

Glossary

Metadata for Digital Collections: A How-to-Do-It Manual by Steven J. Miller

This glossary defines terms according to their usage in the book and in the context of metadata for cultural heritage digital collections. These terms sometimes have different nuances, aspects, or meanings in other contexts. For a few terms, definitions from Wikipedia and other sources have been used and acknowledged. See *also* the separate **Acronym Glossary** for this book and the short list of additional metadata glossaries at the end of the document.

Aboutness	What a resource is about. A term used to refer to the conceptual and other subject content of a resource when analyzing that content and representing it in metadata. Contrasted with <i>Ofness</i> in the subject indexing of images. See <i>also</i> <i>Ofness</i> ; <i>Isness</i> .
Administrative metadata	Data elements or properties used to administer and manage digital resources and collections. Subtypes include technical, preservation, rights, and use metadata.
Aggregation	The merging of metadata from a variety of original sources into a single repository or database. Metadata may be ingested by means of the OAI-PMH or other methods and usually needs some degree of conversion, processing, and normalization in order to function with other metadata in the aggregated context. See <i>also</i> <i>Harvesting</i> .
Analog resource	A resource in a physical or non-digital form. Often refers to the original physical version of a digitized resource. See <i>also</i> <i>Resource</i> .
Application profile	A document specifying an institution's, consortium's, or community's local implementation of a metadata element set or combination of elements from different standardized element sets. In addition to documenting a metadata element set, it may include definitions of, and further comments on, those elements, specifications of obligation, cardinality, required vocabularies and encoding schemes, input or content guidelines, examples, and/or mappings to other schemes. See <i>also</i> <i>Data dictionary</i> ; <i>Dublin Core Application Profile</i> ; <i>Metadata scheme</i> .
Associative relationship	A semantic relationship in some controlled vocabularies that connects related terms, terms that are related in some way other than by equivalence or hierarchical relationships. (Cross)references among <i>related terms</i> have been traditionally designated by the notation <i>See also</i> or <i>RT</i> . See <i>also</i> <i>Controlled vocabulary</i> ; <i>Semantic relationship</i> .
Authority file	A controlled vocabulary documenting the established, controlled, or "preferred" form of a name, term, or code. Usually includes references to variant, synonymous, or "nonpreferred" forms of names or terms, and may also include " <i>See also</i> " references to related names or terms. See <i>also</i> <i>Controlled Vocabulary</i> .
Best practice	A generic term that may refer to community or institutional recommendations, guidelines, and/or other specifications for implementing a metadata scheme. Best practices are typically documented and shared within specific communities of practice and sometimes more broadly. They may overlap with, or be more or less identical to, the kinds of documentation called application profiles or data dictionaries. See <i>also</i> <i>Application profile</i> ; <i>Data dictionary</i> .
Born-digital resource	A digital object/resource originally created in digital form, in contrast to being a digitized format of an originally analog or physical resource. See <i>also</i> <i>Resource</i> .

Cardinality	The number of times a metadata element or field may be repeated in a single metadata record, specified in metadata scheme documentation and in database configuration.
Cataloging	(a) Another term for resource description. (b) The process of creating library catalog records, typically using shared standards such as AACR2, RDA, MARC, LCSH, etc. See <i>also</i> Resource description.
Classification scheme	A controlled vocabulary in which terms are arranged into classes and subclasses or hierarchies of broader and narrower terms. Classification schemes make use of hierarchical semantic relationships, and often also of equivalence relationships, but not normally associative relationships. The term “taxonomy” is sometimes used synonymously with “classification scheme,” while sometimes the two are distinguished. Traditional bibliographic classification schemes in the library science tradition designate classes by means of “notation,” that is, numeric or alpha-numeric shorthand, frequently using decimals to extend the class notation. Bibliographic classification schemes may be enumerative or faceted. See <i>also</i> Controlled vocabulary; Taxonomy.
Complex digital resource	A digital resource/object composed of multiple digital files that together comprise a single logical whole. For example, a 150 page book digitized as 150 separate image files, plus images for the cover, etc., or a single logical sound recording comprised of multiple digital sound files.
Content management system (CMS)	A piece of software that enables the creation, storage, publication, and management of digital content. Most CMS’s designed specifically for digital collections include these functions for both the digital resources and the associated metadata representing those resources, as well as the user interface for searching, browsing, and navigating. CONTENTdm and Greenstone are two examples of such software.
Content standard	A type of data or metadata standard that gives rules or guidelines for entering metadata into specified elements or fields. Taken broadly they may also encompass metadata input guidelines in local metadata schemes and application profiles. Formal international content standards include RDA, CCO, DACS, and AACR2. Contrast with a value standard. See <i>also</i> Standard; Value standard.
Content versus carrier	A distinction between the intellectual or artistic content of a resource and one or more physical or digital vehicles for carrying or conveying that content to users. For example, the exact same textual content may be carried by an analog printed book, a large-print book, a digital e-book in ePub format, a PDF document, and an HTML file. The same image content may be carried by a physical photoprint or slide and by a digitized jpeg image file. The content is identical while the carrier differs.
CONTENTdm	A digital collection management software product available from OCLC. By most accounts the most commonly used digital collection content management system in the USA.
Controlled vocabulary	A pre-established list of terms, names, or codes that standardizes a single form to be used for consistent collocation and linking. Controlled vocabularies may range in complexity from simple flat <i>Lists</i> to complex <i>Ontologies</i> , with <i>Synonym rings</i> , <i>Authority files</i> , <i>Taxonomies</i> , <i>Classification schemes</i> , and <i>Thesauri</i> lying in between. Most controlled vocabularies address the problems of ambiguity and synonymy in natural human language. Controlled vocabularies may include <i>equivalence</i> , <i>hierarchical</i> , and/or <i>associative</i> semantic relationships and their corresponding cross-references, often notated as USE, USED FOR, SEE, SEE ALSO, BT, NT, and RT. See <i>also</i> Semantic Relationship and the names of the individual types of controlled vocabularies and semantic relationships stated in this entry.

Crosswalk	A document that shows the mapping of elements from one metadata scheme or element set to the equivalent elements in a different scheme, or to the closest thing to an equivalent when there is no exact or near-exact correspondence. Often used synonymously with the term <i>Mapping</i> . See also Mapping.
Cultural heritage institution	An umbrella term commonly used to encompass libraries, archives, museums, historical societies, and the like. Also called cultural memory institutions.
Data dictionary	A form of documentation of a metadata scheme, often in the form of a table. In addition to documenting a metadata element set, it may include definitions of those elements, specifications of obligation, cardinality, required vocabularies and encoding schemes, input or content guidelines, examples, and mappings to other schemes. Often used synonymously with the term application profile. See also Application profile; Metadata scheme.
Description	In the DCAM, a collection of metadata statements about a resource. Sometimes used as an equivalent term for metadata <i>record</i> . See also Description set; Record. Note: this technical usage of the term is not to be confused with the Dublin Core <i>Description</i> element or with other generic uses of the term "description."
Description set	In the DCAM, a set of one or more descriptions. A description set constitutes what has traditionally been called a metadata "record." See also Description; Record.
Descriptive metadata	Data elements or properties used to describe and provide access to digital resources and collections for end users, allowing users to search, browse, find, navigate, collocate, identify, interpret, and access those resources.
Digital collection	A collection of digital resources made available online. Resources may be born digital or be digitized versions of analog/physical resources. Availability may be restricted or unrestricted. The terms digital library, digital repository, and digital archive are sometimes used synonymously with digital collection.
Digital object	A resource in a computer-readable format that is either born-digital or is a digitized format of an analog/physical resource. See also Resource.
Document Type Definition (DTD)	A machine-readable technical document or file that formally defines a specific XML or SGML language by specifying the elements, attributes, and other characteristics of that language. In metadata terms, a DTD defines a specific metadata element set such as EAD. See also XML Schema.
DCMI Abstract Model (DCAM)	A specification of the components and constructs used in Dublin Core metadata. "It defines the nature of the components used and describes how those components are combined to create information structures. It provides an information model which is independent of any particular encoding syntax. Such an information model allows us to gain a better understanding of the kinds of descriptions that we are encoding and facilitates the development of better mappings and cross-syntax translations. ... The DCMI Abstract Model builds on work undertaken by the World Wide Web Consortium (W3C) on the Resource Description Framework (RDF)." Source: http://dublincore.org/documents/abstract-model/ .
Dublin Core Application Profile (DCAP)	(a) A document specifying an institution's, consortium's, or community's local implementation of Dublin Core. (b) More formally, a DCAM must conform to the current DCMI specifications that are in accord with the DCAM, RDF, and the Semantic Web, as expressed in http://dublincore.org/documents/profile-guidelines/index.shtml . See also Application profile.
Dublin Core Metadata Element Set	The fifteen original core set of metadata elements or properties comprising the IETF RFC 5013, ANSI/NISO Z39.85-2007, and ISO 15836:2009 standards and formally documented at http://dublincore.org/documents/dces/ . Often also used to

(DCMES)	encompass all current DCMI Metadata Terms, including the fifteen original properties (elements), additional properties, subproperties (refinements), vocabulary and syntax encoding schemes, classes, and terms related to the DCMI Abstract Model, all documented at http://dublincore.org/documents/dcmi-terms/ .
Dublin Core Metadata Initiative (DCMI)	An open organization engaged in the development of interoperable metadata standards that support a broad range of purposes and business models. The mission of the DCMI is to provide simple standards to facilitate the finding, sharing and management of information by developing and maintaining international standards for describing resources; supporting a worldwide community of users and developers; and promoting widespread use of Dublin Core solutions. Sources: http://dublincore.org/ and http://dublincore.org/about-us/ .
DTD	See Document Type Definition.
Element	A name or label for a selected piece of data that represents an attribute or characteristic of a resource judged important to a community or organization for serving various functions. Examples of Dublin Core elements include Title, Creator, Subject, Date, and Language. In a database context, a metadata element is a <i>field</i> . In current DCMI and RDF terminology, a metadata element is a <i>property</i> . In some metadata schemes, an element may include subelements or subproperties. An element may also be designated by, or known as, a <i>tag</i> , as is the case in MARC and XML. See also Element set; Field; Property.
Element set	A structured set of metadata elements or properties established by an organization for recording selected attributes of resources. Different communities and organizations establish different element sets to serve different users, needs, and types of resources. Also known as a <i>metadata scheme</i> or <i>data structure standard</i> . Internationally-established, standardized element sets include the data elements within the DCMES, MODS, VRA, and EAD schemes. See also Metadata scheme; Structure standard.
Encoding and exchange standard	A type of data or metadata standard that establishes formal methods of encoding metadata for machine readability, computer processing and data exchange. Formal international encoding standards include MARC and XML.
Encoding scheme	A general term used especially in the Dublin Core context to refer to a controlled vocabulary or data syntax specification governing the acceptable values for a given metadata element or property. See also Qualifier; Syntax encoding scheme; Vocabulary encoding scheme.
Equivalence relationship	A semantic relationship in some controlled vocabularies that connects equivalent or synonymous terms, relieving users of the need to think of and search on multiple equivalent terms for the same concept, multiple forms of name for the same person, and the like. In some controlled vocabularies one term is selected as the “preferred,” “authorized,” or “established” form to be used to collocate or gather together all instances of the same concept, person, etc. The other terms or forms of name are connected to the “preferred” term by links or cross-references. References and links among <i>equivalent terms</i> have traditionally been designated by the notation <i>See or Use</i> , directing users from the non-preferred to the preferred term, and <i>See from or Used for</i> , showing users the non-preferred terms connected to the preferred term. See also Controlled vocabulary; Semantic relationship.
Facets	Aspects or characteristics shared in common by many resources in a given context and expressed as categories. Common facets include topic, person, place, time period, form, genre, language, resource type. Facets may be used for user search refinements, browsing, and navigation in digital collection interfaces. In library and information science classification theory and thesaurus design, the term and concept

of “facets” have been defined and applied in highly formal, specialized, and more rigorous ways than the more popular usage of the term, but the fundamental underlying concept is the same.

Field	A unit of information in a database. In the context of metadata, a field instantiates a metadata property or element and its value, together comprising a statement about the information resource that the database record represents. <i>See also</i> Element; Property.
Granularity	The level of detail at which a resource is described or represented by a metadata record. For example, a digital collection as a whole can be represented by its own metadata record, as can each individual resource within that collection. A sound recording may include different pieces of music by different composers and performers. The sound recording “album” may be described or represented as a whole, and/or each individual piece of music may also be represented by its own metadata record. A series of books may be described at the series level and/or at the individual book level, and/or, if applicable, at the individual chapter level, with a separate metadata description/record created for each. The crucial thing in metadata creation or resource descriptions is to keep the properties and values in the metadata record consistent with the resource at the level at which it is being described or represented.
Harvesting	The process of acquiring metadata from a source external to the harvesting agent’s repository, usually by an automated method. This normally entails ingesting and processing metadata from various sources into a repository of aggregated metadata. The most common automated method of harvesting metadata is the OAI-PMH. <i>See also</i> Aggregation; Open Archives Initiative Protocol for Metadata Harvesting.
Hierarchical relationship	A semantic relationship in some controlled vocabularies that connects broader and narrower terms. Narrower terms may be generic members of a class or genus, specific instances of a class, or parts of a greater whole. References and links among <i>broader and narrower terms</i> have traditionally been designated by the notation <i>BT and NT</i> . In online systems hierarchies of broader and narrower terms may be used for broadening and narrowing search results, browsing and navigating through the hierarchy, and the like. <i>See also</i> Controlled vocabulary; Semantic relationship.
Identifier	A string of characters, usually in the form of relatively concise numeric or alpha-numeric designation or a uniform resource identifier, that uniquely identifies a resource. Uniqueness may be local or global. Standard identifiers include URI, URL, URN, DOI, ISBN, ISSN. Nonstandard, local identifiers include local accession or identification numbers, call numbers, and the like. <i>See also</i> Uniform Resource Identifier; Uniform Resource Locator.
Indexing	The process and result of analyzing the subject content of a resource and representing it by means of subject terms, names, and other terms that can be arranged in a browsable alphabetical index and otherwise used for various user information retrieval functions. Indexing is sometimes contrasted with “description,” in the sense that “indexed” metadata elements contain data that is formatted for machine processing, while “description” elements contain data in a free text format that may be searched by keyword but is not used for special machine processing.
Interoperability	The ability of two or more systems to work together and to exchange information with minimal loss of meaning. Metadata interoperability depends on conformance to various kinds of shared standards that allow metadata from various sources to operate and function together in ways that are useful to users.
Isness	What a resource <i>is</i> , in contrast with what it is <i>about</i> or <i>of</i> . A term that refers to the

genre or form of a resource. While not found in metadata literature, the term is used in the book as a parallel to the concepts of aboutness and ofness. See also Aboutness; Ofness.

Linked Data	A set of best practices for publishing and connecting structured data on the Web. Key technologies that support Linked Data are URIs, HTTP, and RDF. Source: http://linkeddata.org/faq . The current Web is a network of <i>linked documents</i> , whereas Linked Data structured in RDF enables a network of semantically-meaningful <i>linked data</i> . See also Resource Description Framework; Semantic Web; http://en.wikipedia.org/wiki/Linked_Data .
--------------------	---

Literal	In RDF, a string of lexical characters as raw text that is literally what it is in the sense that it is not a URI that links to a resource about which further statements may be made. Literals may be the objects of an RDF statement, but not the subject or predicate. See also Statement; String; Resource Description Framework. ----- “Literals are used to identify values such as numbers and dates by means of a lexical representation. Anything represented by a literal could also be represented by a URI, but it is often more convenient or intuitive to use literals. A literal may be the object of an RDF statement, but not the subject or the predicate. Literals may be plain or typed: A plain literal is a string combined with an optional language tag. This may be used for plain text in a natural language. As recommended in the RDF formal semantics, these plain literals are self-denoting. A typed literal is a string combined with a datatype URI. It denotes the member of the identified datatype's value space obtained by applying the lexical-to-value mapping to the literal string.” Source: http://www.w3.org/TR/rdf-concepts/#section-Literals
----------------	--

Mapping	The process of converting metadata from one scheme or element set into another. Moving data values from the elements in one scheme to the equivalent, near-equivalent, or best possible elements in the other scheme. Often used synonymously with the term <i>Crosswalk</i> . See also Crosswalk.
----------------	--

MARC (Machine Readable Cataloging)	A group of standards for encoding and processing library cataloging and other related data, including bibliographic, authority, holdings, classification, and community data. The term “MARC” is sometimes used as shorthand to refer to the tags or data elements that comprise the <i>MARC Format for Bibliographic Data</i> , which are based largely on the <i>Anglo-American Cataloguing Rules</i> and the catalog card model of metadata. See http://www.loc.gov/marc/ and http://www.loc.gov/marc/bibliographic/ecbdhome.html .
---	---

Metadata	Data <i>about</i> data; structured data elements or properties and their values that together make statements about resources for purposes of description, retrieval, administration, preservation, authentication, exchange, internal structuring, and other functions. Metadata is logically separate from the data or information resource it is about. See also Element; Property; Statement; Value.
-----------------	--

Metadata Object Description Schema (MODS)	A metadata element set and corresponding XML schema originally based on the MARC Format for Bibliographic Data, but which stands on its own as a metadata scheme. Used predominantly in some larger academic and research libraries and consortia for digital resource description. MODS consists of language-based tags and makes use of hierarchically-nested XML elements and subelements as well as element attributes. See http://www.loc.gov/standards/mods/ .
--	---

Metadata scheme	A general term often used to refer to (a) an internationally standardized and shared metadata element set or metadata structure standard, such as Dublin Core, MODS, VRA, or EAD, or (b) a set of specifications for an institution's, consortium's, or community's local application of a metadata element set, along with further guidelines and specifications, and documented in the form of an application profile,
------------------------	--

data dictionary, or the like. See also Application Profile; Data dictionary; Element set.

MODS	See Metadata Object Description Schema.
Namespace	<p>“In general, a namespace is a container that provides context for the identifiers (names, or technical terms, or words) it holds, and allows the disambiguation of homonym identifiers residing in different namespaces.” Source: http://en.wikipedia.org/wiki/Namespcae.</p> <p>“XML namespaces are used for providing uniquely named elements and attributes in an XML document. They are defined in a W3C recommendation. An XML instance may contain element or attribute names from more than one XML vocabulary. If each vocabulary is given a namespace, the ambiguity between identically named elements or attributes can be resolved.” Source: http://en.wikipedia.org/wiki/XML_namespace.</p>
OAI	See Open Archives Initiative.
Ofness	A term used to refer to the literal content depicted in an image when analyzing that content and representing it in metadata. Contrasted with <i>Aboutness</i> in the subject indexing of images. See also Aboutness; Isness.
One-to-One Principle	<p>The principle that a single metadata record or description should describe or represent one and only one resource. For example, separate records should be created for an original photograph and a digitized image of that photograph, and the two records linked to each other in a database. For various practical reasons, the principle is rarely followed in common practice for digital collections, with the results that properties and values pertaining to both the original and digital formats of a resource are commonly contained in a single record.</p> <p>The One-to-One Principle has been reaffirmed within the DCMI Abstract Model which holds that a <i>description</i> is “one or more statements about one, and only one, resource.” In the RDF, Linked Data, and Semantic Web contexts, a “resource” is anything about which a statement may be made; it may be a person, place, object, concept, and so on, in addition to a book, image, and the like. Statements in the form of RDF triples are processed individually rather than as aggregates in the form of a “record.” In this context, the One-to-One Principle takes on a somewhat different and clearer aspect and can be maintained precisely.</p> <p>See also Record; Resource Description Framework; Statement.</p>
Ontology	<p>A formal model of the entities that exist is a specified domain and the relationships among those entities. Ontologies are typically intended to be machine-actionable in a Web environment and to allow Semantic Web applications to make logical inferences about relationships among entities not explicitly stated by a human. The primary W3C format for encoding ontologies is OWL, the Web Ontology Language.</p> <p>-----</p> <p>“In computer science and information science, an ontology is a formal representation of knowledge as a set of concepts within a domain, and the relationships between those concepts. It is used to reason about the entities within that domain, and may be used to describe the domain. In theory, an ontology is a “formal, explicit specification of a shared conceptualisation”. An ontology provides a shared vocabulary, which can be used to model a domain — that is, the type of objects and/or concepts that exist, and their properties and relations. Ontologies are the structural frameworks for organizing information and are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, and information architecture as a form of knowledge representation about the world or some part of it.” Source: http://en.wikipedia.org/wiki/Ontology_(information_science)</p>

Open Archives Initiative (OAI)	An organization that develops and promotes interoperability standards to facilitate the efficient dissemination of content. OAI has its roots in the open access and institutional repository movements. Its work has expanded to also promote broad access to digital resources for eScholarship, eLearning, and eScience. See http://www.openarchives.org/ . See also Open Archives Initiative Protocol for Metadata Harvesting.
Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)	A computer-actionable, low-barrier mechanism for digital repository interoperability, widely used in the cultural heritage metadata world. Data Providers are repositories that expose structured metadata via OAI-PMH. Service Providers then make OAI-PMH service requests to harvest that metadata. See also Harvesting; Open Archives Initiative.
OWL	See Web Ontology Language.
Preservation metadata	Data elements or properties containing information needed for the long-term preservation of digital resources and collections, including their migration to other digital formats as software and hardware change over time.
Property	A selected attribute or characteristic of a resource that forms the basis for a metadata scheme or element set, and one of the three components of an RDF triple that makes a statement about a resource. Also known as a metadata element, tag or field name. Common cultural heritage digital resource properties include <i>title</i> , <i>creator</i> , <i>subject</i> , <i>date</i> , <i>type of resource</i> , <i>language</i> , and so on. See also Element, Property-value pair; Statement; Value.
Property-value pair	A combination of a metadata property (element, tag, field) and a metadata value that together comprise a metadata statement about a resource. Two of the three components of an RDF triple. See also Property; Statement; Value.
Qualified Dublin Core	The Use of Dublin Core elements or properties with the addition of qualifiers, which include element refinements or subproperties and encoding schemes, whether vocabulary encoding schemes or syntax encoding schemes. See also Qualifier; Refinement; Encoding Scheme.
Qualifier	A refinement or restriction of the meaning or scope of a metadata element giving that element greater specificity, or the specification of a vocabulary or syntax encoding scheme that applies to an element. A legacy term in the DCMI, supplanted by the terms <i>subproperty</i> , <i>vocabulary encoding scheme</i> and <i>syntax encoding scheme</i> . See also Refinement; Encoding Scheme.
RDF	See Resource Description Framework.
Record	The basic unit of metadata in a traditional database context. A record is comprised of a set of elements, fields, or tags and their values that together constitute a set of statements about one or more resources. A metadata record describes or represents an information resource and may function as a surrogate for that resource in a metadata database. Records much be encoded in some machine readable encoding syntax or standard such as OAI Dc XML, MODS XML, or MARC 21 in order to be processed and exchanged by computers.
Refinement	A restriction on the meaning or scope of a metadata element, giving that element greater specificity. For example, the Dublin Core Coverage element may use the Spatial or Temporal element refinement. In current DCMI terminology, an element is called a property and an element refinement is called a subproperty. See also Qualifier.
Registry	A place on the Web in which metadata vocabularies, properties, subproperties,

controlled vocabularies, and the like are formally declared, defined, and maintained, usually by an authorized agency. For compatibility with RDF, Linked Data, and the Semantic Web, each registered vocabulary term must be formally identified by a URI in the form of a URL.

Repeatability	See Cardinality.
Resource	A generic term for any thing which can be identified and about which a statement can be made and a metadata description created. In cultural heritage metadata context, a resource is typically a digital or analog information container such as a text (book, pamphlet, letter, diary, digital text file, etc.), still image (photograph, slide, etching, painting, digital image file, etc.), moving image (film reel, DVD, digital video file), sound recording (vinyl record, cassette tape, CD, digital audio files, etc.) and the like. A collection of resources can itself be regarded as a resource about which a metadata description may be created, as can smaller component parts of any resource. In the RDF, Linked Data, and Semantic Web contexts, a “resource” may be one of the types of things stated above, but it may also be an abstract concept, person, place, object, and so on: anything which can be identified and about which a statement can be made. See <i>also</i> Resource description; Resource Description Framework; String.
Resource description	The process of describing or creating metadata for a resource, also known as cataloging. Metadata is usually created in the form of a database or XML record that represents the resource and consists of a set of properties (elements, fields, or tags) and their values, together comprising a set of statements about the resource. In the RDF, Linked Data, and Semantic Web contexts, statements are not necessarily aggregated into “records.” See <i>also</i> Cataloging.
Resource Description Framework (RDF)	A model or framework for making statements about resources in the form of computer-processable triples consisting of a <i>subject</i> which is a resource, a <i>predicate</i> which is a property, and an <i>object</i> which is a value. A web-based model in which resources are preferably identified by a URI in the form of a URL. An official W3C Recommendation and one of the basic building blocks of the Semantic Web and Linked Data. See <i>also</i> http://en.wikipedia.org/wiki/Resource_Description_Framework and http://www.w3.org/RDF/ . See <i>also</i> Linked Data; Property; Resource; Semantic Web; Statement; Value.
Rights metadata	Data elements or properties containing information about ownership, copyright, restrictions on access and use pertaining to digital resources and collections.
Schema	See Metadata scheme.
Semantic relationship	In traditional controlled vocabularies, semantic relationships denote equivalence, hierarchical, and associative relationships among terms, providing users with synonym control and a network of references and links between nonpreferred terms and preferred terms used for collocation; broader and narrower terms; and related terms. References, also called cross-references, have traditionally used such notation as USE, USED FOR, SEE, SEE ALSO, BT, NT, and RT. In Web-based systems, these references or links may be instantiated in the form of hyperlinks and various search, browse, and navigation structures. See <i>also</i> Associative relationship; Controlled vocabulary; Equivalence relationship; Hierarchical relationship.
Semantic Web	“The Semantic Web is a “web of data” that enables machines to understand the semantics, or meaning, of information on the World Wide Web. It extends the network of hyperlinked human-readable web pages by inserting machine-readable metadata about pages and how they are related to each other, enabling automated agents to access the Web more intelligently and perform tasks on behalf of users.

The term was coined by Tim Berners-Lee, the inventor of the World Wide Web and director of the World Wide Web Consortium, which oversees the development of proposed Semantic Web standards. He defines the Semantic Web as "a web of data that can be processed directly and indirectly by machines." The term "Semantic Web" is often used more specifically to refer to the formats and technologies that enable it. These technologies include the Resource Description Framework (RDF), a variety of data interchange formats (e.g. RDF/XML, N3, Turtle, N-Triples), and notations such as RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain. Many of the technologies proposed by the W3C already exist and are used in various contexts, particularly those dealing with information that encompasses a limited and defined domain, and where sharing data is a common necessity, such as scientific research or data exchange among businesses." Source: http://en.wikipedia.org/wiki/Semantic_web.
See *also* Linked Data; Resource Description Framework.

Shareability	The ability of metadata records and statements to be shared among various metadata communities, repositories, and systems. Requires use of shared metadata element definitions and intended scope, shared vocabularies and schemes, and creation of metadata that can retain its meaning and functionality after being taken out of its original closed-system context. Overlaps with the concept of <i>Interoperability</i> .
Simple (unqualified) Dublin Core	The fifteen original Dublin Core elements or properties, used without qualifiers (element refinements or encoding schemes).
Simple Knowledge Organization System (SKOS)	"A family of formal languages designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary. SKOS is built upon RDF and RDFS, and its main objective is to enable easy publication of controlled structured vocabularies for the Semantic Web. SKOS is currently developed within the W3C framework." Source: http://en.wikipedia.org/wiki/SKOS .
SKOS	See Simple Knowledge Organization System.
Standard	A formal specification of norms or requirements for some kind of application, normally established in the form of a document. In the context of cultural heritage metadata, standards are often divided into data structure, value, content, and encoding/exchange standards. The term "standard" is most commonly used to refer to formal international and national standards, but it may also be used more loosely to encompass less formal but documented community, consortial, and even local institutional standards, guidelines, and best practices. See <i>also</i> Structure, Content, Value, and Encoding/exchange standards.
Statement	An assertion made about a resource, comprised of a property-value pair associated with a resource. The statement is the basic unit of metadata. A collection of statement comprises a <i>description</i> or a <i>record</i> . In RDF a statement is a "triple" comprised of a <i>subject</i> which is a <i>resource</i> , a <i>predicate</i> which is a <i>property</i> , and an <i>object</i> which is a <i>value</i> . See <i>also</i> Property; Property-value pair; Resource; Value; Triple.
String	A string of lexical characters that computers can usually process only as a string of characters and not as a dynamic link. In some cases, a string of characters may function as a dynamic link when they conform exactly to an established form such as a vocabulary or syntax encoding scheme, in which case the character string is "typed," that is, designated as belonging to a specific type of encoding scheme. Most current library catalogs link controlled forms of names and subject terms by

matching exact strings of lexical characters, including punctuation and spacing, between bibliographic and authority records. This is much less efficient and flexible than using a unique identifier instead of a string to represent a concept, person, place, thing, etc. In RDF, URIs in the form of URLs are used whenever possible instead of character strings, so that software applications can process and link them in various dynamic ways. See *also* Literal; Resource Description Framework.

Structure standard	A type of data or metadata standard that specifies a set of <i>properties</i> or data/metadata <i>elements</i> , along with definitions and sometimes additional specifications or recommendations. In the cultural heritage metadata context, structure standards are for all intents and purposes the same thing as an <i>element set</i> . Formal international structure standards include DCMES, MODS, and VRA elements. See <i>also</i> Element set.
Subject heading list	A controlled vocabulary of subject terms, primarily associated with libraries. Subject heading lists may include all three traditional semantic relationships (equivalence, hierarchical, and associative), and in that respect are indistinguishable from the thesaurus as traditionally defined. But they are frequently distinguished from thesauri in that subject heading lists tend to be more global in their scope of subject coverage, make heavy use of subdivisions, and are more likely to include compound, inverted, and phrase-like terms than are most traditional thesauri.
Synonym ring	A controlled vocabulary used to connect synonymous terms, or terms treated as synonymous, for retrieval by a search engine. None of the terms is selected as “preferred” and used for collocation. All terms are therefore equal and thus constitute a kind of continuous “ring.” Synonym rings make use of the equivalence semantic relationship but not the hierarchical or associative relationships. See <i>also</i> Controlled vocabulary.
Syntax encoding scheme	A general term used especially in the Dublin Core context to refer to a data syntax specification governing the acceptable values for a given metadata element or property. A typical syntax encoding scheme is the <i>W3C Date and Time Format</i> that specifies a consistent format for recording dates. Syntax encoding schemes are a form of data value standard and may be distinguished from vocabulary encoding schemes. See <i>also</i> Encoding scheme; Value standard; Vocabulary encoding scheme.
Taxonomy	A general term encompassing many different kinds of arrangements of terms into categories, most commonly hierarchical categories of broader and narrower terms. Often used synonymously with the term “classification scheme.” Sometimes used in current parlance to designate hierarchal website organization and navigation systems, in contrast to traditional bibliographic classification schemes, especially the library science tradition. See <i>also</i> Classification scheme; Controlled vocabulary.
Technical metadata	Data elements or properties containing technical information about digital resources and collections such as type of scanner used, scanning resolution, and image editing specifications.
Thesaurus	A controlled vocabulary of terms employing all three traditional semantic relationships: equivalence, hierarchical, and associative. The most complex of the traditional types of controlled vocabularies. Thesauri are typically created and used for a specific subject domain or information community. See <i>also</i> Controlled vocabulary; Semantic relationship; Subject heading lists.
Triple	In RDF, a data structure consisting of a <i>subject</i> which is a “resource,” a <i>predicate</i> which is a “property,” and an <i>object</i> which is a “value,” which together constitute a formal <i>statement</i> . In RDF, URIs in the form of URLs are usually required for the subject/resource and predicate/property, and as much as possible also for the

object/value in each triple in order to allow Semantic Web and Linked Data applications to most powerfully process and link those triples. *See also* Property; Resource Description Framework; Statement; Value.

Triplestore	<p>“A purpose-built database for the storage and retrieval of Resource Description Framework (RDF) metadata. Much like a relational database, one stores information in a triplestore and retrieves it via a query language. Unlike a relational database, a triplestore is optimized for the storage and retrieval of many short statements called triples, in the form of subject-predicate-object, like "Bob is 35" or "Bob knows Fred". Some triplestores can store billions of triples.” Source: http://en.wikipedia.org/wiki/Triplestore.</p>
Uniform Resource Identifier (URI)	<p>“In computing, a Uniform Resource Identifier (URI) is a string of characters used to identify a name or a resource on the Internet. Such identification enables interaction with representations of the resource over a network (typically the World Wide Web) using specific protocols. Schemes specifying a concrete syntax and associated protocols define each URI. One can classify URIs as locators (URLs), or as names (URNs), or as both. A Uniform Resource Name (URN) functions like a person's name, while a Uniform Resource Locator (URL) resembles that person's street address. In other words: the URN defines an item's identity, while the URL provides a method for finding it. The ISBN system for uniquely identifying books provides a typical example of the use of URNs.” Source: http://en.wikipedia.org/wiki/Uniform_Resource_Identifier. RDF specifies the use of URIs in the form of HTTP-formatted URLs. <i>See also</i> Resource Description Framework; Uniform Resource Locator.</p>
Uniform Resource Locator (URL)	<p>“In computing, a Uniform Resource Locator (URL) is a Uniform Resource Identifier (URI) that specifies where a known resource is available and the mechanism for retrieving it. In popular usage and in many technical documents and verbal discussions it is often incorrectly used as a synonym for URI.” Source: http://en.wikipedia.org/wiki/Uniform_Resource_Locator. RDF specifies the use of URIs in the form of http:// formatted URLs. <i>See also</i> Resource Description Framework; Uniform Resource Locator.</p>
Use metadata	Data elements or properties containing information about the use of digital resources and collections, such as user search logs and number of views.
Value	The content of a metadata property, element, tag, or field that applies to the specific resource being described. For example, the value of a <i>Date</i> property could be “2011” or “1954-07-07”; the value of a <i>Language</i> property could be “English,” “Chinese,” or “eng.” One of the three components of an RDF triple that makes a statement about a resource. <i>See also</i> Property; Property-value pair; Statement.
Value standard	A type of data or metadata standard that specifies the allowable values for specified properties or elements. Value standards are typically either controlled vocabularies or syntax encoding schemes. Contrast with content standards. Formal international value standards include AAT, TGN, LSCH, DCMIType, and W3C-DTF. <i>See also</i> Content standard; Encoding scheme; Standard.
Visual Resources Association Core Categories (VRA)	A metadata element set, called <i>core categories</i> , designed especially to represent important attributes or properties of physical visual resources and analog and digital images of those resources. Especially applicable to museum objects, works of art and architecture, and images of these. Used especially by museums but also many digital libraries. The VRA 3.0 and 4.0 schemes differ in several significant ways. Both make a sharp distinction between works and images of works, and they thereby strictly maintain the One-to-One Principle. <i>See also</i> VRA 3.0 and VRA 4.0.
Vocabulary	A general term used especially in the Dublin Core context to refer to a controlled

encoding scheme	vocabulary specification governing the acceptable values for a given metadata element or property. May be distinguished from a syntax encoding scheme. See <i>also</i> Controlled vocabulary; Encoding scheme; Syntax encoding scheme.
VRA 3.0	A set of elements or properties, called <i>categories</i> , along with qualifiers, specifications for controlled data values, and mappings to Dublin Core. VRA 3.0 is not inherently XML-based, although it may be expressed in XML. See <i>also</i> Visual Resources Association Core Categories; VRA 4.0.
VRA 4.0	A set of elements or properties, called <i>categories</i> , and its associated XML schema, making use of hierarchically-nested XML elements and subelements as well as element attributes. See <i>also</i> Visual Resources Association Core Categories; VRA 3.0.
W3C	See World Wide Web Consortium.
Web Ontology Language (OWL)	“The Web Ontology Language (OWL) is a family of knowledge representation languages for authoring ontologies. The languages are characterized by formal semantics and RDF/XML-based serializations for the Semantic Web. OWL is endorsed by the World Wide Web Consortium (W3C) and has attracted academic, medical and commercial interest.” Source: http://en.wikipedia.org/wiki/Web_Ontology_Language . See <i>also</i> Ontology.
World Wide Web Consortium (W3C)	The primary international standards organization for the World Wide Web. “The W3C mission is to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure the long-term growth of the Web.” Source: http://www.w3.org/Consortium/mission.html .
XML (EXtensible Markup Language)	“A set of rules for encoding documents in machine-readable form. It is defined in the XML 1.0 Specification produced by the W3C, and several other related specifications, all gratis open standards.” Source: http://en.wikipedia.org/wiki/XML .
XML schema	A machine-readable XML document or file that formally defines a specific XML language by specifying the elements, attributes, and other characteristics of that language. In metadata terms, an XML Schema defines a specific metadata element set such as MODS or VRA 4.0. See <i>also</i> Document Type Definition (DTD).

Other Useful Metadata Glossaries

Baca, Murtha, ed. 2008. *Introduction to Metadata*, 2nd ed., v. 3.0. Los Angeles: Getty Research Institute. http://www.getty.edu/research/publications/electronic_publications/intrometadata/glossary.html.

NISO (National Information Standards Organization). 2004. “Understanding Metadata.” Bethesda, MD: NISO Press. <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.

Woodley, Mary S. 2005. “DCMI Glossary.” <http://dublincore.org/documents/usageguide/glossary.shtml>. (This Glossary is under revision and will be replaced with a new version later in 2011.)

Zeng, Marcia Lei, and Jian Qin. 2008. *Metadata*. New York: Neal-Schuman.