

What is digital preservation?

To understand what digital preservation is it is helpful to consider some characteristics of its targets, namely objects “expressed in digital form” (Thibodeau, 2002, p. 6):

- Materially, digital objects consist of bits inscribed on a physical medium, for example, the pits and lands on a CD. Thibodeau calls this part of digital objects the “physical object” (2002, p. 6).
- These inscriptions are not human-readable. However, their pattern can be interpreted by the correct software. Thus, the bits – recognized as a file¹ by the correct software – can be rendered in a specific software-hardware environment. This is the “logical object” according to Thibodeau (2002, p. 7).
- For human users, the meaning of digital objects does not derive from the pattern of the bits, but rests in the information content that the bits encode. Thibodeau refers to this aspect of digital information as the “conceptual object” (2002, p. 8). The conceptual object is accessible to us in the form of a rendering of the bits on a screen.² It can be rendered from one or several files (see Brown, 2013, p. 13).
- While there may be cases where the meaning of the rendered bits is self-evident, often this is not the case – for example, to understand the meaning of a digital photograph we may require information about when and where it was taken. Social science research data requires rich documentation, for example about the methods of data collection and preparation, to be comprehensible.
- We can easily manipulate digital objects and create an unlimited amount of copies.

All of these characteristics can pose “weak spots” when it comes to keeping digital objects alive: the physical inscriptions can be compromised by damage or obsolescence of the storage medium. There may be no suitable software available to interpret the bit sequence, because it was created with a software that has since become obsolete. The meaning of a digital object may no longer be evident to us because we do not have sufficient context information to understand its content. Or we are faced with different versions of an object, but cannot decide which is the authoritative one. All of these issues pose potential threats to accessing and comprehending digital objects, and it is therefore these issues that digital preservation must address. This is reflected in common definitions of the term digital preservation (see box). Both definitions mention maintaining access to digital objects for a defined time-span as the goal of digital preservation. This includes both technical access to the physical and logical object as well as intellectual access to the conceptual object: it is an important task of digital preservation to maintain the comprehensibility of digital objects, especially, by providing sufficient context information in the form of metadata.

¹ “File: A named and ordered sequence of bytes that is known to an operating system” (PREMIS Editorial Committee, 2012, p. 25)

² This rendering is referred to as the “performance” of a digital record by (Heslop, Davis, & Wilson, 2002).

Digital preservation definitions

“Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation [...] refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change” (Digital Preservation Coalition, 2008, “Introduction: Definition and Concepts”)³

“The process of maintaining a digital object for as long as required, in a form which is authentic and accessible to users” (Brown, 2013, p. xii).

The second important goal, mentioned by Brown, is maintaining authenticity of digital objects. Authenticity is defined as “[t]he quality of trustworthiness of a record – [...] the assurance that a record is what it purports to be and has demonstrably not been tampered with or otherwise corrupted” (Brown, 2013, p. xi). As mentioned above, digital objects can easily be manipulated (both accidentally or with malicious intention) or corrupted – e.g. by “data rot” caused by decaying storage media. What is more, in order to preserve digital objects, we will in all likelihood have to change them, for example, by migrating them to a different file format. Thus, digital preservation has to comprise routines and strategies to protect digital assets against unauthorized changes, to make any unauthorized or unintended changes detectable (e.g. with the help of check sums and fixity checks), and to document all authorized, required changes carried out as part of the digital preservation strategy in the metadata.

Recommended introductory resources

Brown, A. (2013). Practical digital preservation. A how-to guide for organizations of any size. London: facet publishing.

Digital Preservation Coalition. (2008). Preservation Management of Digital Materials: The Handbook. <http://www.dpconline.org/advice/preservationhandbook/introduction>

CESSDA Training also provides access to the learning materials of our introductory course “First Steps Towards Digital Preservation”. Please contact us for details.

Case studies that illustrate digital preservation challenges can be found in the “Atlas of Digital Damages” initiated by Barbara Siermann (<http://www.atlasofdigitaldamages.info/v1/>).

Institutional readiness and the different dimensions of digital preservation

To be successful, digital preservation requires more than a functioning technical environment. While it is important to have specific procedures for bitstream preservation (strategies for back-up, redundant storage, media refreshment/migration, error detection, etc.) or migration, these procedures have to be embedded in a suitable organizational framework.

This view is reflected in the metaphor of the three-legged stool. According to this concept, digital preservation can be compared to a stool standing on three legs: technology, organization, and resources (see <http://dpworkshop.org/dpm-eng/conclusion.html>). If one of the legs is missing, the stool cannot stand, just as digital preservation can only be successful if we address all three dimensions. Thus, in addition to technological tools and equipment and preservation strategies digital preservation require a suitable policy framework comprising, among others, a clear commitment to digital preservation by the parenting institution; policies for acquisition, preservation, and access; well-defined and documented procedures and strategies for all relevant

³ <http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts>



workflows. Moreover, digital preservation is a long-term effort which requires sufficient allocation of financial, human, and technological resources (McGovern, 2007).

How well an organization is prepared for the challenges posed by digital preservation can be expressed with the help of the concept of “institutional readiness”. A checklist for assessing institutional readiness is available on the pages of the Digital Preservation Management tutorial, developed by the Cornell University Library (http://dpworkshop.org/dpm-eng/eng_index.html). The checklist covers the following areas and can be used to gain an overview of potential gaps in the capacity to address digital preservation concerns:

- Organizational Infrastructure (mission, policies and procedures, authority, implementation)
- Technological Infrastructure (digital collections, archival storage, storage practice, obsolescence, depository, depository development, security, OAIS Compliance)
- Resources (sustainable funding, staffing, administrative structure) (see <http://dpworkshop.org/sites/default/files/readiness.pdf>)

Another model supporting institutions in an assessment of their “readiness” for digital preservation is proposed in Kenney & McGovern (2003), who distinguish the following “five stages of organizational response to digital preservation”:

1. Acknowledge: Understanding that digital preservation is a local concern;
2. Act: Initiating digital preservation projects;
3. Consolidate: Seguing from projects to programs;
4. Institutionalize: Incorporating the larger environment; and
5. Externalize: Embracing inter-institutional collaboration and dependency. (no pag.)

Preservation vs. curation

Although sometimes the terms digital preservation and digital curation are used interchangeably, often they are assumed to have different meanings. No complete consensus exists how to distinguish the two terms. However, a tendency exists to regard digital curation as an umbrella term for a series of activities performed throughout the entire lifecycle of digital assets, including digital preservation as one type of activity (see Yakel, 2007, p. 338). This is for example suggested by the definition by the Digital Curation Centre (DCC), according to which digital curation “involves maintaining, preserving and adding value to digital research data throughout its lifecycle” (<http://www.dcc.ac.uk/digital-curation/what-digital-curation>). At the same time, however, the DCC web pages often use “curation and preservation” in conjunction to describe the following: “Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle” (<http://www.dcc.ac.uk/resources/curation-lifecycle-model>). The situation is complicated by the advent of the term “digital stewardship” (see Lazorchak, 2011).

It follows that archives and repositories referring to the concepts of digital preservation and/or digital curation to describe their services should also provide a definition and clarify whether the terms are used interchangeably or not.



References

- Brown, A. (2013). *Practical digital preservation. A how-to guide for organizations of any size*. London: facet publishing.
- Digital Preservation Coalition. (2008). *Preservation Management of Digital Materials: The Handbook*. Retrieved from <http://www.dpconline.org/advice/preservationhandbook/introduction>
- Heslop, H., Davis, S., & Wilson, A. (2002). *An Approach to the Preservation of Digital Records*. Retrieved from http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf
- Kenney, A. R., & McGovern, N. Y. (2003). The Five Organizational Stages of Digital Preservation. In Scholarly Publishing Office (University of Michigan) (Ed.), *Digital Libraries: A Vision for the 21st Century: A Festschrift in Honor of Wendy Lougee on the Occasion of her Departure from the University of Michigan*. Ann Arbor: MPublishing, University of Michigan Library.
doi:<http://hdl.handle.net/2027/spo.bbv9812.0001.001>
- Lazorchak, B. (2011). Digital Preservation, Digital Curation, Digital Stewardship: What's in (Some) Names? *The Signal. Digital Preservation*. Retrieved August 07, 2013, from <http://blogs.loc.gov/digitalpreservation/2011/08/digital-preservation-digital-curation-digital-stewardship-what?s-in-some-names/>
- McGovern, N. Y. (2007). A Digital Decade: Where Have We Been and Where Are We Going in Digital Preservation? *RLG DigiNews*, 11(1). Retrieved from http://deepblue.lib.umich.edu/bitstream/handle/2027.42/60441/McGovern-Digital_Decade.html;jsessionid=1789D2837E80E0B5F0A4F588CC4DDEA2
- Thibodeau, K. (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In Council on Library and Information Resources (Ed.), *The State of Digital Preservation: An International Perspective* (pp. 4–31). Retrieved from <http://www.clir.org/pubs/abstract//reports/pub107>
- Yakel, E. (2007). Digital curation. *OCLC Systems & Services*, 23(4), 335–340.
doi:[10.1108/10650750710831466](https://doi.org/10.1108/10650750710831466)



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).