

# Digital Library/ Metadata Terms to Know

**AAT (Art & Architecture Thesaurus):** Thesaurus used primarily by catalogers of visual resources and museum objects; offers a controlled vocabulary of terms used to describe physical attributes of art, materials, art movements, and periods. One of the three (soon to be four) Getty Vocabularies maintained by the Getty Research Institute.

**administrative metadata:** Technical or workflow data that is associated with a record. Examples of administrative metadata can be details on copyright and licensing information, preservation information, and storage format of the file. Typically, at least part of this information is considered 'not public' (for administrative use only), and would not be metadata that would be displayed in a record that the public user views.

**authority file:** An index, created either locally or on a more universal scale by catalogers, that incorporates rules that are used to determine uniform headings. Authority files most commonly deal with uniform titles and name authorities. SEE ALSO LC Authorities, TGN, ULAN.

**back-end server:** Server that stores data and runs a database, as opposed to the front-end server, which typically only provides security and a user interface. Examples of back-end servers are MySQL and Oracle.

**CDWA (Categories of Description for Works of Art):** XML metadata schema that is used specifically to describe works of art, architecture, and material culture. CDWA includes 512 categories (fields), a small group of which are considered 'core' and utilized in CDWA Lite.

**CCO (Cataloging Cultural Objects):** A manual that provides metadata standards on how to catalog and describe cultural works, specifically dealing with a subset of fields utilized in CDWA and VRA Core. SEE ALSO CDWA, VRA Core.

**DAMS (Digital Asset Management Systems):** Computer software or hardware systems that aid in the management of "digital assets," such as digital images, text files, videos, or music. These systems typically deal with the acquisition, storage, cataloging, and retrieval of digital files. Although DAMS are commonly thought of as used in libraries and archives, they are increasingly being used in the business setting as well.

**data dictionary:** A dictionary that describes the field names included in a metadata schema, as well as format rules and requirements for each field.

**descriptive metadata:** Data that describes the intellectual content of an object/record; is typically the data that is visible to the users of a system and is the most valuable in terms of searchability for the user.

**digital repository (also digital archive):** An electronic repository where digital files are archivally stored on servers. Typically these digital files are surrogates or originals of primary sources, and can range in type (text, image, audio, video). SEE ALSO DAMS.

**DTD (Document Type Definition):** An XML schema language that imposes restrictions on an XML document's structure—the elements that are to be used, and the arrangement of those elements. For example, the EAD DTD specifies the elements and arrangement that can be used to describe an archival collection. SEE ALSO EAD, element, XML.

**Dublin Core (DC):** A metadata schema that is arguably the most universal—it can work for the widest range of file types and the widest range of disciplines. However, due to its broad scope, it can be difficult to fully describe certain types of records in detail. Originated out of a workshop hosted by OCLC in 1995. There are two levels of Dublin Core: simple (15 basic fields) and qualified.

**EAD (Encoded Archival Description):** A standard for the encoding of archival finding aids which is endorsed and supported by the Library of Congress and the Society of American Archivists (SAA). EAD is capable of describing a collection from the top-most level (an entire collection) down to an individual item level. Each file is encoded into an XML file using a specified set of tags and rules. SEE ALSO finding aid, tags, XML.

**element:** A field in a metadata schema. SEE ALSO field, tags.

**element set:** Grouping of fields that make up a metadata schema. SEE ALSO metadata schema.

**ERMS (Electronic Records Management Systems):** Computer software or hardware systems that manage electronic files. Similar to DAMS, although typically the term ERMS is also used to describe a system that manages a much broader range of file types than a DAMS. For example, in the academic setting, an ERMS could manage emails (sometimes even text messages) sent by faculty/staff, press releases, newsletters, memos, and similar electronic documents.

**field:** A category for specific data within a metadata schema. SEE ALSO element, tags.

**finding aid:** A document that describes an archival collection. Can include biographical details, container lists, and indexes of names or topics. SEE ALSO EAD.

**folksonomy:** Metadata generated not by experts, but by the users and creators of the content. Controlled vocabularies are usually not employed. Examples of websites that use folksonomic tagging are Flickr, Picasa, del.icio.us, and blogging hosts such as Blogspot and Wordpress.

**institutional repository (IR):** A digital repository that collects and preserves an institution's intellectual output. The content of an IR will vary from institution to institution, but can include journal articles written by faculty/students, theses or dissertations, and even video or audio recordings of speeches or performances.

**LC Authorities (Library of Congress Authorities, also abbreviated LCAUTH):** Publicly searchable Library of Congress database that holds name, subject, and title authority records, mostly for published monographs. SEE ALSO authority file.

**MARCXML:** XML framework that allows users to work with MARC data in an XML form. SEE ALSO XML.

**MeSH (Medical Subject Headings):** Authority thesaurus for the cataloging of medically related files and texts. Maintained by the National Library of Medicine.

**metadata:** Data about data, or information about a particular object, file, or record.

**metadata mapping (metadata crosswalks):** The linking of similar metadata elements between two or more different metadata schemas. This is usually done so that data can be shared between two collections or institutions utilizing different metadata schemas, or so that metadata can be efficiently OAI harvested (metadata would be mapped to Dublin Core in this case). SEE ALSO OAI-PMH.

**metadata schema:** Set of elements and rules that together form a standard for the description of a record or set of records. SEE ALSO element set.

**METS (Metadata Encoding & Transmission Standard):** Schema that is used for encoding descriptive, administrative, and structural metadata for all types of digital objects. Expressed in XML, and maintained by the Library of Congress. Previously known as Making of America II (MOA).

**MODS (Metadata Object Description Schema):** Schema used for a bibliographic element set, based on a subset of MARC fields but uses language-based tags as opposed to numeric tags, as is the case with MARC. Expressed using the XML schema language.

**OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting):** Rules that provide for successful and efficient metadata harvesting by developing repository interoperability. Currently, a repository must have their metadata formatted into Dublin Core in order for their records to be able to be harvested using OAI. SEE ALSO metadata mapping.

**open source:** Software or a computer application that is available free of charge to the consumer, as opposed to vendor software where there is usually a use and licensing fee associated with the product.

**RDA (Resource Description & Access):** Newest edition (not yet released) of AACR2, which focuses much more on the description of digital objects than previous versions.

**social tagging:** SEE folksonomy.

**tags (tagging):** 1. The act of assigning a subject or topic to a particular text or image.  
2. In the XML and HTML languages, a tag is a formatted way of expressing an element:  
<bioghist></bioghist>

**TEI (Text Encoding Initiative):** Standard for the encoding of electronic texts.

**TGM (Library of Congress Thesaurus for Graphic Materials):** Thesaurus that includes terms for describing visual materials according to subject and genre. Maintained by the Library of Congress.

**TGN (Thesaurus for Geographic Names):** Authority file of international geographic names. One of the three (soon to be four) Getty Vocabularies, maintained by the Getty Research Institute.

**ULAN (Union List of Artist's Names):** Authority file of names of creators and patrons of the fine arts. One of the three (soon to be four) Getty Vocabularies, maintained by the Getty Research Institute.

**VRA Core:** A descriptive metadata element set created for use when cataloging cultural objects. Can function as an extension schema for METS. Developed by the Visual Resources Association. SEE ALSO METS.

**W3C (World Wide Web Consortium):** International consortium devoted to developing guidelines and protocols for the World Wide Web. Their mission is to ensure long-term growth for the Web.

**Web 2.0:** The current trend to enhance information sharing and creativity via the Internet. Technologies include the use of blogs, wikis, and social networking sites. Specific applications include del.icio.us, Flickr, Picasa, MySpace, Facebook, Twitter, Google Earth, and Craigslist.

**XML (eXtensible Markup Language):** A specification for creating markup languages, where users can specify their own elements. Enables in the sharing of structured data across the Internet. A stylesheet must be applied to an XML file before it can be rendered into HTML or any other browser-readable format. SEE ALSO XSL.

**XSL (eXtensible Stylesheet Language):** A transformation language that can be applied to an XML file which allows the file to transform into HTML or other various formats. SEE ALSO XML.