Digital Readiness Glossary

Term	Definition
3-2-1 Rule	The 3-2-1 rule informs digital preservation and storage strategies. Maintain three copies of your digital files on two different storage media with at least one copy stored off site. See: Storage Diversification, Digital Storage
Access	In archives, access refers to the ability to locate and retrieve archival information for use within applicable restrictions.
Access Copy	A copy made from a digital object that is intended for use, such as online display or transmission over email
Accessibility	Digital accessibility is the ability of a website, mobile application or electronic document to be easily navigated and understood by a wide range of users, including those users who have visual, auditory, motor or cognitive disabilities.
Administrative Metadata	Administrative metadata is information needed to help manage the digital object, such as copyright and preservation information. See: Metadata.
Analog	Analog refers to information that exists in nondigital format such as printed or manuscript text, audio tapes or films, photographs or other graphics, or 3-D objects. Digitization is the conversion of analog information into digital information. Analog items in the archive might also be known as physical content.
Appraisal	In an archival context, appraisal is the process of determining whether records and other materials have permanent (archival) value.
Archival Copy	An archival copy in digital collections refers to digital content, targeted for preservation, that is considered the archival version of the intellectual content of a digital resource. Archival copies/preservation copies generally do not undergo significant processing or editing, and are often used to make other copies including reproduction and access copies. See: Master Copy, Preservation File.
Audit	An audit is an independent review and examination of records and activities to test for compliance with established policies or standards, often with recommendations for changes in controls or procedures.
Backup	A backup copy is an additional copy of a digital asset made to protect against loss due to unintended destruction or corruption of the primary set of digital assets.
Best Practices	Best practices are procedures and guidelines that are widely accepted because experience and research has demonstrated that they are optimal and efficient means to produce a desired result.
Bit	A bit is the smallest unit of information that a computer can work with. Each bit is either a "1" or a "0".
Bit Preservation	Bit-level preservation is the basic level of preservation of a digital resource (literally, preservation of the bits forming a digital resource). Bit-level preservation may include maintaining onsite and offsite backup copies, virus checking, fixity checking, and periodic refreshment to new storage media.
Bit-depth	Bit depth is determined by the number of bits used to define each pixel. The greater the bit depth, the greater the number of tones (grayscale or color) that can be represented. Digital images may be produced in black and white (bitonal), grayscale, or color.
Born-digital	Born-digital content has never had an analog form. Born-digital materials differ from analog documents, films, sound recordings and photographs that were digitized; that is, scanned or converted to a digital format.
Byte	A byte is a unit of digital information and measure of data volume, normally equivalent to eight bits. Bytes are the smallest operable units of storage in computer technology. Kilobyte (KB) = 1,000 bytes Megabyte (MB) = 1,000 kilobytes Gigabyte (GB) = 1,000 megabytes Terabyte (TB) = 1,000 gigabytes

Calibration (for scanners)	Calibration refers to aligning a scanner's color profile with its attached computer's color profile; a process that uses a color target. Calibration ensures true capture of the original colors in a digital format See: Color Target
Checksum	A checksum is a unique numerical signature derived from a file. Checksums are used in fixity checking in order to compare copies. See: Fixity
Cloud-based storage	Cloud storage is a way to save data securely online so that it can be accessed anytime from any location and easily shared with those who are granted permission. Cloud storage also offers a way to back up data to facilitate recovery off-site. Cloud storage services include Google Drive, Dropbox, Box, etc.
Collection	A general term to describe a body of records, and may include documents, photographs, audio/visual material, maps, etc., in both physical and electronic forms.
Collection Development Policy	Guidelines outlining the scope and selection of materials that support a repository's mission. Generally, a collecting policy defines the scope of existing collections and also describes processes such as deselection, retention, preservation, and storage. It provides guidance for archives staff, organizations and individuals interested in donating, and other collecting repositories.
Collection Inventory	A collection inventory includes, at a minimum, a list of items in a collection or a list of collections maintained by an organization. See: Inventory
Collections Management System (CMS)	A Collections Management System (CMS), sometimes called a Collections Information System, is software used by the collections staff of a collecting institution or by individual private collectors and collecting hobbyists or enthusiasts. Collections Management Systems (CMSs) allow individuals or collecting institutions to organize, control, and manage their collections' objects by "tracking all information related to and about" those objects. May also be referred to as a Discovery Platform. See: CONTENTdm, Mukurtu, Omeka, PastPerfect, Preservica
Color Target	A color target is a type of measuring table that calculates the exact color recognition capability of a scanner and identifies the singularities of that scanner. The color target is a small card with a range of colors printed on it that the scanner can scan during the calibration process. See: Calibration.
Color/bi-tonal/gray scale	A bitonal (bi-tonal, or two bits) image is black and white. A grayscale image represented by multiple bits of information, typically ranging from 2 to 8 bits or more. A color image is typically represented by a bit depth ranging from 8 to 24 or higher.
Community of Practice	A community of practice is a way to learn by working together. As described by Etienne Wenger, Richard McDermott and William M. Snyder in their 2002 book Cultivating Communities of Practice, a Community of Practice (COP) is "a group of people who share a common concern, set of problems, or passion about a topic and deepen their knowledge and expertise in this area by interacting on an ongoing basis."
Compression (compressed/uncompressed)	Compression is a process that reduces the amount of space necessary for data to be stored or transmitted. Compression alters digital image quality.
Content	Content refers to the intellectual substance of a document, including text, data, symbols, numerals, images, and sound. For A/V material, the content is the data encoded in a recording. For a book or other publication, it is the text and accompanying illustrations. For a photograph, it is the image itself, not the medium the image is held on (e.g., paper, glass or plastic.) For a digital photograph, it is the image and embedded metadata. For multimedia, it is the digital files and embedded metadata, not the hard drive or disc it is stored on. See: Digital Content
Content Migration	Content migration is the process of transferring content between storage types, formats, or computer systems.
Content Refreshing	The act of copying digital content to a new physical carrier, typically of the same media type. This is done to prevent the loss of content due to media degradation. See: Content Migration
Content Statement	A content statement might also be known as a harmful content statement. It is a brief introduction to materials that may be traumatic, triggering, hurtful or harmful to an unaware patron.

Controlled Vocabulary	A controlled vocabulary is a standardized, pre-determined list of terms, such as the Library of Congress Subject Headings.
Conversion	Conversion usually refers to some form of analog-to-digital conversion, such as digitizing VHS tapes or film reels, including scanning paper documents to create digital images or rekeying paper text into a computer. Conversion is more than copying files. It involves a change in media internal structure, such as from diskette to tape, from one version of an application to a later version, or from one application to another.
Copyright	Copyright refers to the legal rights protecting the interests of creators or their assignees by granting them control over the reproduction, publication, adaptation, exhibition, or performance of their works in fixed media.
Creative Commons License	A Creative Commons license is a type of license, built on copyright, that provides a standardized way for creators to give others the right to share and use their work.
Crowd-sourced projects	Crowdsourcing uses a large number of people to complete a specific task. Crowdsourcing in archives and special collections can take the form of transcribing handwritten documents, indexing genealogical records, identifying people and places in photos, correcting optical character recognition (OCR) errors in digitized newspaper collections, tagging or captioning historical images, adding pictorial content to maps, transcribing oral histories, and much more.
Cultural Heritage (organization, collection)	Cultural heritage refers to the legacy of physical artifacts and intangible attributes of a group or society that is inherited from past generations. Not all legacies of past generations are "heritage", rather heritage is a product of selection by society. Cultural heritage includes tangible culture (such as buildings, monuments, landscapes, books, works of art, and artifacts), intangible culture (such as folklore, traditions, language, and knowledge), and natural heritage (including culturally significant landscapes, and biodiversity).
Cultural Property Rights	The concept that a society, especially that of indigenous peoples, has the authority to control the use of its traditional heritage. Cultural property rights are roughly analogous to copyright, but the rights are held by a community rather than an individual and the property protected was received by transmission through generations rather than being consciously created. Cultural property rights have not been generally established or codified by statute in the United States, although the Native American Graves Preservation and Repatriation Act (NAGPRA) may be seen as recognizing those rights. Other countries, notably Australia, have begun to codify cultural property rights. See: NAGPRA
Dark Archive	A dark archive is a repository that stores archival resources for future use but is accessible only to its custodian. A dark archive does not grant public access and only preserves the information it contains. The information can be released for viewing depending on its donor and organizational restrictions, at which time it is no longer considered "dark."
Data Dictionary	A data dictionary is a collection of names, definitions, and attributes about data elements that are being used or captured in a database, information system, or part of a research project. It describes the meanings and purposes of data elements within the context of a project, and provides guidance on interpretation, accepted meanings and representation.
Data Integrity	Refers to the trustworthiness of system resources over their entire life cycle
DCMI (Dublin Core)	The Dublin Core, also known as the Dublin Core Metadata Element Set, is a set of core elements for describing resources. DCMI stands for "Dublin Core Metadata Initiative." Dublin Core is a widely used metadata standard.
De-duplication	De-duplication refers to techniques for eliminating duplicate copies of repeating data.
Deed of gift	The deed of gift is a form that confirms a legal relationship between the donor and repository that is based on a clearly articulated and common understanding.

Description	Description is the process of analyzing, organizing, and recording details about the formal elements of a record or collection of records, such as creator, title, dates, extent, and contents, to facilitate the work's identification, management, and understanding. Description can be done at the collection level or the item level.
Descriptive Metadata	Descriptive metadata is information used to search for and locate an object such as title, author, subjects, keywords, and publisher. Descriptive metadata allows users to locate, distinguish, and select materials on the basis of the material's subjects or 'aboutness.' It is distinguished from information about the form of the material, or its administration. See: Metadata
Digital	Digital records, archives, media, projects, activities, responsibilities, etc. are those involving or making use of computer devices, data, or media.
Digital Asset	A digital asset is a single computer file, or group of computer files, the content of which is valuable to your organization.
Digital Content (Digital Materials)	Digital content refers to any item created, published or distributed in a digital form, including, but not limited to, text, data, sound recordings, photographs and images, motion pictures and software. Born-digital content has never had an analog form, and differs from analog documents, movies and photographs that were digitized - that is, scanned or converted to a digital format. This term is used interchangeably with digital materials. See: Content
Digital Curation	Digital curation is the act of maintaining and adding value to a body of digital information for future and current use; specifically, the active management and appraisal of data over the entire life cycle.
Digital Millennium Copyright Act (DMCA)	The Digital Millennium Copyright Act (DMCA) addresses important parts of the relationship between copyright and the internet. It established a notice-and-takedown system, among other provisions.
Digital Object	A digital object is an item, either born digital or analog, which has been targeted for digital preservation and its accompanying metadata
Digital Obsolescence	Digital obsolescence refers to a situation where a digital resource is no longer readable because of an archaic format: the physical media, the reader (required to read the media), the hardware, or the software that runs on it is no longer available. See: Obsolescence
Digital Preservation	The term digital preservation encompasses all of the activities, policies, strategies and actions required to ensure that the digital content designated for long-term preservation is maintained in usable formats, for as long as access to that content is needed or desired, and can be made available in meaningful ways to current and future users, for as long as necessary regardless of the challenges of media failure and technological change. Digital preservation goals include ensuring enduring usability, authenticity, discoverability, and accessibility of content over the very long term.
Digital Preservation Plan	A Digital Preservation Plan describes actionable steps to be taken to preserve digital resources within an organization (the Action Plan), and documents how this Action Plan was chosen.
Digital Preservation Policy	A digital preservation policy is the mandate for an archive to support the preservation of digital records through a structured and managed digital preservation strategy.
Digital Stewardship	Digital stewardship encompasses all activities related to the care and management of digital objects over time. Proper digital stewardship addresses all phases of the digital object lifecycle: from digital asset conception, creation, appraisal, description, and preservation, to accessibility, reuse, and beyond.
Digital Storage	Digital storage refers to a digital method of keeping data, electronic documents, images, etc. in a digital storage location, usually a hard drive or in cloud-based storage. Archival digital storage is not the same as a backup — archival storage keeps content accessible for future users and computers, while backups keep your computer files working safely and securely. See: 3-2-1 Rule, Redundancy

Digitization	Digitization is the process of creating digital copies or "surrogates" by scanning or otherwise converting analog materials. Digitization is the conversion of analog information into digital information. See: Analog
Disaster Threat (level, area)	Disaster risk zones show the likelihood of various natural disasters affecting a particular geographic area. It is advisable to have digital storage options in various disaster risk zones different from your own; for instance, if your area is prone to earthquakes, choose cloud-based backups in an area not prone to earthquakes (and ideally not prone to natural disasters at all).
Diversity	Diversity refers to: 1) Understanding and valuing characteristics of those who demonstrate a wide range of characteristics. 2) Audiences whose members exhibit a wide variety of characteristics. In both definitions those characteristics include different ethnic and racial backgrounds, age, physical and cognitive abilities, family status, sexual orientation, socioeconomic status, religious and spiritual values, and geographic location.
DPI (dots per inch)	Dots per inch, or DPI, refers to the number of printed dots contained within one inch of an image printed by a printer. It is a measure of the resolution of a printed document or digital scan -- a higher DPI is a sharper image; a lower DPI is a fuzzier image.
Electronic Records	Electronic records are those created digitally in the day-to-day business of an organization, such as word processing documents, emails, databases, or intranet web pages.
FADGI	FADGI stands for the Federal Agencies Digital Guidelines Initiative. Their Technical Guidelines for Digitizing Cultural Heritage Materials are best practices for cultural heritage imaging for still images.
Fair Use	In its most general sense, a fair use is any copying of copyrighted material done for a limited and "transformative" purpose, such as to comment upon, criticize, or parody a copyrighted work. Such uses can be done without permission from the copyright owner. In other words, fair use is a defense against a claim of copyright infringement. If your use qualifies as a fair use, then it would not be considered an infringement.
File Format	Digitally, a file format is a standard way that information is encoded for storage in a computer file. A file format is often indicated by a file name extension e.g. .tif, .pdf, .jpg.
File Naming Convention	A file naming convention is a set of rules used to create consistent names across a set of files.
Fixity	Fixity refers to the "unchangedness" of data, usually evidenced by identical and persistent checksums generated from the same file over time. Fixity refers to the stability of a digital object over time. See: Checksum
Format Migration	Format migration refers to a means of overcoming technological obsolescence by transferring digital resources from one hardware/software generation to the next.
Full (digital) Preservation	Full digital preservation is the use of format migration, emulation, digital forensics, and other strategies to ensure that the content of digital materials, rather than just the original bits and bytes, remain protected and accessible over time despite technology obsolescence and the need for refreshed storage media. See: Format migration, Emulation, Digital forensics
Gap Analysis	Gap analysis is the comparison of actual performance with potential or desired performance. In libraries, archives and museums, this can refer to gaps in collections, procedures, documentation or other work areas.
Hard Disk Drives	Hard disk drives are a form of magnetic media that have magnetic platters read by spinning arms.
Hard Drive (external, portable, SSD)	An external hard drive plugged into a computer port rather than installed inside a computer. Used for storage and backups. SSD=Solid State Drive
Harmful Content Statement	A content statement might also be known as a harmful content statement. It is a brief introduction to materials that may be traumatic, triggering, hurtful or harmful to an unaware patron. See: Content Statement

Inclusion Gaps	In archives, "inclusion gaps" refers to voices or materials that may not be in your collections but perhaps should be. For instance, Native Americans occupied your geographic location long before your current organization began collecting records; are Native American voices respectfully represented anywhere? Likewise, do women, people of color, people with disabilities, etc. have voices in your materials?
Intellectual Property Rights	Intellectual property rights are the rights given to persons over their literary or artistic works. They usually give the creator an exclusive right over the use of his/her creation for a certain period of time. Intellectual property rights are governed by copyright restrictions.
Inventory	In archives, an inventory refers to a lists of holdings: archival items, collections, photographs, or recordings. See: Collection Inventory
Legacy Media	Legacy media are carriers of digital information that are either obsolete or becoming obsolete soon. Files on legacy media should be given higher prioritization in digital preservation to prevent their permanent loss.
Legacy System	A legacy system is software or hardware that was built using methods that are outdated or obsolete.
LOCKSS	Digital preservation principle that Lots Of Copies Keep Stuff Safe.
Lossless format	Lossless formats are file formats that are stable and therefore compatible with long-term preservation efforts. In general, these formats have the following characteristics: openly documented; supported by a range of software platforms; widely adopted; lossless data compression or no compression; non-proprietary; and does not contain embedded files or embedded programs.
Machine-readable	Machine-readable means that the information is in a medium or format that requires a mechanical device to make it intelligible to humans. Machine-readable is commonly used to refer to digital computer data files.
Master Copy	Master copy is a term used to describe an original, unmodified analog or digital file. Due to the negative connotations of the term "master," archival copy or preservation copy are preferred terms. See: Archival Copy, Preservation Copy
Media Deterioration or Degradation	Deterioration or degradation is the breakdown of an analog object that holds digital objects, potentially causing the digital objects on the media to no longer be retrievable.
Metadata	Metadata is a Latin term meaning "information about information." In the digital realm, metadata is data that describes key information about digital objects (image files, text files, digital audio/video) and, when appropriate, the original objects they represent. Types of metadata include administrative, descriptive, structural, and technical.
Metadata Elements	Metadata elements are defined data points that are used to capture information about a resource. Some of these data points might include a title, an identifier, a creator name, or a date.
Metadata standard	A metadata standard is a system or a set of rules that ensure descriptive information is applied consistently across your items. DublinCore is a commonly-used metadata standard in digital archives.
Monitoring	Monitoring refers to logging or recording various aspects of digital storage configuration, including hardware, activity, and data integrity.
NAGPRA	The Native American Graves Protection and Repatriation Act passed in 1990 requires public and private institutions that have received federal funds to inventory Native American objects in their care, notify the appropriate cultural or tribal representatives, and return the objects if requested. See: Cultural Property Rights
OAI-PMH	The Open Archives Initiative Protocol for Metadata Harvesting is a protocol for harvesting (collecting) descriptive metadata records from a repository so that services can be built using metadata from many sources. Simply put, OAI-PMH is a way for repositories to structure and exchange information in the same formats.
OAIS	OAIS is an acronym that stands for Open Archival Information System. The system gives the digital preservation community a common language and outlook for talking about digital preservation.

Obsolescence	Format or technology obsolescence occurs when a piece of software or hardware is no longer in wide use or available at all. This causes it to be difficult or impossible to use any files that depend on this software or hardware. See: Digital Obsolescence
Open Source Software	Open source software is developed through public collaboration and distributed without charge. Because open source software is free, it is more likely to continue to be usable longer than paid software, which may discontinue supporting the programs at any time.
Optical media	Optical media refers to any data storage device or equipment that uses optical data storage and retrieval techniques to read and write data. It stores data digitally on a media device and uses a laser to read data from it. Optical media is also referred to as optical storage.
Organizational Mission	The organizational mission defines the purpose of the institution, in The organizational mission defines the purpose of the institution, including who it serves, how it provides those services, and what unique assets it uses to provide the service. The mission statement guides all activities of the institution.
Personally Identifiable Information (PII)	Personally Identifiable Information can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context.
Pixel	A pixel is an element in an array that forms an image, a tiny dot. It is a unit of measure used to describe the size or resolution of an image, i.e. pixels per inch.
PPI (Pixels Per Inch)	Image resolution is typically described in PPI, which refers to how many pixels are displayed per inch of an image. Higher resolutions mean that there more pixels per inch (PPI), resulting in more pixel information and creating a high-quality, crisp image. Images with lower resolutions have fewer pixels, and if those few pixels are too large (usually when an image is stretched), they can become visible.
PREMIS	PREMIS is an acronym that stands for Preservation Metadata: Implementation Strategies. PREMIS metadata structures and describes what sort of preservation actions have been done to a digital object.
Preservation File	A preservation copy refers to digital content targeted for preservation that is considered the archival version of the intellectual content of a digital resource. Preservation copies generally do not undergo significant processing or editing. Preservation copies are often used to make other copies including reproduction and distribution copies. See: Archival Copy, Access Copy, Master Copy
Preview	A preview file is a reduced size or length audio and/or visual representation of content, in the form of one or more images, text files, audio files, and/or moving image files.
Provenance	Provenance refers to the origin or source of something. In archives, it is information regarding the origins, custody, and ownership of an item or collection.
Public Domain	The term "public domain" refers to creative materials that are not protected by intellectual property laws such as copyright, trademark, or patent laws. The public owns these works, not an individual author or artist.
Quality Assurance / Quality Control	Quality control (QC) or Quality assurance (QA) is a process used to verify the quality, accuracy, and consistency of digital projects. A regular systematic QC process allows you to check for files that do not meet the standards of your project plan, and identify any problems.
Redundancy	Redundancy refers to the creation and retention of multiple near-identical copies of the same data, stored in different digital locations. See: Digital Storage
Replication	Replication is the automated copying of data from one primary storage location to another or several other storage locations. Replication is distinct from redundancy in that it dynamically updates the secondary storage locations.
Retention Schedule	Also known as "digital asset retention and disposition schedule." A document that identifies digital assets (typically by asset groups) and the date on which their disposal must take place. Often, the date will be recorded as "indefinitely," but this can be revised as an institution's inventory is reviewed.

Rights Management	Rights management refers to a system that identifies intellectual property rights relevant to particular works and that can provide individuals with access to those works on the basis of permissions to the individuals.
Rights Statements	A simple, standardized system of labels that clearly communicate the copyright and re-use status of digital objects to the public, which improves usability and access for users.
Scanner	An image scanner is a device used to scan images, printed text and objects into a digital format.
Storage	See: Digital Storage
Storage Diversification	Storage diversification, also known as the Geographically Dispersed Data Storage Model, keeps more than one copy of the object in more than one geographical region. See: 3-2-1 Rule
Sustainability	In this context, sustainability refers to activities to ensure your project can continue, for example: 1) Creating and documenting policies, procedures and workflows, 2) Creating training materials for future project staff, 3) Developing a digital preservation plan, 4) Building organizational or community support for the project, 5) Pursuing additional grants or more permanent funding to support the project work.
Technical Metadata	Technical metadata refers to information about aspects of the object related to its file format or the software used to create the file. This may include things like the scanning equipment used to create a digital object and the settings used to create or modify it. See: Metadata
Traditional Knowledge (TK) Labels	TK Labels are a tool for Indigenous communities to add existing local protocols for access and use to recorded cultural heritage that is digitally circulating outside community contexts.
Transcription	Transcription is the process of making a written copy of a recording or document. For audio or video recordings, a transcription is a written copy of the spoken material. For handwritten archival artifacts, a transcription is a typed, usually digital, version of the handwriting.
User	A user is an individual who uses the collections and services of a repository; a patron; a reader; a researcher; a searcher.
Virus Scan	A virus scan checks for malicious programs and macros on a computer or electronic device.
Web Archiving	Web archiving is the process of gathering up data that has been recorded on the World Wide Web, storing it, ensuring the data is preserved in an archive, and making the collected data available for future research.
Workflow	A workflow consists of the tasks, procedural steps, organizations or people, information and tools needed for each step in a process.
Working Copy	A working copy is a digital asset derived from an archival copy. Working copies can be modified to suit the needs of the project at hand.
XML	XML stands for eXtensible Markup Language. XML is one of the most common ways to represent metadata.

Definitions provided by or adapted from the following sources:

Cornell Digital Imaging Tutorial	Canada: Digital preservation recommendations for small museums
Digital Preservation Coalition	Curating Community Digital Collections Glossary
Indigitization Toolkit	Developing a Digital Preservation Policy
International Council on Archives	Digital Millennium Copyright Act (DCMA)
Investopedia: Cloud Storage	Digital Preservation Plan Framework for Cultural Heritage Institutions
LOCKSS	Digital Preservation Recommendations for Small Museums
Mukurtu	Digital Preservation Research (POWRR)
Network World	Digital Public Library of America (DPLA)
OAI-PMH	Federal Agencies Digital Guidelines Initiatives (FADGI)
Purdue University LibGuide	How to calibrate your scanner (lifewire.com)
SilverFast	Outagamie Waupaca Library System (OWLS) Digitization Workflow

Sustainable Heritage Network	Society of American Archivists (SAA) Dictionary of Archival Terms
Techopedia	University of California Merced Library, "What is a data dictionary?"
University of Michigan	University of California Santa Cruz Library
What is digital accessibility?	World Trade Organization
Wikipedia	Yale University