Digital Archives Terminology[1]

administrative metadata – Administrative metadata elements describe information needed to manage or use the object. For example, administrative metadata for an image may include: creation date, copyright permissions, required software, provenance (history), and file integrity checks.

analog – Analog is used to described physical media. Examples of analog formats include photographs and films made with light-sensitive media, NTSC and PAL video recordings, and phonograph records and older magnetic sound recordings on tape.

audit trail – Information in records that track a transaction from beginning to end, making it possible to review whether it was done according to relevant policies and standards. An audit trail typically includes the time of transaction, the parties involved, and actions taken.

bit rot/bit loss – Corruption of the bits in a digital item during transmission or storage. Bit rot may render an object unusable or may just not be visible.

born digital – Refers to materials that originate in a digital form. This is in contrast to digital reformatting, through which analog materials become digital.

checksum – A mathematical value used in a simple error-detection method to verify data.

cloud computing – The practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server. Typically used to describe third-party tools, but internal organizational servers can also be hosted on the "cloud."

collection – A unifying group of materials. In the digital environment, items may belong to one or more collections.

controlled vocabulary – A limited set of terms and phrases for selected metadata fields, used in searching and browsing.

descriptive metadata – Information that refers to the intellectual content of an item or a collection. Descriptive metadata allows users to locate, distinguish, and select materials on the basis of the material's subjects or 'aboutness.' Descriptive metadata elements consist of information about the content and context of an object. For example, descriptive metadata for an image may include: title, creator, subject (tags), and description (abstract).

digital file – An individual electronic file such as a document or image. This file definition differs from the tradition use of a "file" in archives to represent a group of documents.

---

[1] Selected definitions adapted from: Pearce-Moses. (2005). "A Glossary of Archival Records and Terminology." The Society of American Archivists, http://www.archivists.org/glossary/ and *Wikipedia*, http://wikipedia.org

electronic record – An official record of an organization captured through electronic means, and which may or may not have a paper record to back it up.

emulation – The use of one system to reproduce the functions and results of another system. Emulation is the replicating of functionality of an obsolete system.

folder (directory) – A sub-division of a computer's hard disk into which you put files.

item – In digital collections an item may represent and individual file, such as a single image, or a multiple page document composed of multiple files.

lossless – Undegraded; retaining all information.

lossy – Degraded by a loss of information.

migration – The process of moving data from one information system or storage medium to another to ensure continued access to the information as the system or medium becomes obsolete or degrades over time.

OCR – Optical Character Recognition. OCR "reads" an image to create a machine-readable text of the text on the image.

obsolescence – Obsolescence is the state of being which occurs when an object, service or practice is no longer wanted even though it may still be in good working order.

persistent identifiers – Identifiers assigned to an object that will continue to point to the object if the object is moved or changed. Many digital collections have moved from having separate persistent identifiers and URLs to having one persistent identifier that also functions as a URL.

pixel – A pixel contains information about the brightness and, for color images, the hue of the element. A pixel may be represented by a single bit, restricting the image to bitonal values. In monochromatic images, a pixel is at least a byte, capable of representing 256 shades. For color images, a one-byte pixel can represent only 256 colors; a three-byte pixel can represent more than 16 million colors. The height and width of an image are often measured in terms of pixels, and the resolution may be measured in terms of pixels per inch or centimeter.

preservation metadata – Information about an object used to protect the object from harm, injury, deterioration, or destruction. Preservation metadata may be used to store technical information supporting preservation decisions and actions; document preservation actions taken, such as migration or emulation policies; record the effects of preservation strategies; ensure the authenticity of digital resources over time; [and] note information about collection management and the management of rights.

refreshing – Refreshing is the transfer of data between two types of the same storage medium so there are no bitrate changes or alteration of data.

replication – Creating duplicate copies of data on one or more systems is called *replication*. Data that exists as a single copy in only one location is highly vulnerable to software or hardware failure, intentional or accidental alteration, and environmental catastrophes like fire, flooding, etc. Digital data is more likely to survive if it is replicated in several locations. Replicated data may introduce difficulties in refreshing, migration, versioning, and access control since the data is located in multiple places.

resolution – The number of dots or pixels per inch (sometimes per centimeter or millimeter) used to create the screen image.

schema – A formal description of a data or metadata structure.

technical metadata – Technical metadata elements describe the format, process, and inter-relatedness of objects. For example, technical/structural metadata for an image may include: camera, aperture, exposure, file format, and set (if in a series).