc.edu/wisclinc/metatool/ or http://www.w3.org/XML/Schema.

Reference

Piasecki, M., L. Bermudez, B. Beran, S. Islam, Y. Choi, X. Liang, and S. Jeong. (2005) Hydrologic Metadata, in *Hydrologic Information System Status Report*, ed. by D. Maidment, CUAHSI, Web access http://www.cuahsi.org/docs/HISStatusSept15.pdf.