
GLOSSARY

For additional terms not defined here, see:

Pearce-Moses, R. (2012). *Society of American Archivists: Glossary of archival and records terminology*. Retrieved from <http://www2.archivists.org/glossary>

ADMINISTRATIVE METADATA

metadata used for the administration of a collection or digital object. May include acquisitions information, provenance, use restrictions, access restrictions, copyright ownership, citation information, and general processing information.

API (APPLICATION PROGRAMMING INTERFACE)

a set of functions or routines that can be used to allow communication between applications. In a Web environment, published APIs allow the creation of “mashups” incorporating elements of one Web service into another unrelated site.

ARCHIVAL OBJECT

a digital object intended for long-term preservation. Current best practices for archival objects with respect to still images call for uncompressed TIFF (Tagged Image/Interchange File Format) or another uncompressed, nonproprietary format, such as JPEG2000. Sometimes called the “digital master” or “master image.”

COLLECTION

a group of resources brought together by shared content, provenance, ownership, subject, etc. Digital objects can simultaneously belong to multiple collections.

COMPRESSION

the re-encoding of data to make it smaller. May be “lossy,” discarding information from the original file, or “lossless,” which preserves the original information.

CONTENT STANDARD

community-specific guidelines for metadata creation. Content standards may address the type of information included in specific metadata elements as well as the format and punctuation of those elements. Examples include the AACR2r (Anglo-American Cataloging Rules, 2nd edition, revised), RDA (Resource Description and Access), and DACS (Describing Archives: A Content Standard).

CONTROLLED VOCABULARY

a list of terms that may be used to describe resources in order to ensure consistency and improve information retrieval. Examples of controlled vocabularies include Library of Congress Subject Headings and Name Authorities, the Getty Art and Architecture Thesaurus, and the Nomenclature for Museum Cataloging. Controlled vocabularies can also be developed and maintained at a local level.

DERIVATIVE IMAGE

an image that has been created from another image. Usually involves a loss of information. Techniques to create derivative images include sampling to a lower resolution, using lossy compression techniques, or altering an image with image manipulation software during image processing.

DESCRIPTIVE METADATA

metadata primarily intended to serve the purposes of discovery, identification, and selection. Includes elements similar to traditional bibliographic information, such as author/creator, title, subject, and date, but may also include descriptions, transcriptions, coordinates, and user-generated elements such as tags.

DIGITAL ASSET MANAGEMENT SYSTEM (DAMS)

a software system, frequently in the form of a database, used to organize and process digital assets like images, documents, and audio and video files. See attached list for common systems.

DIGITAL OBJECT

an item or resource in digital format. May be the result of digitization or may be “born digital.”

DIGITIZATION

the conversion from printed paper, film, or other media formats to an electronic format where an object is represented as either black and white dots, color or grayscale pixels, or 1s and 0s.

DPI (DOTS PER INCH)

a measure of print resolution, i.e., how many individual dots are produced within one linear inch.

DUBLIN CORE

set of metadata elements created to facilitate the discovery of electronic resources, maintained by the Dublin Core Metadata Initiative (DCMI). The Dublin Core Element Set consists of the original 15 core elements developed for use with HTML and has been defined in ANSI/NISO Standard Z39.85-2012 and ISO standard 15836:2009. Additional metadata terms are defined at <http://dublincore.org/documents/dcmi-terms/>

ENCODED ARCHIVAL DESCRIPTION (EAD)

a standard used to mark up (encode) archival finding aids in a manner that reflects the hierarchical nature of archival collections and that provides a structure for describing the whole of a collection, as well as its components

FILE FORMAT

the underlying structure of a digital file, often characterized by its extension (e.g., .pdf, .jpg, etc.). Different types of files, such as text, images, audio, etc. require different formats. Some formats use compression to minimize file size, while others preserve all the information in the original.

GRAYSCALE

an image mode that captures 256 shades of gray (8-bit). Every pixel of a grayscale image has a brightness value ranging from 0 (black) to 255 (white). Grayscale mode is commonly used in scanning black and white photographs.

HARVESTER

a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from repositories.

INSTITUTIONAL REPOSITORY

software and associated tools used to capture, structure, provide access to, and preserve digital materials produced by an organization or community.

INTEROPERABILITY

the ability of multiple systems, using different hardware and software platforms, data structures, and interfaces, to communicate, exchange, and share data. See also OAI-PMH and Z39.50.

JPG, JPEG

Joint Photographic Experts Group, or an image file format commonly used for photographs, based on a standard developed by that group. JPEG2000 is a newer standard based on different technologies.

MARC

Machine-Readable Cataloging. Data structure standard used in Integrated Library Systems (ILS) for Online Public Access Catalogs (OPACs).

METADATA

structured information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource.

METADATA CROSSWALK

an authoritative mapping from the metadata elements of one scheme to the elements of another.

METADATA HARVESTING

a technique for extracting metadata from individual repositories and collecting it in a central catalog to facilitate search interoperability. See also OAI-PMH

METADATA SCHEME

a set of metadata elements and rules for their use that has been defined for a particular purpose. A metadata scheme only dictates what metadata elements are allowed, not what they contain, and so is necessary but not sufficient to ensure interoperability. Examples include MARC, MODS, Dublin Core, EAD, and many others.

MIGRATION

a digital preservation technique to preserve the integrity of digital files by transferring them across hardware and software configurations and subsequent generations of computer technology.

OAI-PMH

The Open Archives Initiative's Protocol for Metadata Harvesting, an application-independent interoperability framework based on metadata harvesting, maintained by the Open Archives Initiative.

OCR

Optical Character Recognition, a process that produces a page of text from an image file. Accuracy rates vary greatly depending on the quality of the image.

PERSISTENT URL (PURL)

a Uniform Resource Locator (URL) associated with a service that will redirect the URL to a resource, even if the resource's location changes.

PIXEL

short for picture element, the building block of digital images. Each pixel can represent a number of different shades or colors, depending on how much storage space is allocated for it.

PPI (PIXELS PER INCH)

resolution of a computer display, image scanner or digital camera, measured in terms of the number of pixels displayed per linear inch of display.

PRESERVATION METADATA

metadata primarily intended to help manage the process of ensuring the long-term preservation and usability of digital information resources.

QUALITY CONTROL

techniques used to ensure that high quality is maintained through the various stages of digitization.

REPOSITORY

a storage system for digital resources and/or metadata, such as a DAMS. In OAI-PMH, a network accessible server that can process OAI-PMH requests.

RESOLUTION

the number of pixels (in both height and width) making up an image. The more pixels, the higher the resolution; the higher the resolution, the greater the file size. Can be expressed as a ratio (640 x 480 pixels) or in terms of pixels per inch (ppi).

RIGHTS METADATA

metadata primarily intended to enable the management of rights related to information resources; a type of administrative metadata.

STRUCTURAL METADATA

metadata that describes the internal organization of a resource and its place in an external organization, including any relationships it has with other resources.

SURROGATE

a secondary object meant to substitute for the original, such as a metadata record used in place of the original physical object or a digital image used in place of a three-dimensional object.

TECHNICAL METADATA

metadata primarily intended to document the creation and characteristics of digital files.

TIF, TIFF

Tagged Image File Format, an industry standard image file format. Uncompressed, originally developed for desktop publishing. 1 to 64 bit depth, used mostly for high quality imaging and archival storage.

UNIQUE IDENTIFIER

a name or number that unambiguously identifies an item within a repository or database. Unique identifiers can be self-describing (e.g., ISBNs) or assigned.

URI

Uniform Resource Indicator, a formatted string that serves as an identifier for a resource, typically on the Internet. The most common type of URL is the URL (Uniform Resource Locator) used in HTML.

COMMON DIGITAL ASSET MANAGEMENT SYSTEMS

ArchivalWare

Web-based, full-text search and retrieval system for digital libraries. Proprietary. Developed and sold by Progressive Technology Federal Systems, Inc. (PTFS). Previously licensed by SirsiDynix and sold under the names SirsiDynix Digital Library and Horizon Digital Library.

URL: <http://www.archivalware.net/>

ArchivesSpace

Web application to manage descriptive information for archives, manuscripts, and digital objects designed to incorporate the best features of Archivist's Toolkit (AT) and Archon. Open source; currently in alpha release (version 1.0 originally scheduled for delivery in September 2013). Primary institutional partners are the New York University Libraries, UC San Diego Libraries, and the University of Illinois Urbana-Champaign Libraries. Management services provided by LYRASIS. Hosting options are expected to be provided by LYRASIS.

URL: <http://www.archivesspace.org/>

Archon

Archon is a unified platform for archival description and access. It provides both a way to record descriptive information about collections and digital objects and a means to view, search, and browse that information in a fully-functional public web site. Open source. Developed by the University of Illinois at Urbana-Champaign Library. As of 2013, scheduled to be deprecated in favor of ArchivesSpace.

URL: <http://www.archon.org/>

Blacklight

An open source Ruby on Rails gem that provides a discovery interface for any Solr index. Blacklight provides a default user interface which is customizable via the standard Rails (templating) mechanisms. Blacklight accommodates heterogeneous data, allowing different information displays for different types of objects. The University of Virginia, Stanford University, Johns Hopkins University, and WGBH are the principal contributors to the code base. Active user community.

URL: <http://projectblacklight.org/>

CONTENTdm

High-performance storage and retrieval software for multimedia collections. Proprietary. Developed and sold by OCLC. Active user community.

URL: <http://www.contentdm.org/>

DigitalCommons

Hosted institutional repository software for universities, colleges, law schools, and research centers. Proprietary. Developed and sold by bepress.

URL: <http://digitalcommons.bepress.com/>

DSpace

Turnkey institutional repository software. Open source. Available as a local installation or, as of 2013, as a hosted service (DSpaceDirect) supported and sold by DuraSpace. Active user community.

URL: <http://www.dspace.org/>

URL: <http://www.dspaceirect.org/>

Fedora Commons

Framework for building institutional repositories. Open source. Available as a local installation; interface options include Hydra and Islandora. Supported by DuraSpace. Active re-development in progress as part of the Fedora Futures project. Active user community.

URL: <http://www.fedora-commons.org/>

Hydra

Components that let institutions deploy robust and durable digital repositories (the body) supporting multiple “heads”: fully-featured digital asset management applications and tailored workflows. Its principle platforms are the Fedora Commons repository software, Solr, Ruby on Rails and Blacklight. Available as a local installation. Open Source. Key partners include DuraSpace, Stanford University, the University of Hull, the University of Virginia, and MediaShelf.

URL: <http://projecthydra.org/>

Islandora

Institutional repository system based on Fedora Commons, Drupal, Solr, and additional applications. Open source. Available as a local installation. Originally developed by the University of Prince Edward Island's Robertson Library. Active user community.

URL: <http://islandora.ca/>

MWeb

Family of products for creating online catalogs of data, images, and media, with searching optimized for complex cultural data. Proprietary. Developed and sold by Selago Design, Inc. Licensed to Museum Software, Inc. and re-sold as PastPerfect Online.

URL: <http://www.systemsplanung.com/mweb/default.asp>

OpenJournalSystems (OJS)

Journal management and publishing system developed by the Public Knowledge Project. Open source. Available as a local installation; hosted options available from PKP Publishing Services. Key participating institutions include Simon Fraser University, the School of Education at Stanford University, the University of British Columbia, the University of Pittsburgh, the Ontario Council of University Libraries, and the California Digital Library.

URL: <http://pkp.sfu.ca/?q=ojs>

Omeka

Web-publishing platform for the display of library, museum, archives, and scholarly collections and exhibitions. Open source. Available as a local installation; hosted options available from the Corporation for Digital Scholarship (omeka.net). Developed by the Roy Rosenzweig Center for History and New Media, George Mason University. Active user community.

URL: <http://omeka.org/>

PastPerfect

Collection management and cataloging software primarily developed for museums. Proprietary. Includes four separate catalog options - Objects, Photos, Archives, Library. The PastPerfect Basic Program does not include the ability to attach images. The Multi-Media upgrade must be purchased separately. Developed and sold by PastPerfect Software, Inc. A partnership with the American Association for State and Local History allows the *Nomenclature 3.0* object classification system to be incorporated in the software, and AASLH members receive a discount on software licenses. Active user community.

URL: <http://www.museumsoftware.com/>

PastPerfect Online

Web publishing export tool that allows PastPerfect catalogs to be exported to hosted MWeb websites. Proprietary. Setup fee plus annual hosting fee.

URL: <http://www.museumsoftware.com/pponline.html>

PastPerfect Virtual Exhibit

Virtual Exhibit converts selected collection records and images from PastPerfect into HTML pages for posting online.

URL: <http://www.museumsoftware.com/ve5.html>

Plinkit

Template-based web site creation toolkit based on Plone content management system. Open source. Available as a local installation; hosted options are available from various state library collaborative. Online Photography Exhibits can be included.

URL: <http://www.plinkit.org>