Near-term metadata challenges

M. Gould, J. Rocha, S. Nativi, J. Nogueras, M. Manso

1 Information Systems, Universitat Jaume I, Castellón, Spain
2 Computer Science, Universidade de Minho, Braga, Portugal
3 Inst. Methodologies for Environ. Analysis, CNR-IMAA, Italy
4 Computer Science and Engineering, Universidad de Zaragoza, Spain
5 Topographic Engineering and Cartography, Universidad Politécnica de Madrid, Spain

Based on experience from the AGILE Working Group on interoperability, SDI component development projects, and from the INSPIRE metadata Implementation Rules Drafting Team, we describe our personal views on near-term metadata challenges in support of a fully functional and useful SDI under the INSPIRE umbrella. We address both current (and legacy) issues as well as future needs, focusing on advances that are judged as possible within the coming 5 years.

Our views and recommendations fall under the following seven categories.

Shift from cartographic to informatics viewpoint
Geodata are collections of digital objects. Particular assemblies of these objects, when displayed, become maps. Spatial metadata need to facilitate discovery and description of historical and current data in the form of map series and even individual paper maps, however this will soon become the rare legacy case. Modern metadata need to be able to provide discovery and description of objects: features encapsulating geometry, associated thematic attributes, self-description, and behaviour rules. Other related multimedia fields of study have much to offer here.

Automated production and extraction
Metadata text editors are also destined to become legacy applications. The geodata collection or creation process needs to explicitly include the metadata creation process, as is now common in the remote sensing and other communities. The geodata community needs to change its mindset, through education initiatives, so that the next generation of geodata specialists naturally expects metadata to be present and attached or associated with the geodata payload. Automatic extraction from within a GIS environment currently can collect an estimated 50% of the core metadata necessary to provide discovery-level interoperability. For this practice to become commonplace mindsets also need to change to accept the fact that users may, and will, create their own metadata, being geodata experts or not, as they become user-providers.

Separation of discovery and description
Current metadata standards include the ability to discover and describe geodata resources. However it is clever implementation practice, and not the standards themselves, that creates innovative, useful services based on metadata. Much work is needed in order to educate the community on the key differences between discovery metadata—determining what is available—from the secondary description of what is discovered. Current standards mix these two metadata types among the (often) hundreds of elements in the same document, and this causes confusion as it mixes the user/discovery and provider/cataloguing communities.

Linking metadata to data to services
The three worlds of creation and publication of metadata, data, and services continue to exist in parallel rather than in an integrated form. Work on the creation and use of identifiers linking these three aspects, facilitating the use of registries, is direly needed.

Enabling an optimal use of thesauri to aid multilinguality
True pan-European network services need to support automated multilingual support. This will involve linking various parts of the user experience to multilingual thesauri and gazetteers, permitting queries in one language to be handled using metadata in another language, with responses in perhaps yet another. The SDIGER (INSPIRE pilot) project, among others, has provided interesting input to this issue.

Treatment of imagery and other Earth Science data
As observational and model output datasets in the Earth Sciences (ES) increase in resolution, there is a growing demand for information systems that interoperate between GI and ES domains. However, differences in the way the two communities think about and describe their data can give rise to difficulties in integrated analysis and display of datasets from the two disciplines. Improved geospatial data integration and GI Management is especially important for the European GMES initiative, which aims to provide society with certified and documented data from Earth observation sources and in situ measurements and surveys.

The GI community has been working on solutions for treating ES datasets. These efforts lead to the definition of “more general” models for geospatial information. Such models distinguish two kinds of geospatial information: boundary and coverage data. Boundary data is often called "vector data" and is almost always feature oriented. Generally, ES datasets are thought of as imagery or coverages and they often involve grid-oriented data. GI data and metadata models have been reshaped and extended. A valuable example is represented by the ISO 19115 Part 2: Metadata for imagery and gridded data; it extends the existing geographic Metadata standard by defining the schema required for describing imagery and gridded data.

In order to understand to what extent GI data and metadata models are suited for representing ES datasets, there are significant questions to be addressed, such as: 1. How well is time modeled? 2. How much of ES semantics are effectively captured? 3. How important is the documentation of acquisition process or measuring equipment for discovering and evaluating ES data?

In the Web era, the GI and ES different data and metadata models produce diverse content models generating disciplinary Markup Languages (e.g. GML, ncML, ESML, etc.). Mediation approaches, such as crosswalks languages, represent a valuable solution to harmonize GI and ES models.

Testing Onsrud’s Geodata Commons
GI professor, lawyer and GSDI president Harlan Onsrud has advocated a geodata commons, with set rules for creating geodata, documenting it with metadata, publishing both geodata and metadata (in a semi-automated manner) and also preserving rights while openly sharing the geodata with the community. This model, which is not at all incompatible with INSPIRE, needs to be tested to determine viability and possible benefit to the wider SDI community.

12th EC & EC GIS Workshop

Innsbruck, Austria

21-23 June 2006

Edited by

Karen Fullerton
and
Katalin Tóth

JRC IES

MISSION STATEMENT
The mission of the Institute for Environment and Sustainability is to provide scientific and technical support to the European Union's policies for protecting the environment and the EU Strategy for Sustainable Development.

European Commission
Joint Research Centre (DG JRC)
Institute for Environment and Sustainability (IES)
Spatial Data Infrastructures Unit
I-21020 Ispra (VA), Italy

Contact information
Tel.: +0039 0332 786491
Fax: +0039 0332 789803

E-mail: ies@jrc.it
Website: http://ies.jrc.cec.eu.int/

Editor: Karen Fullerton, Katalin Tóth
Cover: José-Joaquín Blasco

Legal Notice
The contents of this document do not necessarily reflect the official opinion of the European Commission or the European Community Institutions. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of the information contained in this production.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server
http://europa.eu.int

Luxembourg: Office for Official Publications of the European Communities

EUR LB-X1-06-024-EN
ISBN 92-79-02083-8

© European Communities, 2006

Reproduction is authorized provided the source is acknowledged

Printed in Austria by Hernegger Offsetdruck GmbH, Innsbruck

Contents

SESSION: POSTERS 147