

Collaboration in the Digital Age Glossary

compiled by the Institute of Museum and Library Services and Heritage Preservation

June 2008

authentication

A mechanism that attempts to establish the authenticity of digital materials at a particular point in time. For example, digital signatures.⁴

authenticity

The digital material is what it purports to be. In the case of electronic records, it refers to the trustworthiness of the electronic record as a record. In the case of "born digital" and digitized materials, it refers to the fact that whatever is being cited is the same as it was when it was first created unless the accompanying metadata indicates any changes. Confidence in the authenticity of digital materials over time is particularly crucial owing to the ease with which alterations can be made.⁴

born digital

Digital materials which are not intended to have an analog equivalent, either as the originating source or as a result of conversion to analog form.⁴

compression

The re-encoding of data to make it smaller. Most image file formats use compression because image files tend to be large and consume large amounts of disk space and transmission time over networks.⁶

controlled vocabulary

Formal limits on a vocabulary, useful for consistent use of vocabulary terms.⁶

content file

A file that is either born digitally or produced using various kinds of capture application software. Audio, image, text, and video are the basic kinds of content files.²

digital assets

A collection of computer files that contain intellectual content (images, texts, sounds, video) and/or descriptive metadata of the content and its digital format. They represent an investment for the depositor and an information resource for the researcher.²

digital materials

A broad term encompassing digital surrogates created as a result of converting analog materials to digital form (digitization), and "born digital" for which there has never been and is never intended to be an analog equivalent, and digital records.⁴

digital object

An abstraction that can refer to any type of information. The object may be simple or complex, ranging from values used in databases to graphics and sounds. Objects are not necessarily self-contained, for example, a graphics object may require an external piece of software to render the image. In addition to the data that makes up the fundamental content, the object often includes metadata that describes the resource in a manner that supports administration, access, or preservation.⁵

digital preservation

Combination of policies, strategies and actions to ensure the access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time.¹

digital provenance

A record of all migrations, transformations, or translations performed on a digital object from its original creation to the present time.²

digitization

The process of creating digital files by scanning, audio transfer, or otherwise converting non-digital materials.

documentation

The information provided by a creator and the repository, which provides enough information to establish provenance, history, and context and to enable its use by others.⁴

electronic records

Records created digitally in the day-to-day business of the organization and assigned formal status by the organization. They may include word processing documents, emails, databases, or intranet web pages.⁴

finding aid

A tool used to communicate the contents of an archival collection, the finding aid typically includes administrative information, contextual information, scope and content information, intellectual organization and physical location information for archival and manuscript materials.⁶

fixity

The quality that a digital file has not changed in any way between two points in time.

image capture

Using a scanner, digital camera, or other device to create a digital representation of an object.⁶

image file format

A standardized way of storing digital image data. Different file formats commonly use different methods of compression. Some image file formats include JPEG, JPEG2000, TIFF, PNG, BMP and GIF.

ingest

The process by which a digital object or metadata package is absorbed by a different system than the one that produced it.²

interoperability

The ability of multiple systems, using different hardware and software platforms, data structures, and interfaces, to communicate, exchange, and share data.⁶

JPG, JPEG

Joint Photographic Experts Group. An 8-24 bit image file format that is best suited for photographs. It supports "lossiness," which means that it will throw away some detail in order to achieve better compression. It has variable amount of compression to vary quality and file size. It does not work well for text. Widely used as a delivery format.⁶

JPEG2000

An image file format based on an image compression standard that supports both lossy and lossless compression. While not yet supported as widely as JPEG or TIFF, JPEG2000 is becoming more popular both as a delivery format and as an archival format.

life-cycle management

Proactive approach to the management of digital materials that recognizes, whatever their form or function, digital materials need to be actively managed at each stage of its life and that each stage of a digital materials life affects another stage so digital preservation activities should begin early as practicable. This is different from traditional preservation, where management is largely passive until detailed conservation work is required, typically, many years after creation and rarely, if ever, involving the creator.⁴

LOCKSS (Lots of Copies Keep Stuff Safe)

Technology designed to preserve copies of electronic publications by ensuring that multiple copies are stored at different organizations. See <http://lockss.stanford.edu/>.

lossless

Characteristic of compression schemes where information is not permanently lost in compression.

lossy

Characteristic of compression schemes where information is permanently lost in compression.

metadata

Information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource.⁶

administrative metadata

Metadata primarily intended to facilitate the management of resources.⁶

descriptive metadata

Metadata primarily intended to serve the purposes of discovery, identification, and selection.⁶

metadata harvesting

A technique for extracting metadata from individual repositories and collecting it in a central catalog to facilitate search interoperability.⁶

metadata scheme

A set of metadata elements and rules for their use that has been defined for a particular purpose.⁶

preservation metadata

Metadata primarily intended to help manage the process of ensuring the long-term preservation and usability of digital information resources.⁶

rights metadata

Metadata primarily intended to enable the management of rights related to information resources; a type of administrative metadata.⁶

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,⁶ e.g., the sequence of pages for a group of images of a diary or of detailed images of a larger image.²

technical metadata

Metadata primarily intended to document the creation and characteristics of digital files.⁶

migration

A means of overcoming technological obsolescence by transferring digital resources from one hardware/software generation to the next. The purpose of migration is to preserve the intellectual content of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology. Migration differs from the refreshing of storage media in that it is not always possible to make an exact digital copy or replicate original features and appearance and still maintain the compatibility of the resource with the new generation of technology.⁴

persistent identifier

A unique code that identifies a digital object within the central repository.

OAI (Open Archives Initiative)

An organization that has developed interoperability standards to facilitate the efficient dissemination of online content. For more information about its two major projects, the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and the Open Archives Initiative Object Reuse and Exchange (OAI-ORE), see <http://www.openarchives.org/>.⁵

OAIS (Open Archival Information System)

An archival method based on the Reference Model for an Open Archival Information System (OAIS) (<http://public.ccsds.org/publications/archive/650x0b1.pdf>). An archive that uses this method consists of an organization of people and systems that has accepted the responsibility for information deemed to need long term preservation and make it available for a designated community.

PDF

Portable Document Format, 4-64 bit depth. Uncompressed. Used mainly to image documents for delivery. Need plug-in or Adobe application to view. Adobe's Portable Document Format, the term Adobe uses to describe Acrobat files.⁶

quality control

Techniques used to ensure that high quality is maintained through the various stages of digitization.⁶

reformatting

Copying information content from one storage medium to a different storage medium (media reformatting) or converting from one file format to a different file format (file reformatting).⁴

refreshing

Copying information content from one storage media to the same storage media.⁴

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 its clarity and definition and the greater the file size. Can be expressed as a ratio (640 x 480 pixels) or in terms of dots per inch (dpi).⁶

scanning

See digitization.

server

Host computer for web pages, applications, or services.⁶

surrogate

A secondary object meant to substitute for the original, such as a photograph of an artwork used in place of the artwork.⁶

tag

A keyword used to describe a digital object.

thesaurus

A controlled vocabulary with syndectic structure in which all allowable terms are given and relationships between terms are shown.⁶

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. Generally non-compressed and high quality, including large file sizes. Most TIFF readers only read a maximum of 24-bit color.⁶

trusted repository

An entity whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. Trusted digital repositories may take different forms: some institutions may choose to build local repositories while others may choose to manage the logical and intellectual aspects of a repository while contracting with a third-party provider for its storage and maintenance.⁷

validation

A process to check one or more aspects of a submission for schema errors, file format problems, and ingest parameter inconsistencies that might affect its suitability for preservation. Results of a validation may include any combination of structural analysis information, warning messages, or fatal errors that prevent an object from being ingested.²

watermark (or digital watermark)

The process of embedding information in a digital image, audio, or video file. In visible watermarking, the information is typically text or a logo, which identifies the owner of the file. In invisible watermarking, information is added as digital data but it cannot be perceived as such, a technique often used in copyright protection systems.

References

- ¹ American Library Association, Association for Library Collections & Technical Services, Preservation and Reformatting Section, *Definitions of Digital Preservation*, 2008. <http://www.ala.org/ala/alcts/newslinks/digipres/index.cfm>
- ² Colvin, Jennifer, *Glossary*, California Digital Library, updated as of May 15, 2006. <http://www.cdlib.org/inside/diglib/glossary/>
- ³ Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS)*, 2002. <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- ⁴ Digital Preservation Coalition, Maggie Jones, and Neil Beagrie. "Introduction: Definitions and Concepts" from *Preservation Management of Digital Materials: A Handbook*, 2001 (maintained online). <http://www.dpconline.org/text/intro/definitions.html>
- ⁵ Moses, Richard Pearce. *A Glossary of Archival and Records Terminology*. Society of American Archivists, 2005. <http://www.archivists.org/glossary/index.asp>
- ⁶ North Carolina Exploring Cultural Heritage Online, *Glossary*, 2007. <http://www.ncecho.org/guide/glossary.htm>
- ⁷ RLG, *Trusted Digital Repositories: Attributes and Responsibilities*. RLG-OCLC, 2002. <http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>