Glossary of Metadata Related Terms, Concepts, and Standards

(as compiled by Thom Shepard, WGBH Educational Foundation)

Metadata Related Terms

Application Profile	A set of metadata elements, policies, and guidelines defined for a particular application. The elements may be from one or more element sets, thus allowing a given application to meet its functional requirements by using metadata from several element sets including locally defined sets. For example, a given application might choose a subset of the Dublin Core that meets its needs, or may include elements from the Dublin Core, another element set, and several locally defined elements, all combined in a single schema. An Application profile is not complete without documentation that defines the policies and best practices appropriate to the application. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Authority Control	A set of rules or procedures that maintain consistency for accessing names or terms within a database. Means of establishing a consistent form of the name or concept through authority records. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Authority File	Collection of records that show the preferred form of a personal or corporate name, geographic region or subjects. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Catalog List	The simplest type of controlled vocabulary is a high-level categorization (or classification) scheme. At the time of input, one or more categories must be selected from the scheme and added to the document metadata. At the time of seeking information, the user does not have to think of keywords, but simply browses the list of categories and subcategories. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm**
Classification Scheme	A logical scheme for arrangement of knowledge, usually by subject. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Controlled Vocabulary	1. A prescribed set of consistently used and carefully defined terms. 2. Dictionary of terms related to a particular subject field. 1. **http://library.csun.edu/mwoodley/dublincoreglossary.html**

Term	Definition
Copyright	A bundle of exclusive rights conferred by a government on the creator of original literary or artistic works such as books, articles, drawings, photographs, musical compositions, recordings, films, and computer programs. International in scope, copyright grants the creator reproduction, derivation, distribution, performance, and display rights. **http://usinfo.state.gov/topical/econ/ipr/ipr-glossary.htm**
Crosswalk	A table that maps the relationships and equivalencies between two or more metadata formats, supporting the ability of search engines to search effectively across heterogeneous databases, promoting interoperability. Also referred to as Mapping. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Data Modeling	Process by which the data requirements within a given scope are grouped into objects called entities, relationships between those entities are documented, and a graphical representation called an entity-relationship diagram is produced. Logical data modeling is done during the analysis phase of a project. Physical data modeling is done during the design phase of a project. **http://www.state.tn.us/finance/oir/qa/stds/glossary/glossary.htm**
Descriptive Metadata	Metadata that supports the discovery of a digital object. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Dumb-down Principle	A rule for the application of Interoperability Qualifiers, which stipulates that qualifiers can refine but not extend the meaning of the element to which they are applied. The result is a loss of precision but still useful for the user or an application. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Element	One of the items that collectively form a metadata structure. Common elements are 'title', 'creator', 'date', and 'publisher'. Dividing data into elements allows users to carry out more accurate searches by searching on one element only. For instance, when looking for documents by Jennifer Green, searching the 'creator' field only will retrieve items by Jennifer Green only. It avoids items where the word 'green' appears in other contexts, as a subject, location etc. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm**
Element Set	A collection of discrete units of data or metadata. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Embedded Metadata	Metadata that is maintained and stored within the object it describes; the opposite of stand-alone metadata. **http://dublincore.org/documents/2001/04/12/usageguide**
Encoding Scheme	A scheme that controls the content, or 'value' of an element or element refinement, in order to clarify the meaning or improve resource discovery. These schemes include controlled vocabularies and formal notations or parsing rules. A value expressed using an encoding scheme will thus be a token selected from a controlled vocabulary (e.g. a term from a classification system or set of subject headings) or a string formatted in accordance with a formal notation (e.g., "2000-01-01" as the standard expression of a date). Encoding schemes are designed to be interpreted by machines or by human readers. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm**
Fair Use	Codified in the 1976 U.S. Copyright Law and frequently used by scholars, journalists, and librarians, the fair use provision permits the limited use of copyrighted scientific and artistic material to supplement or briefly illustrate oral or written commentary, literary or artistic criticism, or teaching materials. In determining that a use is fair, four factors must be considered: (1) the purpose and character of the use -- whether it is commercial or nonprofit; (2) the nature of the copyrighted material; (3) the amount of the total work used; and (4) the effect of the use upon the potential market -- whether or not the author is deprived of sales.

	http://usinfo.state.gov/topical/econ/ipr/ipr-glossary.htm
Field	Commonly used in database applications to describe a space in which data of the same type is entered (e.g. 'title' or 'price'), 'field' is a similar concept to 'element'. http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm
Flat file	Data files that contain records with no structural relationships. http://databases.about.com/library/glossary
Granularity	The level of detail at which an information object or resource is viewed or described. http://library.csun.edu/mwoodley/dublincoreglossary.html
Harvest	A system providing software architecture for gathering, indexing and accessing Internet information with certain protocol and descriptive metadata schema. http://www.geocities.com/metagateway/glossary.html
Information management	Information Management (IM) describes the measures required for the effective collection, storage, access, use and disposal of information to support agency business processes. The core of these measures is the management of the definition, ownership, sensitivity, quality and accessibility of information. These measures are addressed at appropriate stages in the strategic planning lifecycle and applied at appropriate stages in the operational lifecycle of the information itself. http://www.records.nsw.gov.au/publicsector/rk/glossary
Intellectual control	The control established over the informational content of records and archives resulting from ascertaining and documenting their provenance, and from the processes of arrangement and description.
Intellectual Property	Creative ideas and expressions of the human mind that possess commercial value and receive the legal protection of a property right. The major legal mechanisms for protecting intellectual property rights are copyrights, patents, and trademarks. Intellectual property rights enable owners to select who may access and use their property, and to protect it from unauthorized use. http://usinfo.state.gov/topical/econ/ipr/ipr-glossary.htm
Interoperability	The ability of different types of computers, networks, operating systems, and applications to work together effectively, without prior communication, in order to exchange information in a useful and meaningful manner. http://library.csun.edu/mwoodley/dublincoreglossary.html
Learning Object	Any entity, digital on non-digital, which can be used or referenced in technology-supported learning. http://library.csun.edu/mwoodley/dublincoreglossary.html
Meta Tag	The HTML element used to demarcate metadata on a Web page. http://library.csun.edu/mwoodley/dublincoreglossary.html
Metadata	A summary of information about the form and content of a resource. The term 'metadata' has been used only in the past 15 years, but has become particularly common with the popularity of the World Wide Web. The underlying concepts have been in use for as long as collections of information have been organized. Of particular interest to this Framework are the facets of metadata intended to support resource discovery and records management. http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm
Metadata Model	Structured information that expresses the intellectual content, intellectual property and/or an identifiable occurrence (instantiation) of an information resource. http://library.csun.edu/mwoodley/dublincoreglossary.html
Metadata Schema	Provides a formal structure designed to identify the knowledge structure of a given discipline.

Namespace	A unique name that identifies an organization that has developed an XML schema, identified by a Uniform Resource Identifier (URL or URN). **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Ontology	The working model of entities and interactions in some particular domain of knowledge or practices, such as electronic commerce or "the activity of planning." In artificial intelligence (AI), an ontology is, according to Tom Gruber, an AI specialist at Stanford University, "the specification of conceptualizations, used to help programs and humans share knowledge." In this usage, an ontology is a set of concepts - such as things, events, and relations - that are specified in some way (such as specific natural language) in order to create an agreed-upon vocabulary for exchanging information. **http://whatis.techtarget.com/definition/0,,sid9_gci212702,00.html**
Persistent Uniform Resource Locator	A public alias for a document. A PURL remains stable, even when the document's background URL changes. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Physical control	The control established over the physical aspects (such as format, quantity and location) of the archives and records in custody. **http://www.records.nsw.gov.au/publicsector/rk/glossary/**
Project	Projects are performed by people, constrained by limited resources, planned, executed, controlled, temporary, and unique. A project is a temporary endeavor undertaken to create a unique product or service; meaning it has a definite beginning and a definite end. A project is unique in that the product produced is different in some distinguishable way from all similar products. **http://www.state.tn.us/finance/oir/qa/stds/glossary/glossary.htm**
Project Plan	A project plan is a comprehensive plan that pulls together all the outputs of project planning activities, which include: project scope, project activities, activity sequence, activity durations, resources required for activities, project schedule, cost estimate, spending plan, and a quality plan. **http://www.state.tn.us/finance/oir/qa/stds/glossary/glossary.htm**
Protocol	A standard procedure or set of rules with which software and hardware systems must comply in order to be compatible. A network protocol must be observed by all users to allow data communications. Protocols deal with error handling, framing, and line control in transmitting and receiving packets. Dictionary of Multimedia by Brad Hansen
Provenance	1. The agency, office or person of origin of records, i.e. the entity which created, received or accumulated and used the records in the conduct of business or personal life. Also referred to as records creator. 2. The chain of custody which reflects the office(s) or person(s) that created, received or accumulated and used the records in the conduct of business or in the course of personal life. Identifying and documenting the provenance of records is an essential part of establishing their authenticity and integrity as evidence. 3. In archival theory, the principle of provenance requires that archives of an agency or person not be mixed or combined with the archives of another, i.e. the archives are retained and documented in their functional and/or organizational context. **http://www.records.nsw.gov.au/publicsector/rk/glossary/**
Qualifier	Something that describes or characterizes an object. It refines an element's meaning. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Record	1. In a database, a group of related items or fields, treated as a single unit of information. 2. Document containing data or information of any kind and in any form, created or received and accumulated by an organization or person in the transaction of business or the conduct of affairs and subsequently kept as evidence of such activity through incorporation

	into the record keeping system of the organization or person. Records are the information by-products of organizational and social activity. 1. Dictionary of Multimedia by Brad Hansen 2. **http://www.records.nsw.gov.au/publicsector/rk/glossary**
Related Rights	Also known as neighboring rights. Generally included under the umbrella of copyright in the United States, these are a bundle of exclusive rights provided to performers and producers of a sound recording or audio-visual work. **http://usinfo.state.gov/topical/econ/ipr/ipr-glossary.htm**
Relational	A collection of data files with at least one field in common, so that there is a means of relating any number of data units by using a common factor. Dictionary of Multimedia by Brad Hansen
Repository	A collection of resources that can be accessed to retrieve information May consist of several databases tied together by a common search engine. **http://databases.about.com/library/glossary**
Schema	1. A systematic, orderly combination of elements. A set of rules for encoding information that supports a specific community of users. 2. The place where definitions and restrictions of usage for properties are documented. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Semantics	pertaining to or arising from the meaning of words or other symbols
Specification	the detailed description of requirements, dimensions, materials, etc of a proposed standard
Stand-alone Metadata	Metadata that is created, maintained and stored independently of the object it describes. **http://dublincore.org/documents/2001/04/12/usageguide**
Standard	A rule that sets the baseline to measure and to provide quality and/or consistency. **http://www.state.tn.us/finance/oir/qa/stds/glossary/glossary.htm**
Sub-element	Term sometimes used to refer to an element refinement. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm**
Syntactic interoperability	Achieved by marking up our data in a similar fashion so we can share the data and so that our machines can understand and take the data apart in sensible ways. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Syntax	The form and structure with which metadata elements are combined. **http://library.csun.edu/mwoodley/dublincoreglossary.html**
Taxonomy	1. The science of classification, traditionally used to describe a hierarchical scheme for classifying plants and animals. More recently it has been borrowed to describe a classification scheme for organizing networked resources and supporting user-friendly navigation among them. Some taxonomies incorporate thesaurus features to augment the hierarchical structure. 2. Taxonomy (from Greek taxis meaning arrangement or division and nomos meaning law) is the science of classification according to a pre-determined system, with the resulting catalog used to provide a conceptual framework for discussion, analysis, or information retrieval. In theory, the development of a good taxonomy takes into account the importance of separating elements of a group (taxon) into subgroups (taxa) that are mutually exclusive, unambiguous, and taken together, include all possibilities. In practice, a good taxonomy should be simple, easy to remember, and easy to use. 1. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm** 2. **http://whatis.techtarget.com/definition/0,,sid9_gci331416,00.html**
Thesaurus	A controlled vocabulary designed to support information retrieval by guiding both the person

	assigning metadata and the searcher to choose the same terms for the same concept. A thesaurus conforming to ISO 2788 (=BS 5723) supports navigation and term selection by showing relationships between terms that are close in meaning. **http://www.e-envoy.gov.uk/publications/frameworks/metadata/glossary.htm**
UMID	Unique Material Identifier
Virtual	Describes the existence of an object, entity, or relationship in software or via networks, rather than in a tangible, physical condition.
XML	Extensible Markup Language. Subset of SGML designed to be transmissible over the Internet in such a way that document browsers do not need to access the document type definition to validate the document before display. As well as requiring all elements to be "well-formed", e.g. have both start and end tags present, the specification provides XML specific attributes and processing instructions that can be used to control the way documents are processed. **http://www.diffuse.org**

Digital Object Identifier: DOI

Digital Object Identifier	A system developed by the International DOI Foundation to identify and exchange intellectual property in the digital environment. It provides a framework for managing intellectual content, for linking customers with content suppliers, for facilitating electronic commerce, and enabling copyright management for all types of media. **http://www.doi.org/**

Dublin Core

Audience	Expresses the intended age or grade level of the resource; for example, a book intended for primary school readers. May include interest groups or special needs. Proposed element or extension, not part of basic 15 elements. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Contributor	An entity responsible for making contributions to the content of the resource. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Coverage	The extent or scope of the content of the resource. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Creator	An entity primarily responsible for making the content of the resource. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Date	A date associated with an event in the life cycle of the resource. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Description	An account of the content of the resource. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**

Format	The physical or digital manifestation of the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Identifier	An unambiguous reference to the resource within a given context.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Language	A language of the intellectual content of the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Publisher	An entity responsible for making the resource available.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Relation	A reference to a related resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Rights	Information about rights held in and over the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Source	A reference to a resource from which the present resource is derived.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Subject	The topic of the content of the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Title	The name given to the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
Type	The nature or genre of the content of the resource.
	http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml

METS

Administrative Metadata	1. how files were created and stored intellectual property rights original source object from which the digital library object derives provenance of the files comprising the digital library object may be either external to the METS document, or encoded internally. 2. Metadata used in managing and administering information resources, e.g., location or donor information. Includes rights and access information, data on the creation and preservation of the digital object. 1. **http://www.loc.gov/standards/mets/** 2. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
Descriptive Metadata	points to descriptive metadata external to the METS document or contains internally embedded descriptive metadata, or both. (Dublin Core, MARC, etc) Multiple instances of both external and internal descriptive metadata may be included in the descriptive metadata section. **http://www.loc.gov/standards/mets/**
File Groups	lists all files comprising all electronic versions of the digital object. Allows for nesting, to provide for subdividing the files by object version. **http://www.loc.gov/standards/mets/**

Structural Map	outlines a hierarchical structure for the digital library object links the elements of that structure to content files and metadata that pertain to each element. **http://www.loc.gov/standards/mets/**

MPEG-7

DDL	Description Definition Language. Allows the creation of new Description Schemes and, possibly, Descriptors and to allows the extension and modification of existing Description Schemes. **http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm**
Description Scheme	Specifies the structure and semantics of the relationships between their components. These components may be both Descriptors and Description Schemes. **http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm**
Descriptors	Representations of Features, that define the syntax and the semantics of each feature representation. **http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm**

SMEF

Data Dictionary	a set of metadata definitions for the information required in production, distribution and management of media assets. **http://www.bbc.co.uk/guidelines/smef/**

Standards

DRM	Digital Rights Management. Involves the description, layering, analysis, valuation, trading and monitoring of the rights over an enterprise's assets; both in physical and digital form; and of tangible and intangible value.
Dublin Core	A 15-element metadata element set intended to facilitate discovery of electronic resources. The Dublin Core has been in development since 1995 through a series of focused invitational workshops that gather experts from the library world, the networking and digital library research communities, and a variety of content specialties. **http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml**
IMS	"The original name of this initiative was the Instructional Management Systems (IMS) project. Over time, it became clear that the term 'instructional management systems' raised more questions than it answered." **http://www.imsproject.org/**

ISAN	The International Standard Audiovisual Number is a voluntary numbering system for the identification of audiovisual works, providing a unique, internationally recognized and permanent reference number for each audiovisual work. It consists of 16 hexadecimal digits divided into two segments: a 12-digit root segment followed by a 4-digit segment for the identification of episodes or parts when applicable. http://www.nlc-bnc.ca/iso/tc46sc9/isan.htm
MARC	Machine-Readable Cataloging Record. The MARC formats are standards for the representation and communication of bibliographic and related information (authority, holdings, classification, community information) in machine-readable form. MARC 21 grew out of the harmonization of USMARC and CAN/MARC, formerly national standards, and has emerged as an international standard. MARC21 is an implementation of the American National Standard, Information Interchange Format (ANSI Z39.2) and its international counterpart, Format for Information Exchange (ISO 2709). http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml
METS	The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation. http://www.loc.gov/standards/mets/
MPEG-7	MPEG-7 standardizes Descriptors (D) that define the syntax and the semantics of each "feature" represented by an audiovisual presentation, Description Schemes (DS) that specify the structure and semantics of the relationships between components, and a Description Definition Language (DDL) that allows the creation of new DSs and Ds and allows the extension and modification of existing DSs. MPEG-7 content may be delivered independently or together with the content it describes. http://www.diffuse.org
OAI	Open Archives Initiative metadata harvesting protocol uses XML to record metadata that distinguishes between data providers who represent the creators and service providers who meet the requirements of end users. Metadata can be based on the Dublin Core metadata set or alternative forms such as MARC library records. http://www.diffuse.org
ODRL	Open Digital Rights Language. Provides the semantics for DRM in open and trusted environments whilst being agnostic to mechanisms to achieve the secure architectures. http://xml.coverpages.org/odrl.html
RDF	(Resource Description Framework. The basic language for writing metadata; a foundation which provides a robust flexible architecture for processing metadata on the Internet and capable of enabling metadata exchange among application communities. http://library.csun.edu/mwoodley/dublincoreglossary.html
SCORM	Sharable Content Object Reference Model to incorporate many of emerging standards and specifications (e.g. LOM, IMS CP) into one content model. SCORM will provide a technical infrastructure that will allow content objects to be easily shared across multiple learning delivery environments. http://www.diffuse.org
SMEF	The BBC has defined a Standard Media Exchange Framework (SMEFTM) to support and enable media asset management ("MAM") as an end-to-end process across its business areas, from commissioning to delivery to the home. http://www.bbc.co.uk/guidelines/smef/
UPID	Universal Program Identifier for a single selected piece of completed content. For example, each episode and each version of that episode of a specific program would be associated

with a unique UPID. The UPID is intended to be applied after post-production, which differentiates it from a Universal Material Identifier (UMID) – a numbering system developed specifically for production and post-production use.