GEO-Interchange:

XML Standards for Geospatial Applications

A Pathfinder for LIS Professionals and Geospatial Data Managers

By

Karen A. Vagts

May, 2007

Introduction

Extensible Markup Language (XML), a tag-based language standard developed by the World Wide Web Consortium (W3C), provides great possibilities for data exchange and interoperability in a wide range of applications. One such application involves geographic information systems (GIS): the application of data to activities involving geography, cartography, and related sciences that describe the Earth.

Since the 1990s, a number of standards for geospatial XML have emerged. Already they have facilitated data exchange and manipulation. The purpose of this pathfinder is to summarize the major standards and to point to locations of further information about them in hope is that this will facilitate the selection of appropriate XML standards for a particular purpose.

XML is infinitely extensible and almost every physical information object has the potential to be geo-referenced. Therefore, a great many standards could be included here, ranging from XML schemas for telephone directories to restaurant surveys to census taking. But this document focuses on metadata standards designed primarily for GIS and cartographic materials (maps, gazetteers, and satellite images, among other objects). A few standards that were not developed specifically for geospatial applications are included because of the possibilities for extending geo-xml that they demonstrate. The metadata standards presented here share two XML elements: geographical location coordinates and place names.

This pathfinder assumes a basic familiarity with the concepts of both XML and GIS. Some general sources on these topics are provided in the section on General References.

Criteria for evaluation and selection of a Geospatial XML standard

Geospatial XML metadata standards vary by purpose, scope, standard type, authority, and rigidity. These features should be considered carefully when selecting an XML standard to work with; the selection of a particular standard determines the current and future interoperability of an XML-based application with other data sources. The same considerations also apply to the development of a new XML standard; specifying an XML vocabulary that parallels existingt existing standards helps ensure extensibility.

XML standards are not mutually exclusive. Multiple schemas or grammars can be combined or, alternatively, applied selectively. Some of the more complex standards are modularized, so that only certain portions need be applied. XML transformation tools provide flexibility in the choice and use of standards, although key elements should always be used in order to enable future cross walking of data among different metadata structures.

- **Generic versus Application Specific**

Generic standards, such as MathML for mathematics or ChemML for chemistry, contain basic tags that are recognized by a large professional community and for which a standard for general communications exists. In the geospatial community, such a standard is Geography Markup Language (GML).

Application-specific standards are more focused and may be tied to a proprietary product, such as GIS software, to a particular data collection, or to specific government or industry requirements. Such standards often build upon general schemas, adding additional elements or attributes or data entry requirements to respond to more specific needs.

- **Scope/Focus**

Broadly scoped XML standards such as GML and CSDG are comprehensive schemas that specify elements for all aspects of a geospatial systems, including graphic elements (expressed as polygons or shapes, points, and lines); these typically are complex and detailed standards designed for major projects with long time-frames. They often form the basis of, or are used with, GIS software or large, shared data depositories. Other standards, such as KML, are more scalable, featuring an extensive set of elements but *requiring* only a subset. The Dublin and Darwin Core standards are intended to be minimal elements sets although both have demonstrated a propensity to expand.

In the same way that traditional cartographic materials often are classified as maps, atlases, or gazetteers, geospatial XML standards also fulfill a particular function. The Getty Thesaurus of Geographic Names (TGN) and GNIS databases function as searchable gazetteers, being databases with an underlying DTD or schema. General georeferencing standards provide a full range of elements used to produce maps. Others, such as WMS, are protocols or services that repurpose geospatial data for other functions.

- **Standard Type**

XML standards vary by construction. Some are preset Document Type Definition (DTD) or schema (namespace) files that can be downloaded from a ftp site. Others are merely a set of described elements that that users can embed in an XML document type of their own design.

A important aspect of a standard's type is its strictness. A basic XML file of data elements can be quite loose, with minimal requirements about order, data typing, or repeatability of elements. DTDs and, even more so, schemas are strongly hierarchical and specific, which may seem restrictive but are necessarily for use with specific applications.

- **Licensing / Accessibility**

Many XML standards are open source, freely available from websites and available for use. Others require that you pay a fee to access them (particularly if they contain data, not just element definitions) or to use them. Some are free but require a initial acceptance of a use agreement or may request that the standard be credited wherever it is used.

- **Numeric Location Referencing**

In geolibraries, the most critical point of access is numeric location referencing or the establishment of "footprints." Numeric referencing usually is more specific and accurate than referencing by names and therefore is fundamental for interoperability.

Most numeric location referencing is expressed in longitude and latitude, using degree minutes or decimal degrees. In the contemporary world, most sources apply the Greenwich mean time or GMT system where the 0° prime meridian (longitude) is set at Greenwich, England and 0° latitude set at the equator. All other locations are relative to this 0°, 0° point of origin, using a reference grid.

Despite the seemingly universal application of this reference grid, many grids co-exist. Many older materials as well as some contemporary sources use other locations for the prime meridian (Greenwich only became a common standard in the 19th century). Map projections (adjustments to allow for the distortions when representing a globe on a flat surface) create

variable grids as do grid systems such as the Universal Transverse Mercator (UTM) and the British National Grid systems. XML standards vary in whether they specify a particular one. The issue can, however, be critical as these systems produce major differences in measuring linear distances and areas.

XML standards also vary as to how they specify footprints (Figure 1). Some use a single set of coordinate points for latitude and longitude. Others use bounding boxes (also known as minimum bounding boxes or rectangles) that enclose a set of location points; the bounding box describes the furthest most east, west, north, and south coordinates. Yet others may describe all points on a polygon or line. With regards to bounding boxes, a critical difference exists between bibliographic oriented standards and GIS: in fields 034 and 255c of MARC 21 catalog records, the coordinates used represent the extent of a printed map whereas in other standards, they describe the boundaries of a designated location.

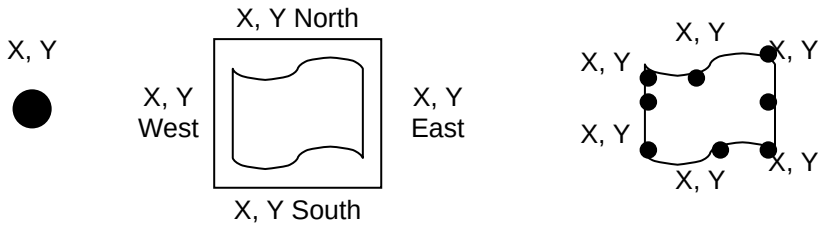

Figure 1: Comparison of various approaches to location referencing: A single point, usually in the center of the location (left); a bounding box, outlining the furthest extent of each side (center); a multipoint approach, essentially describing the location through multiple coordinate references (right).

- **Scale**

Geospatial data may be coded by map scale, e.g. 1:24,000 or 1:50,000 with attendant units, such as metric or imperial. Scale is most important with regards to printed maps; in GIS, scale is more tied to resolution or fineness of detail as well as output considerations for print or other media. Scale matters in selecting among XML standards and data sources because of the implications for georectifying (aligning) data collected from different sources and formats.

- **Accuracy**

Since many geospatial data applications involve precise measuring of areas and distances among them, accuracy becomes extremely important, especially for large-scale maps; the standards set by many government agencies specify a minimum level of accuracy and provide an element tag to record the level provided. While the standards themselves do not generate the appropriate degree of accuracy, consideration of this factor is important if when accessing, or contributing to, a geolibrary.

- **Place Names**

Along with location coordinates, place names traditionally have been the primary means of accessing spatial data. Unlike mathematical coordinates, however, place names are inexact and prone to change, controversy, and spelling variants in spelling. For this reason, many geospatial XML standards indicate a preferred place name authority system or gazetteer standard, such as the ones described here. Another option is to use a place reference number or code, such as

those used by the International Standards Organization (ISO) Standard ISO 3166 or the United States Geological Survey Geographical (USGS) Names Information System (GNIS) location reference numbers.

- **Authorities**

Some XML standards, like the Getty TGN or the USGS GNIS, are themselves authorities, e.g. standards for place names, providing content for use in other applications. Other standards do not contain data but may make recommendations on authorities to use; for example, the Dublin Core recommends using the TGN while the Library of Congress (LC) MODS schema prefers the GNIS and the MARC Code List for Geographic Areas. The TGN itself is an example of an authority that references many other authorities, including the GNIS.

- **Temporal Data**

Temporal data, or the specification of time periods, are hard to separate from place data. Time is inextricably tied to place names, cartographic projections, reference systems, and even location coordinates (which may change in response to a major geological, climatic, or political event). As with locations, temporal references do not conform to a single standard: many approaches to calendars, date ranges, and time spans exist. XML elements for this information constitute important access points and consequently some XML standards recommend the use of authorities for the naming of historical time periods.

- **Compliance with National and International Standards**

Most geospatial XML standards purport to comply with requirements established by government agencies or standards bodies such as the American Society for Testing and Materials (ASTM), the ISO, and National Information Standards Organization (NISO). Some geospatial XML standards are themselves national (CSDGM) or international (ISO) standards. The ability to comply with such standards can effect project funding or the ability to contribute to a geospatial database. Compliance is also helpful in ensuring long-term migration to other shared XML-based data sources.

Creating XML Standards

Because XML is a content-neutral, open-source standard for data, anyone can create an XML-based standard for geospatial applications. In doing so, however, it makes sense to consider existing applications so as to ensure future extensibility and interoperability. These standards also reflect significant research and experience, thereby eliminating the need to come down a time-consuming learning curve in developing a new standard.

The following list details basic XML elements that might be included in a geospatial XML DTD or schema. A more detailed generic model is offered by Hill (2006, pp. 177-181).

Suggested Element	**Considerations**
Source	Source of the information about the element, particularly its location features. Typical sources include thesauri, geodatabases, and field documentation.
Location Coordinates	One or more types of location footprints (single point coordinates, a bounding box, etc) can be used. These can be subdivided into child elements for different presentations (degrees minutes or decimal degree) of longitude and latitude or different referencing systems. A third coordinate or dimension, for vertical elevation or depth, also may be specified.
Place Name	Place name elements should reference a location name standard. They can be subdivided into child elements for contemporary names, ISO or GNIS location codes, historical versions, and spelling variants.
Scale	Attributes or child elements can be used to express units of measurement (metric or imperial or pixels).
Projection	Map projections should reference standard projection names (Mercator, Peters, Lambert).
Accuracy	Accuracy statements can include both quantitative levels based on a given system and comments about accuracy measures pertaining to the element.
Date	Separate date tags may be needed for both the date of the object described (descriptive data) and the creation and revision dates of the containing data record (administrative data). Temporal information can be subdivided into child elements for differencing calendaring systems (Georgian, Islamic), historical eras, and other temporal features.

In developing a custom XML standard, important decisions involve:

- The use of attributes or additional elements
- The extent of hierarchy
- The ability to repeat elements multiple times
- The need to "data type" elements (require data to be entered in a specific format, such as a date or number)

XML Standards for General Georeferencing

The following standards cover general geospatial elements. They usually are XML-based schemas or Document Type Definitions (DTDs) that describe the structure and scope of data elements rather than data files.

The standards shown here are a representative sampling of major geo-oriented XML standards, with a focus on applications used in the United States. Many of these standards serve as the basis for schemas tailored for specific locations, industries, or academic applications. Other countries have developed their own standards, although most are based on GML (described below).

Geography Markup Language (GML)

Author / Sponsor	Open Geospatial Consortium Inc (OGC) http://www.opengeospatial.org/
XML Standard Type	Schema
Date Introduced	2000
Current Version	GML 3.1.1
Documentation	http://www.opengeospatial.org/standards/gml/
XML Declaration / Specification Location	http://schemas.opengis.net/gml/3.1.1/base/
Purpose / Focus	To provide a vendor-neutral interface for the "modeling, transport, and storage of geographic information" for applications used by companies, government agencies and universities.
Description	The OGC is an international, voluntary standards organization comprised of corporations, educational institutions, and government agencies. Founded in 1994, the OGC is an outgrowth of the OpenGIS Project. It develops its standards in conjunction with W3C and ISO and is one of the major developers of geodata standards. GML is part of a detailed suite of OGC standards that the OGC has developed for interoperability. It is a complex and detailed schema, embedded in many software packages (Oracle, ARCGis, etc). that involve GIS. It offers a certification program for vendors which have developed products that comply with GML and other OGC standards. In developing GML, the OGC worked with the ISO ISO/TC 211 Geographic information/Geomatics standards committee (see below) and the ISO is slated to adopt GML as ISO Standard 19136 in 2007. GML also has been adopted by the United States National Information Exchange Model (NIEM), an XML-based metadata clearinghouse. Due to its scope and widespread adoption, GML is the benchmark against which other geospatial standards should be compared although its full implementation is primarily suitable for complex, long-term projects.
See Also	Lake, R. (n.d.). *Introduction to GML: Geography Markup Language*. http://www.w3.org/Mobile/posdep/GMLIntroduction.html

Content Standard for Digital Geospatial Metadata (CSDGM)

Author / Sponsor	Federal Geographic Data Committee (FGDC) http://www.fgdc.gov/
XML Standard Type	Schema
Date Introduced	1994
Current Version	Vers. 2 (FGDC-STD-001-1998)
Documentation	Federal Geographic Data Committee. (2000, May 1). *Content Standard for Digital Geospatial Metadata Workbook*. (For use with FGDC-STD-001-1998). Version 2.0. http://www.fgdc.gov/metadata/documents/workbook_0501_bmk.pdf
XML Declaration / Specification Location	http://www.fgdc.gov/metadata/fgdc-std-001-1998.xsd
Purpose / Focus	To provide a framework for metadata contributing to the U.S. National Spatial Data Infrastructure (NSDI) and for projects funded by U.S. Federal agencies.
Description	Since 1994, under Executive Order 12906, all digital geospatial data produced by federal agencies should be presented in the CSDGM format. The intent is to facilitate data exchange among many different agencies (state, local, tribal, and international communities as well as public, private, and academic sectors) and to integrate data used in programs such as the National Data Map. The FGDC, an interdisciplinary agency of the U. S. Federal Government, oversees this data coordination effort. The CSDGM is highly detailed and complex. All components that make up a GIS project or drawing can be recorded as separate data records; each element in a GIS database could warrant a separate metadata record. To assist in the creation of a record, the FGDC publishes a detailed workbook (referenced above). The complexity of the standard has generated proposals for a CSDGM "Lite" version and several state agencies have developed their own variants. Nonetheless, many federal agencies and state agencies have accumulated substantial datasets using this format, much of which is downloadable from their websites. The Federal portal Geospatial One Stop (http://gos2.geodata.gov/) offers one path to this data. The FGDC is working towards harmonizing the CSDGM with ISO 19115 (see below), which includes efforts to build a more integrated standard for North America along with Canada and Mexico.
See Also	Federal Geographic Data Committee (FGDC). (2006, August). *Geospatial Metadata Standards*. http://www.fgdc.gov/metadata/geospatial-metadata-standards

ISO 19115 – Geographic information

Author / Sponsor	Information Standards Organization (ISO) ISO/TC 211 Geographic information/Geomatics Committee http://www.isotc211.org/
XML Standard Type	Schema
Current Version	19115:2003
Documentation	http://www.iso.org/iso/en/prods-services/ISOstore/store.html (enter 19115 in the Search field)
XML Declaration / Specification Location	To provide an international standard for describing "objects or phenomena that are directly or indirectly associated with a location relative to the Earth."
Purpose / Focus	The ISO is the major international standards organization (of which American National Standards Institute or ANSI is the U.S. representative). Its committee TC 211, comprised of OGC, UN agencies, and various international professional bodies, oversees the development of ISO Standard *19115:2003 - Geographic information – Metadata*. This in turn is one of the approximately 40 components of ISO 19100, a set of standards for geographic information. ISO 19115 is as comprehensive as CSDGM in specifications (see above) but has a broader scope that goes beyond describing GIS functions; it is intended for use in all range of technical documents. ISO has adopted GML (see above) as the standard for the forthcoming ISO *19136 Geographic information – Geography Markup Language*. The ISO publishes many other standards that, while not strictly focused on geospatial issues, have a bearing on geographic document and metadata production. These range from specifications for country, region, and language codes to character sets as well as general XML standards. The ISO is one of the few standards authors referenced here that requires purchase of its documentation on standards.
See Also	Østensen, O. (2001, July). *The expanding agenda of Geographic information standards*. ISO Bulletin. http://www.iso.org/iso/en/commcentre/pdf/geographic0107.pdf

ADN Metadata Framework

Author / Sponsor	Digital Library for Earth System Education (DLESE) http://www.dlese.org/library/index.jsp
XML Standard Type	XML schema with controlled vocabularies
Date Introduced	1999
Current Version	0.6.50
XML Declaration / Specification Location	http://www.dlese.org/Metadata/adn-item/0.6.50/docs/xml-info.htm
Purpose / Focus	To describe resources used in learning environments that focus on Earth system education.
Description	DLESE was developed in partnership with the National Science Digital Library (NSDL), which has funded several DLESE projects. Its target audience consists of a distributed community of educators, scientists, and others and anticipated uses include lesson plans, learning modules, visualizations, and activities. ADN consists of a metadata set (framework) consisting of a main schema plus schemas for controlled vocabularies and categories. schemas. The geospatial aspect of the ADN was developed with respect to Dublin Core and FDGC. It is distinctive, however, for including specifications for planets and events (floods, hurricanes, and other phenomena) in its schemas. The standard's acronym of ADN stands for ADEPT/DLESE/NASA or Alexandria Digital Earth Prototype / Digital Library for Earth System Education / National Aeronautics and Space Administration. The standard reflects guidelines established by the American Association for the Advancement of Science (AAAS); Dublin Core; Gateway to Educational Materials (GEM); Learning Object Metadata (LOM); National Council for Geographic Education (NCGE); National Council of Teachers of Mathematics (NCTM); National Educational Technology Standards Project (NETS) from the International Society for Technology in Education (ISTE); and NSES: National Science Education Standards. The framework can be used with the XML-based DLESE Collection System (DCS) to catalog educational resources http://dcs.dlese.org/preview
See Also	Digital Library for Earth System Education. (2003, September 18). *ADN Geospatial Overview*. http://www.dlese.org/Metadata/adn-item/geospatial.htm DLESE Collections list. http://www.dlese.org/dds/histogram.do?group=subject

Darwin Core Metadata Georeferencing Elements

Author / Sponsor	Biodiversity Information Standards http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome
XML Standard Type	Schema
Date Introduced	1990s
Current Versions	1.2 – Generic version. 1.21 - Used by FishNet2 (for fish specimens), HerpNet (herpetological collections), ORNIS (birds) and MaNIS (Mammal Networked Information System) networks OBIS - Used by the Ocean Biogeographic Information System Version 2.0 is under development
Documentation	http://wiki.tdwg.org/twiki/bin/view/DarwinCore/DesignAndPurpose
XML Declaration / Specification Location	http://digir.net/schema/conceptual/darwin/2003/1.0/darwin2.xsd http://rs.tdwg.org/dwc/tdwg_dw_record_tapir.xsd
Purpose / Focus	To provide a minimal, unordered schema to describe natural history collection objects.
Description	Biodiversity Information Standards was formerly known as the Taxonomic Database Working Group (TDWG), a non-profit focusing on standards for data recording and exchange about organisms. The Darwin Core originally was developed by the University of Kansas Natural History Museum and Biodiversity Research Center. The Darwin Core's purpose is to support information exchange about geographic locations of species and the existence of specimens in natural history collections. An original focus of the standard was for biological data exchange by the Z39.50 protocol but this has been superseded in some implementations by the open source protocol Distributed Generic Information Retrieval (DiGIR). Although the standard's original intent was to be simple and unordered, the TDWG is considering additional elements to make it more compatible and interoperable with other standards. Proposals have been extended for curatorial and paleontological functions. A proposed expanded version of the Geographical Elements is available at http://wiki.tdwg.org/twiki/bin/view/DarwinCore/GeospatialExtension

Dublin Core Metadata Initiative Georeferencing Elements

Author / Sponsor	Dublin Core Metadata Initiative http://dublincore.org/
XML Standard Type	RDF Namespace
Date Introduced	1999
Current Version	DC Recommendation, Version 1.1 (December, 2006) Formal endorsements: ISO Standard 15836-2003 (February 2003): http://www.niso.org/international/SC4/n515.pdf NISO Standard Z39.85-2001 (September 2001): http://www.niso.org/standards/resources/Z39-85.pdf
Documentation	http://dublincore.org/documents/usageguide/elements.shtml
XML Declaration / Specification Location	http://dublincore.org/documents/dc-xml/ http://dublincore.org/2006/12/18/dces.rdf#coverage
Purpose / Focus	To extend the basic Dublin Core (DC) vocabulary for geospatial elements.
Description	The Dublin Core Metadata Initiative is an international group focusing on interoperable metadata standards to improve resource discovery and description. Although an interdisciplinary effort, it is hosted by the LIS-oriented Online Computer Library Center (OCLC) and is of particular interest to the LIS community as a means of providing interoperability among different bibliographic data sets and software structures. The most common aspect of DC is the Dublin Core Metadata Element Set, which contains fifteen elements for describing resources. The Georeferencing Elements extend the Coverage element, which defines the spatial or temporal aspects of a resource. The preferred thesaurus for this element is the Getty Thesaurus of Geographic Names (TGN) (see below).
See Also	The Electronic Cultural Atlas Initiative (ECAI) is a major implementation of DC which extends the core element set. The ECAI Digital Atlas is a federated, networked collection of scholarly and cultural projects linked through a metadata clearinghouse; the collections are visualized through a GIS viewer. http://ecaimaps.berkeley.edu/clearinghouse/

Geodatabase XML

Author / Sponsor	ESRI (Environmental Systems Research Institute, Inc.) http://www.esri.com/
XML Standard Type	Schema ((also published as a relational database model) Industry-specific extensions published as DTDs
Date Introduced	c. 2004 (with the introduction of ArcGIS 9.0)
Documentation	ESRI. (2007, January). *Geodatabase XML*. http://www.esri.com/software/arcgis/geodatabase/about/xml.html ESRI. *XML Schema of the Geodatabase*. http://downloads.esri.com/support/whitepapers/ao_/XML_Schema.pdf
XML Declaration / Specification Location	http://support.esri.com/index.cfm?fa=downloads.dataModels.gateway
Purpose / Focus	To facilitate information exchange between ArcGIS geodatabases and external systems.
Description	ESRI is the developer of ArcGIS, one of the major software developers of GIS software. The model for the data collected in an ArcGIS project is Geodatabase, which can be implemented either as XML files or as part of a relational database system. Typical data contents include geographic features, satellite images, surface modeling information, and surveying information. Geodatabase XML is intended to facilitate lossless data exchange between ArcGis data files organized within the Geodatabase framework and external applications. Although ArcGIS is a proprietary software system running on PC-based systems, the XML standard is publicly available. ESRI provides DTDs focused in particular applications, ranging from addressing to waterworks. Geodatabase XML reflects international geospatial metadata standards such as FGDC and ISO.
See Also	ESRI. (2006, January). *ArcGIS: Engineered for Interoperability*. An ESRI White Paper. http://www.esri.com/library/whitepapers/pdfs/arcgis-engineered-for-interoperability.pdf

Keyhole Markup Language (KML)

Author / Sponsor	Google http://code.google.com/apis/kml/
XML Standard Type	Schema
Date Introduced	2000s
Current Version	2.1
Documentation	Google, Inc. KML 2.1 Reference. http://code.google.com/apis/kml/documentation/kml_tags_21.html
XML Declaration / Specification Location	http://code.google.com/apis/kml/schema/kml21.xsd
Purpose / Focus	To display geographic data in Google Earth, Google Maps, and Google Maps for mobile, and other compatible applications.
Description	KML is a schema used for mapping applications produced by Google. It uses a hierarchical tree structure. Only a subset of the elements is strictly required. Google Maps and Google Maps for mobile use subsets of the full KML standard and not all elements work in these applications. KML increasingly is being used for applications outside of Google products; implementations have been done by NASA WorldWind, the National Geographic, the Smithsonian, and UNESCO.
Additional Sources	Brown, M. (2006). *Hacking Google Maps and Google Earth*. Indianapolis, IN: Wiley Pub. Google Earth. *Google Earth User Guide*. Version 4.0. http://earth.google.com/userguide/

LC Metadata Object Description Schema (MODS) Georeferencing Elements

Author / Sponsor	Library of Congress' Network Development and MARC Standards Office http://www.loc.gov/standards/marcxml/
XML Standard Type	Schema
Date Introduced	2002
Current Version	3.3
Documentation	MODS Schema, & Documentation. Documentation: http://www.loc.gov/standards/mods/mods-outline.html
XML Declaration / Specification Location	http://www.loc.gov/standards/mods/v3/mods-3-2.xsd
Purpose / Focus	To create original resource description records and to enable transportation of MARC 21 bibliographic catalog records from MARC to other metadata frameworks and back.
Description	MODS is one of several efforts by the LC to provide interoperability and to find a balance between the complexity of MARC records and the demand for simpler bibliographic metadata schemes such as DC (see above). Unlike MARCXML, which is a full XML-based variant of MARC21, MODS is abbreviated and does not guarantee round-trip conversion with MARC records. It is more user-oriented than MARC and more detailed than DC. The MODS Georeferencing elements are a subset of MODs elements, which in turn form a subset of MARC fields. The tag elements are language-based rather than numeric and are partially rearranged. MODS reflects its bibliographic origins, so its geospatial references tend to reflect standards such as LC place headings and MARC location codes. Geospatial coordinates usually refer to the extent of a printed map, not the outlines of a defined place. The MODS Location element refers to the physical storage location of the information object while {lace refers to the location of publication, rather than being a geospatial reference of content. Descriptive geographic information is found in the Subject element and in MODs sub-element categories for geographicCode, hierarchicalGeographic, and cartographics (scale, projection, coordinates). LC provides documentation and conversion tools for cross-walking data among MARC, MODS, DC, and other data standards. For place name referencing, the LC has traditionally maintained place name authority records but increasingly recommends use of standard references such as the GNIS (see below) and MARC or ISO location codes.

XML Standards for Gazetteers

Gazetteers are place name dictionaries or vocabularies. A number of online gazetteers have been developed using an XML-based schema as their fundamental structure or as a data exchange mechanism. Many gazetteers are used with specific gazetteer databases; the data often is available for downloading (sometimes for a fee). It is also possible to access the underlying gazetteer DTD or schema in order to analyze it or to use it to develop additional compatible records.

XML-based gazetteer schemas typically include elements for place names, location coordinates, place type (city, state, natural feature), place codes, and time period. They usually link to other place name entries as many locations have multiple identities. Some, like the TGN, use a taxonomic tree that cross references broader, narrower, and related terms expressed in a visual hierarchy.

Gazetteers are not in themselves GIS applications but can be used as reference lists for data field contents. These are often presented as pull-down lists in GIS applications. Many of the XML georeferencing standards described above recommend or require these vocabularies in their implementation and the place name elements are designed accordingly.

ADL Gazetteer Content Standard (GCS)

Author / Sponsor	Alexandria Digital Library Project (ADL) http://www.alexandria.ucsb.edu/
XML Standard Type	Schema (also published as a relational database model)
Current Version	3.2
Documentation	Guide to the ADL Gazetteer Content Standard version 3.2 (2004) http://www.alexandria.ucsb.edu/gazetteer/ContentStandard/version3.2/GCS3.2-guide.htm
XML Declaration / Specification Location	http://www.alexandria.ucsb.edu/gazetteer/ContentStandard/version3.2/ADL_gazetteer_entry_schema3.2.xsd
Purpose / Focus	To provide a comprehensive framework for recording descriptions of named geographic places.
Description	The ADL GCS schema was developed for use with the ADL Gazetteer, part of the comprehensive geospatial Alexandria Library Project at the University of California. Its scope is international but references U.S. specific standards such as GNIS and the National System for Geospatial-Intelligence (NSG). Because of the comprehensive nature of the ADL, GCS is designed to accommodate multiple names, allowing for multiple language variants and temporal developments. Unlike other gazetteer standards, it was developed in conjunction with an actual implementation of a gazetteer database and related GIS application. GCS is accompanied by the ADL Gazetteer Protocol, a XML-based query and response structure for querying distributed gazetteers
See Also	Janée, G. (2004, October). *The Alexandria Digital Library (ADL) Gazetteer Protocol*. http://www.alexandria.ucsb.edu/gazetteer/protocol/

Getty Thesaurus of Geographic Names (TGN)

Author / Sponsor	J. Paul Getty Trust Research Institute http://www.getty.edu/research/conducting_research/vocabularies/
XML Standard Type	XML Data File
Documentation	Getty Research Institute. (n.d.) TGN XML Data Dictionary. http://www.getty.edu/research/conducting_research/vocabularies/tgn/tgn_xml_dd06.pdf
XML Declaration / Specification Location	http://www.getty.edu/research/conducting_research/vocabularies/download.html
Purpose / Focus	To cataloging and reference geographic spatial data of cultural materials, such as art and architecture.
Description	Part of the Getty Vocabulary program, the TGN is fully populated thesaurus of over 912,000 records, each of which focus on a name place. Each record contains detailed information about a place term, including historical names, alternate spellings, preferred names, spatial coordinates. The data is organized in a faceted hierarchical structure. The content is based on extensive research of major authorities in all subjects, making the TGN a valuable tool for research as well as information organization and retrieval. The TGN website enables full online searching of the database. Downloadable data is available in XML, relational tables, and MARC formats for a fee. The TGN is among the most prominent of XML-compatible vocabularies and is recommended or required for use with many other XML-based standards for georeferencing, such as DC and ECAI (described above). Part of its strength is the extensive number of authorities of all types it references. Other place name vocabularies containing place names have been developed for specific purposes, ranging from natural history to education; the TGN is a good source to use alongside these sources to ensure maximum interoperability. The TGN documentation indicates that its geospatial accuracy is not perfect and is not to be used as a GIS tool.

U.S. Geographic Names Information System (GNIS)

Author / Sponsor	U.S. Geological Survey (USGS) http://www.usgs.gov/
XML Standard Type	XML Document
Documentation	http://geonames.usgs.gov/domestic/metadata.htm
XML Declaration / Specification Location	http://geonames.usgs.gov/docs/metadata/gnis.xml
Purpose / Focus	To maintain the official Federal government standard for gazetteer information on U.S. geological information.
Description	Managed by the USGS, the GNIS database contains information about physical and cultural geographic features in the United States, Antarctica, and related areas (excluding roads and highways). The data contains place name authorities developed by the Federal inter-agency, the U.S. Board on Geographic Names (BGN). In addition to providing the XML standard referenced above, the GNIS is also a fully-populated geographic reference, available as a publicly-accessible searchable database. Its contents are widely used in other gazetteers, including the Getty TGM and LC authorities; the GNIS Feature ID is often a reference element. The GNIS also is embedded in numerous efforts by the USGS and other Federal agencies to create, unify, and connect data models for geographically-related data, such as the Geospatial Blue Book and the National Map A parallel gazetteer for Foreign Place Names is the GEOnet Names Server (GNS), maintained by the National Geospatial-Intelligence Agency (NGA) and the BGN (http://Earth-info.nga.mil/gns/html/index.html). The GNIS XML standard is formatted in accordance with CSDGM requirements (see above).

Other XML Geospatial Applications

Reflecting a general awareness of and interest in geospatial data exchange, a number of other XML-based have emerged in recent years. Some are experimental and not yet implemented but all show interesting possibilities and fresh ways to extend the use of geospatial XML applications.

GeoClef

Author / Sponsor	The Cross-Language Evaluation Forum http://ir.shef.ac.uk/geoclef/
Purpose / Focus	To evaluate cross-language Geographic Information Retrieval (GIR) systems
Description	Geoclef is an experimental standard being developed by a group of researchers to combine cross-language information retrieval (CLIR) with GIR. GeoClef is an effort of The Cross-Language Evaluation Forum (CLEF), a multi-year European Community research initiative. Progress on Geoclef is presented at annual symposia. Work published to date indicates an interest in testing XML-based data.

Scalable Vector Graphics (SVG)

Author / Sponsor	World Wide Web Consortium (W3C) http://www.w3.org/
XML Standard Type	DTD
Date Introduced	1.1
Current Version	http://www.w3.org/TR/SVG11/
Documentation	SVG Working Group (2007, January). *Scalable vector graphics (SVG)* http://www.w3.org/Graphics/SVG/
XML Declaration / Specification Location	http://www.w3.org/TR/SVG11/
Purpose / Focus	To provide an XML-based interchange standards for vector based graphics
Description	SVG is the W3C's schema for describing two-dimensional vector and combined vector/raster graphics in XML. It is a generic standard, not specific to any industry or application, but is listed here because it has been used for mapping applications. SVG files are plain text files, so they can provide links between graphic elements and text-based data using coordinate reference points. SVG profiles are available for mobile devices (SVG Basic and SVG Tiny) and print applications. SVG 1.2 is currently under development; plans are to release it to a schema rather than a DTD.
See Also	Eisenberg, J. (2002). *SVG Essentials*. Sebastopol, CA: O'Reilly. Guo, Z., Zhou, S., Xu, Z. and Zhou, A. (2003). G2ST: A Novel Method to Transform GML to SVG. *Proceedings of the 11th ACM international symposium on advances in geographic information systems*, (pp. 161-168). Available: Association for Computing Machinery (ACM) database.

Web Map Service (WMS)

Author / Sponsor	Open Geospatial Consortium http://www.opengeospatial.org/
XML Standard Type	Schema
Current Version	1.3.0 (OGC 06-042)
XML Declaration / Specification Location	http://schemas.opengis.net/wms/1.3.0/
Purpose / Focus	To provide a protocol to create “maps of spatially referenced data dynamically from geographic information”
Description	WMS is not a standard for storing geospatial data, rather it is a protocol for converting GIS data to other formats, in this case graphic maps, for transfer over the internet. Examples of WMS implementation are NASA’s World Wind server (http://worldwind.arc.nasa.gov/). WML also used with GoogleEarth and Microsoft TerraServer.
See Also	A catalog of WML implementations is available at http://wms-sites.com/catalog

General References

General Note: All websites referenced in this document were accessed in April, 2007.

Arctur, D. & Zeiler, M. (2004). *Designing geodatabases: case studies in GIS data modeling*. Redlands, CA: ESRI Press.

Bartolo, L., Cole, T. W., Giersch, S., & Wright, M. (2005). NSF / NSDL Workshop on Scientific Markup Languages. *D-Lib Magazine, 11*(11). Available: http://www.dlib.org/dlib/november05/bartolo/11bartolo.html

Caplan, P. (2003). *Metadata fundamentals for all librarians*. Chicago: American Library Association.

Haas, S. C. (2004). X Marks the Spot: The Role of Geographic Location in Metadata Schemas and Digital Collections [Electronic Version]. *RLG DigiNews, 8*. Available: http://www.rlg.org/en/page.php?Page_ID=20492&Printable=1&Article_ID=1690

Hill, L. L. (2006). *Georeferencing : the geographic associations of information*. Cambridge, Mass.: MIT Press.

O'Reilly Press. (n.d.). *XML.com*. Retrieved from http://www.xml.com/

Phillips, Hugh. (2006, October). *Metadata Tools for Geospatial Data*. Available: http://sco.wisc.edu/wisclinc/metatool/

Ray, E. T. (2001). *Learning XML*. Sebastopol, CA: O'Reilly.

Woodley, M. S. (1998). Crosswalks The Path to Universal Access? In M. Baca & Getty Information Institute. (Eds.), *Introduction to metadata : pathways to digital information*. Getty Information Institute website. Available: http://www.getty.edu/research/conducting_research/standards/intrometadata/path.html

World Wide Web Consortium (W3C). (n.d.). *Extensible Markup Language (XML)*. Available: http://www.w3.org/XML/