Interoperability, Standards, and Metadata

David M. DANKO, United States

Key words: geographic information, interoperability, standards, metadata

SUMMARY

Interoperability has helped humans advance to our position in the world today. Interoperability becomes more complex and important as the world becomes more integrated. There are many types of interoperability two important types are technical and semantic. The need for interoperability for GIS increases as GIS moves into mainstream information technology (IT) applications and with the increased use of web services' loosely couple networks. There are many factors that are required to make interoperability happen; two major factors are standards and metadata. Standards, criteria which document agreement between a provider and a consumer, enable both technical and semantic interoperability. ISO TC 211 and OGC are developing standards and specifications in the field of geographic information. Metadata has always played an important role in cartography; for centuries it has provided users with an understanding of maps. Metadata is equally important as we have moved into the digital environment. Because digital data is an imperfect representation of the real world, and with the proliferation of data from an ever-widening array of sources and producers, we need knowledge provided by metadata to understand, control and manage geographic information. Metadata adhering to the international standards will allow global networks to operate, provide a common global understanding of geographic data, and promote global interoperability.

INTEROPERABILITY

Mankind has prospered and become the dominate species because of the ability to combine the intellect and efforts of many by working together - as a family, as a village, as a tribe, as a team. Through the ages various societies have prospered and advanced ahead of others partly because they were better able to work together, to communicate and to interoperate enabled by common goals, language, and/or ideals. Each individual can focus on what they do best, their share of what it takes for a society to prosper and advance (imagine where mankind would be if we each had to supply all of our needs on our own). If each individual's part can interoperate well, then the outcome is much larger than the sum of all the parts. This is truer today than ever. We depend on others for most of our needs in an integrated and mostly congruent society. Increasingly societies are interoperating globally. "Think globally, act locally"- things that happen globally affect us locally and things we do locally affect the global community. This does not happen spontaneously it takes coordination, communication, planning, laws, and awareness – all of these things make interoperability possible.

There are several types of interoperability: “technical interoperability” where things physically work together, for instance the size and thread count for nuts and bolts, in information technology (IT) it’s the ability of computers and peripheral devices and software modules to interact; “Semantic interoperability” is the ability to communicate concepts, a common understanding of terminology, this is the type of interoperability with which we struggle the most; a third type of interoperability is “political or human interoperability” which involves societal customs, government regulation, education and training and legal interoperability involving laws, protection of intellectual property, and ownership.

There are many things that are needed to make interoperability happen. We need an infrastructure to support interoperability, a common architecture, and compatible technologies. We need authorization – both authorization to share our data and services with others, and authorization to uses other’s data and services. We need to insure individual’s and organization‘s intellectual property rights are not infringed; we need good copyright laws. We need business agreements and a business model; if it’s of unequal benefit to both sides than there is no need to exchange information, no need for interoperability. Of course we need quality assurance; if the information in an exchange is no good then there is no reason to be interoperable. We need **standards**; standards allow us to interoperate both technically – standards for physical sizes, shapes, frequencies, and in IT standards to ensure hardware and software work together; and semantically – standards for languages, dictionaries, and information models to ensure we are using and understanding the same terms for the same concepts. And of course we need to comprehend other’s data and services, for true interoperability we need - **metadata**. Metadata provides a vehicle to locate and understand data which may be produced by one community and applied by another.

Geographic information systems have (GIS) always required interoperability. GIS uses data from multiple sources and from multiple distributed organizations within a community. For years GIS has been merging different information types: raster, vector, text, and tables. As the use of GIS grows and moves into varied disciplines the need for interoperability increases. Today GIS has to interoperate with a broad array of IT applications and is applied across diverse information communities. Web Services carry this need to new heights with loosely coupled, distributed networks.

STANDARDS

The International Organization for Standardization (ISO) defines a standard as a documented agreement between a provider and a consumer. These are reference documents that may be used in public contracts or international trade. They provide definitions of characteristics, technical specifications, precise criteria, rules, and/or guidelines, which will ensure materials, products, processes and services, are fit for purpose. Their primary function is to ensure interoperability, promote innovation, competition, commerce and free trade. Standards are usually a political compromise, something organizations, consumers and providers, or nations can agree to. Typically they are developed as a consensus solution, not the most advanced tech-

nology but something everyone can agree to. Standards also serve as a democratic mechanism and technology transfer; many organizations participating in standards development are there to learn as much as provide input.
Standards are developed by organizations with a common interest. There are many standards organizations: specific industry, national, trading block (typically regional – Europe, North America, etc), and global. The broader the community to which a standard applies the broader the interoperability, so in most cases international standards provide the broadest interoperability.
ISO is one of the primary international standards organizations. The term ISO stands for "equal" and is not a contraction of International Organization for Standardization. ISO has been around since 1947 with 146 members – one member from each nation. Nations are represented by their national standards originations like ANSI or DIN so members do not necessarily represent the government, so unlike the UN, many ISO representatives have roots in the private sector and industry associations. To date ISO has developed approximately 14000 standards. Each ISO member has one vote – so each nation is on equal footing regardless of economic strength or size. ISO uses a consensus process to ensure widespread applicability and uses a voluntary process to ensure the standards are market driven. ISO itself has no legal authority; the standards are adopted by nations which provide the legal mandate for their use. ISO standards benefit trade by creating wide acceptance of products and services; companies are free to compete in broader markets with technical barriers removed. ISO standards provide support for political trade agreements. They benefit governments by providing technical and scientific underpinnings for legislation and benefit consumers by providing assurances with respect to quality, safety and reliability.

Standards and geographic information

There are many organizations developing standards for geographic information. Two of the primary organizations are ISO TC 211 developing an integrated suite of de jour standards to address both technical and semantic interoperability; and the Open Geospatial Consortium developing de facto specifications focusing, now primarily, on geospatial application program interfaces for World Wide Web applications.

ISO/TC 211 - Standardization in the field of digital geographic information

ISO formed a Technical Committee, TC211, to establish standards for Geographic Information. The official Scope of ISO TC 211:

- "This work aims to establish a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth.

- These standards may specify, for geographic information, methods, tools and services for data management (including definition and description), acquiring, processing, analyzing, accessing, presenting and transferring such data in digital/electronic form between different users, systems and locations.

- The work shall link to appropriate standards for information technology and data where possible, and provide a framework for the development of sector-specific applications using geographic data."

The goals of the committee are to develop a family of international standards that will support the understanding and usage of geographic information; increase the availability, access, integration, and sharing of geographic information, enable interoperability of geospatially enabled computer systems; contribute to a unified approach to addressing global ecological and humanitarian problems; ease the establishment of geospatial infrastructures on local, regional, and global levels; and contribute to sustainable development.

ISO TC 211 has been productive and to-date has produced 24 International Standards and 3 Technical Reports with many more waiting in the wings, as Draft International Standards, soon to be completed. Many of these early standards can be considered foundational for defining geographic information such as Spatial Schema, Temporal Schema, Metadata, Quality, and Spatial Referencing by Coordinates. Others are based on these foundational standards; examples are Simple Feature Access, Web Map Services, and Geographic Mark-up Language (GML). Presently ISO TC 211 is focusing on additional web service specifications, location based services and imagery related standards, and specific information community needs.

ISO 6709:1983	Standard representation of latitude, longitude and altitude for geographic point locations
ISO 19101:2002	Geographic information -- Reference model
ISO/TS 19103:2005	Geographic information -- Conceptual schema language
ISO 19105:2000	Geographic information -- Conformance and testing
ISO 19106:2004	Geographic information -- Profiles
ISO 19107:2003	Geographic information -- Spatial schema
ISO 19108:2002	Geographic information -- Temporal schema
ISO 19109:2005	Geographic information -- Rules for application schema
ISO 19110:2005	Geographic information -- Methodology for feature cataloguing
ISO 19111:2003	Geographic information -- Spatial referencing by coordinates
ISO 19112:2003	Geographic information -- Spatial referencing by geographic identifiers
ISO 19113:2002	Geographic information -- Quality principles
ISO 19114:2003	Geographic information -- Quality evaluation procedures

ISO 19114:2003/Cor 1:2005	
ISO 19115:2003	Geographic information -- Metadata
ISO 19115:2003/Cor 1:2006	
ISO 19116:2004	Geographic information -- Positioning services
ISO 19117:2005	Geographic information -- Portrayal
ISO 19118:2005	Geographic information -- Encoding
ISO 19119:2005	Geographic information -- Services
ISO/TR 19120:2001	Geographic information -- Functional standards
ISO/TR 19121:2000	Geographic information -- Imagery and gridded data
ISO/TR 19122:2004	Geographic information / Geomatics -- Qualification and certification of personnel
ISO 19123:2005	Geographic information -- Schema for coverage geometry and functions
ISO 19125-1:2004	Geographic information -- Simple feature access -- Part 1: Common architecture
ISO 19125-2:2004	Geographic information -- Simple feature access -- Part 2: SQL option
ISO/TS 19127:2005	Geographic information -- Geodetic codes and parameters
ISO 19128:2005	Geographic information -- Web map server interface
ISO 19133:2005	Geographic information -- Location-based services -- Tracking and navigation
ISO 19135:2005	Geographic information -- Procedures for item registration

Table 1
ISO/TC 211 completed standards

Open Geospatial Consortium

The Open Geospatial Consortium (OGC®) is an international industry consortium of 310 government agencies, research organizations, universities, and companies working together in a consensus process to develop publicly available interface specifications which support interoperable solutions "geo-enabling" the Web.

The primary goal of the consortium is to create “a world in which everyone benefits from geographic information and services made available across any network, application, or platform.” Unlike ISO, OGC is organized as a business with a president, a board of directors, and a strategic advisory committee, which oversee three programs: the specification program, the interoperability program, and the outreach and adoption program. In most cases the interoperability program runs test-beds to test and develop interoperability technology and to test draft specifications. The specification development is done in the Technical Committee which is overseen by a Planning Committee. Members participate and vote according to their level of membership. To date OGC has developed 19 specifications:

Specification	**Version**	**Date**
GML simple features profile (GMLsf)	1.0.0	2006-05-08
OpenGIS® Catalogue Service Implementation Specification (CAT)	2.0.1	2004-08-02
OpenGIS® Coordinate Transformation Service Implementation Specification (CT)	1.0	2001-01-12
OpenGIS® Filter Encoding Implementation Specification (Filter)	1.1	2005-05-03
OpenGIS® Geographic Objects Implementation Specification (GO)	1.0.0	2005-05-04
OpenGIS® Geography Markup Language (GML) Encoding Specification (GML)	3.1.1	2004-04-19
OpenGIS® GML in JPEG 2000 for Geographic Imagery Encoding Specification (GMLJP2)	1.0.0	2006-01-20
OpenGIS® Grid Coverage Service Implementation Specification (GC)	1.0	2001-01-12
OpenGIS® Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture (SFA)	1.1.0	2005-11-30
OpenGIS® Implementation Specification for Geographic information - Simple feature access - Part 2: SQL option (SFS)	1.1.0	2005-11-30
OpenGIS® Location Service (OpenLS) Implementation Specification: Core Services (OLS Core)	1.1	2005-05-02
OpenGIS® Simple Features Implementation Specification for CORBA (SFC)	1.0	1999-06-02
OpenGIS® Simple Features Implementation Specification for OLE/COM (SFO)	1.1	1999-05-18
OpenGIS® Styled Layer Descriptor Implementation Specification (SLD)	1.0	2002-08-19
OpenGIS® Web Coverage Service Implementation Specification (Corrigendum) (WCS)	1.0.0	2006-03-31
OpenGIS® Web Feature Service Implementation Specification (WFS)	1.1	2005-05-03

OpenGIS® Web Map Context Implementation Specification (WMC)	1.1	2005-05-03
OpenGIS® Web Map Service Implementation Specification (WMS)	1.3.0	2006-03-15
OpenGIS® Web Service Common Implementation Specification (Common)	1.0	2005-05-03

Table 2
OGC completed specifications

METADATA

Everyday researchers may:
- be required to perform a critical analysis that requires the use of many data samples, from many sources around the world;
- need to pick the perfect dataset to fill a special requirement from thousands of datasets available through an on-line catalog;
- work with 50 datasets at the same time, covering the same area of interest;
- have data so old that no one in the organization will remember anything about it;
- be required to make a life-or-death decision using someone else's geospatial data.

In all of these situations researchers will need to know:

- What geographic data is available?
- Where is it?
- How to obtain it?
- Is it the best data available to make a decision?
- Is it up to date?
- Is it accurate?

These and many other questions require a good understanding of data. They require that data be well documented; they require complete and correct metadata. As we move into the age of spatial data infrastructures, knowledge about data is essential, allowing users to locate, evaluate, extract, and employ geospatial data. Diverse communities with a common understanding of metadata will be able to manage, share, and reuse each other's geographic data, making global interoperability a reality. The ISO Standard for Geographic Information - Metadata (ISO 19115) will provide this common understanding.

Metadata is not new; it is used every day in library card catalogs, Compact Disc (CD) jackets, user's manuals, and in many other ways. Geographic data has a long history using metadata. The marginalia on maps and charts are, of course, metadata. The title, source, scale, accuracy, producer, symbols, navigation notices, warnings, and all of the information found in the borders of maps and charts are metadata. This metadata is very user oriented; just about anyone can pick up a map, understand the metadata, and use the map. Map catalogs are another traditional use of metadata. Typically, map catalog metadata is limited to information such as area coverage, series identifiers (subject matter and scale), publication dates, and distribution information.

Why document geographic information?

Non-geographers using geospatial data*:* A revival in the awareness of the importance of geography and how things relate spatially, combined with the advancement in the use of electronic technology, have caused an expansion in the use of digital geospatial information and geographic information systems (GIS) worldwide. Increasingly, individuals from a wide range of disciplines outside of the geographic sciences and information technologies are capable of producing, enhancing, and modifying digital geospatial information. As the number, complexity, and diversity of geospatial datasets grow, a method for providing an understanding of all aspects of this data grows in importance.
Geospatial data is imperfect: Digital geospatial data is an attempt to model and describe the real world. Any description of reality is always an abstraction, always partial, and always just one of many possible "views". This view, or model, of the real world is not an exact duplication; some things are approximated, others are simplified, and some things are ignored - there is no such thing as perfect, complete, and correct data. To insure that data is not misused, the assumptions and limitations affecting the collection of the data must be fully documented. Metadata allows a producer to fully describe a dataset; users can understand the assumptions and limitations and evaluate the dataset's applicability for their intended use.
Increasingly, the producer is not the user*:* Most geospatial data is used multiple times, perhaps by more than one person. Typically, it is produced by one individual or organization and used by another. Proper documentation provides those not involved with data production with a better understanding of the data and enable them to use it properly. As geospatial data producers and users handle more and more data, proper documentation provides them with a keener knowledge of their holdings and allows them to better manage data production, storage, updating, and reuse.

Where should geographic information be documented?

Metadata is required in at least four different circumstances and perhaps in different forms to facilitate its use: in a catalog for data discovery purposes; imbedded within a dataset for direct use by application software; in a historical archive; and in a human readable form to allow users to understand and get a "feel" for the data they are using.
Catalogs: Metadata for cataloging purposes should be in a form not unlike a library card catalog or on-line catalog. Metadata in a catalog should support searches by subject matter/theme, area coverage/location, author/producer, detail/resolution/scale, currency/date, data structure/form, and physical form/media.
Historical Records: Metadata should support the documentation of data holdings to facilitate storage, updates, production management, and maintenance of geospatial data. Historical records provide legal documentation to protect an organization if conflicts arise over the use or misuse of geospatial data.

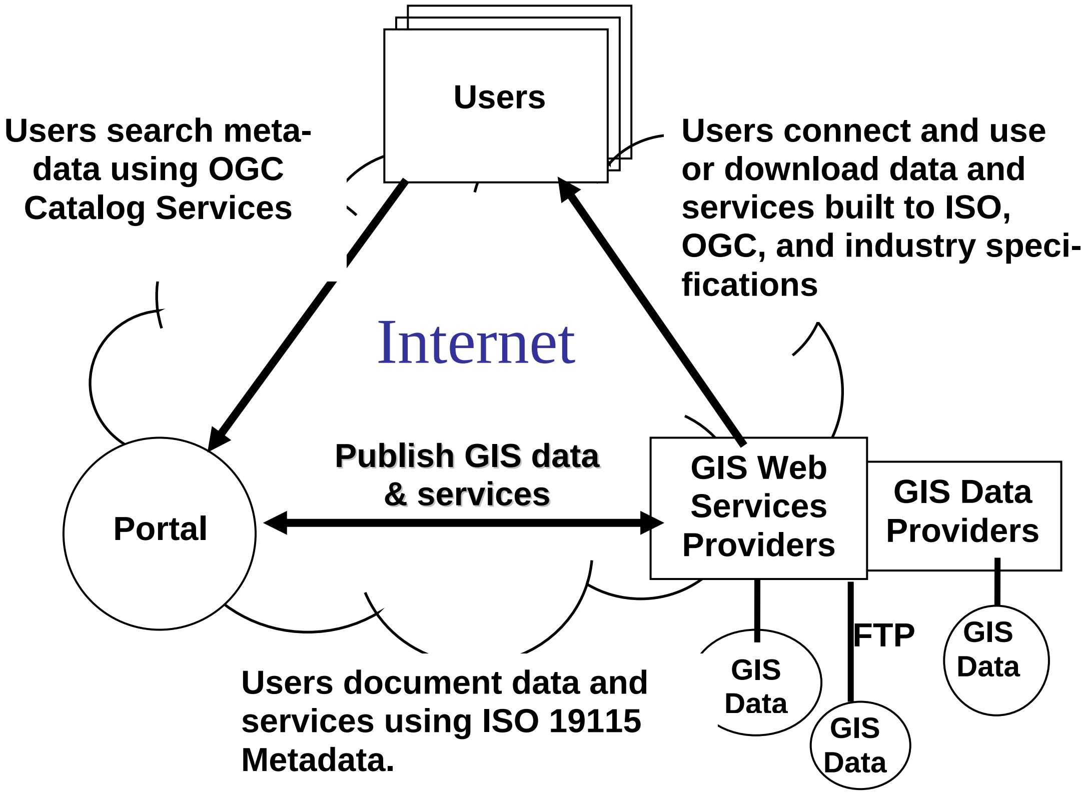
Within a geospatial dataset: Metadata should accompany a dataset and be in a form to support the proper application of geospatial data. GIS and other application software using data need to evaluate data as it applies to a situation. In this form the metadata may be incorporated into the structure of the data itself.
In a human readable form: Metadata in a form in which a computer can locate, sort, and automatically process geospatial data greatly enhance its use, but eventually a human must understand the data. One person's, or organization's, geospatial data is a subjective abstract view of the real world, it must be understood by others to ensure the data is used correctly. Metadata needs to be in a form which can be readily and thoroughly understood by users.

Applying geographic metadata

Metadata supports many applications; these can be classified into four primary functions (see Table 3):
Locate: Metadata enables users to locate geospatial information and allows producers to "advertise" their data. Metadata helps organizations locate data outside the organization and find partners to share in data collection and maintenance.
Evaluate: By having proper metadata elements describing a dataset, users are able to determine its "fitness for an intended use." Understanding the quality and accuracy, the spatial and temporal schema, the content, and the spatial reference system used, allows users to determine if a dataset fills their needs. Metadata also provides the size, format, distribution media, price, and restrictions on use, which are also evaluation factors.
Extract: After locating a dataset and determining if it meets users needs, metadata is used to describe how to access a dataset and transfer it to a specific site. Once it has been transferred, users need to know how to process and interpret the data and incorporate it into their holdings.
Employ: Metadata is needed to support the processing and the application of a dataset. Metadata facilitates proper utilization of data, allowing users to merge and combine data with their own, apply it properly, and have a full understanding of its properties and limitations.

	Catalog	Within Dataset	Historical Record	Human Readable Form
Locate	X		X	X
Evaluate	X	X	X	X
Extract	X	X		
Employ		X		X

Table 3
Metadata Usage Reference Matrix



Service Oriented Architecture

Putting it all together - Interoperability, Standards and Metadata.
Producers publish metadata for users to locate, evaluate, access and understand the data and services that meet their requirements. Interoperating through standards based architecture using metadata

BIOGRAPHICAL NOTES:

David M. Danko has over 40 years experience in the fields of cartography, aerial triangulation, remote sensing, GIS, and geospatial standardization. He served as Project Leader in ISO TC 211 for the development of several ISO metadata standards: ISO 19115:2003, ISO/TS 19139, and as editor for ISO 19115 Part 2 Imagery metadata. He is presently employed by ESRI, Inc as a Senior Consultant-GIS Standards. Mr. Danko is on the Board of Directors of the Cartography and Geographic Information Society (CAGIS) and serves as its representative to the American Congress of Surveying and Mapping; he is also a member of the American Society of Photogrammetry and Remote Sensing (ASPRS).

CONTACT INFORMATION:

David M. Danko
Environmental Systems Research Institute, Inc.
8620 Westwood Center Drive
Vienna, VA 22182 USA
+1-703-506-9515
ddanko@esri.com
http://www.esri.com