

Information Management Resource Kit

Module on Digitization and Digital Libraries

UNIT 4. CREATION AND MANAGEMENT OF DIGITAL DOCUMENTS

LESSON 7. PRESERVATION OF DIGITAL MATERIAL

NOTE

Please note that this PDF version does not have the interactive features offered through the IMARK courseware such as exercises with feedback, pop-ups, animations etc.

We recommend that you take the lesson using the interactive courseware environment, and use the PDF version for printing the lesson and to use as a reference after you have completed the course.



© FAO and UNESCO, 2005

Learning Objectives

At the end of this lesson you will be able to:

- understand the importance of long term preservation and access to digital material; and
- recognize the issues and factors involved in long term preservation.



Why preserve digital material?



Digital materials are acquired by our libraries from various sources.

They come **from external sources** (such as networks - especially Internet - or from physical media like audio CD, DVD, CD-ROM publications, etc.) or were **produced by ourselves**, as born digital documents or documents created through the digitization of original analogue objects.

Analogue and digital documents

Documents exist once written on an information carrier such as paper, magnetic tape, computer disk, parchment, etc.

If this writing is recorded as digital information stored on computer media, we refer to them as **digital or electronic documents**, but if this writing is materially unified with the information carrier (as ink on paper, for example), then we refer to them as **classical or analogue documents**.

Why preserve digital material?



Digital documents are often created through digitization of classical documents. They then start to have an existence that is **independent of the originals** on which they are based.

Once digitized, the originals can be properly preserved, and repeated exposure of original material to the factors accompanying the digitization process is not needed.

Thus, the **original documents have increased chances of survival**, while our **digital collections can be expanded and managed efficiently**.

Issues and factors affecting preservation

All documents, whether analogue or digital, are understandable provided that:

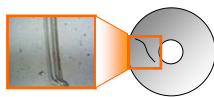
-
- A diagram consisting of three green rectangular blocks. The top block is a stack of three pages. The middle block is a single page with a diagonal slash through it. The bottom block is a single page with a magnifying glass over it. To the left of these blocks are three horizontal lines.
- they have been preserved as **objects**,
 - the information contained in them can be clearly distinguished on the **information carrier**, and
 - our information environment enables us to **decode** this information.

Issues and factors affecting preservation

All information carriers disintegrate in time; only the **speed of decay** differs.



Some information carriers are very durable (stone or parchment), while others do not last for long and lose their integrity (such as acid paper).



The carriers may also lose some of their most important physical features, which support the **readability** of written information; for example, scratches on compact discs will affect the reflectance of the layer on which the information is written.

When the carriers have been heavily damaged, it may **no longer be possible to read the information** and in this case the information ceases to exist and has not been preserved.

Data quality in analogue and digital

The quality of the information can also deteriorate when copies of the original are made. In your opinion, what kind of information is most exposed to this risk?



Analogue



Digital

Please click on the answer of your choice

Data quality in analogue and digital



Analogue information deteriorates with each copy made as it cannot be separated in its entirety from the carrier and there is no exact method of measuring the whole process. We can only **slow down** deterioration by appropriate storage and good preservation measures.

On the contrary, the digital information itself is immaterial, because it consists of a combination of digits. If we manage to copy digital files before the readability of source becomes critical, we are able to preserve everything. It is not important which carrier we use for this, but it is important to **do it on time**.

Preservation dangers in analogue and digital

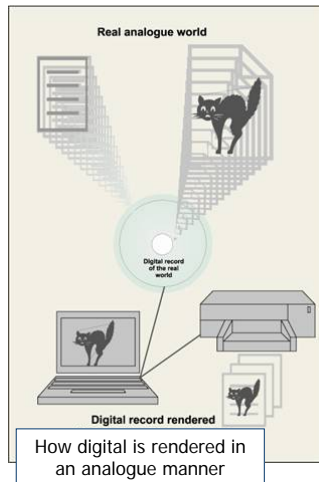
Imagine you find an 8 mm film containing an interesting documentary. It is well preserved, so you decide to watch it. Unfortunately, you realize that you cannot watch the movie, as it would require an obsolete hardware device which is not available to you.

What does the problem you have encountered relate to?

- Carrier deterioration.
- Information readability.
- Decoding.

Please click on the answer of your choice

Preservation dangers in analogue and digital



Even if the carrier has been preserved and the information is still evident, there is no guarantee that we will be able to **decode the message properly** and understand it.

This is valid for both the analogue and digital worlds. With digital objects the **speed of change** is much higher: we may lack the hardware or software to decode them.

Remember that digital information must be presented in an analogue format to be perceived by human senses, and that digital information must be rendered in an analogue manner to be read and understood.

Preservation solutions and best practices



Now that you know the factors affecting preservation, you can consider adopting some solutions for reducing these risks and preserving your digital materials.

Hardware, software, media, and formats with their parameters are continuously developing and changing. You should always be aware about what is going on around you; therefore, you should:

- monitor this development, and
- refresh or migrate your digital objects if necessary.

Let's look at these practices...

Continuous readability of media



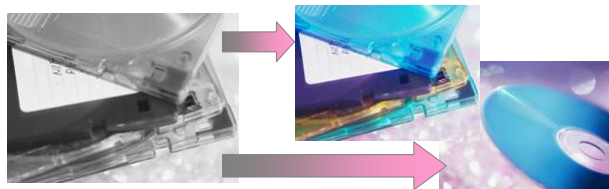
The CD measurement process

In order to ensure the preservation and readability of all your media, you should set up and undertake **monitoring** and other preservation measures, especially:

- measurement of media quality,
- storage in appropriate conditions,
- storage of the digital objects in multiple copies and locations, and
- parallel storage of digital documents in clearly defined structures and formats for use in other applications.

Refreshment

If your digital records risk losing their integrity, you should **refresh them**, that is copy them onto new fresh media of the same type or of a more modern type.



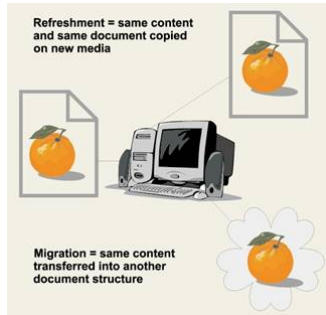
Refreshment must also be undertaken if the media containing the digital document becomes obsolete: in some cases the hardware used to read a certain format ceases to be supported and the media disappears from the marketplace and is no longer in frequent use. In this case, the refreshment should be to a newer well-supported media format, preferably a type that can be easily monitored and measured from the preservation standpoint.

Old hardware should be preserved until all the information has been refreshed to a new format.

Migration

It is not only advancements in computer hardware or media that make old solutions obsolete. The same thing happens with:

- data and metadata formats,
- software platforms (operating systems as well as access software, e.g. viewers), and
- content description standards and rules.



Often, digital objects must be **migrated to newer supported formats**. This may sometimes also require changes in descriptive elements as they need to follow new rules. In this case, you may even think about reproduction of complex digital objects. Migration often requires additional manual corrections to be done in the target document.

Migration



If we copy the whole content of a CD-ROM to a hard disk, the operation is termed data migration.

- True
- False

Please click on the answer of your choice

Migration



Obsolete media and hardware components

If the digital objects are not migrated onto new reliable structures or they are tightly bound to SW/HW platforms, the only hope is that one day new technologies **will be able to emulate** the old environment.

This is very unlikely, so do not rely on this; keep up with refreshment and migration.

If your digital objects are of great interest for users, you must keep them available. This is done, for example, with the catalogue records in a library, where the oldest records have been migrated several times already.

Emulation

Emulation involves the re-creation of the technical environment required to view and use a digital collection. This is achieved by maintaining information about the hardware and software requirements so that the system can be reengineered.

Summary

All digital objects are important component parts of our cultural memory, and we should aim at **preserving and keeping them readable**.

Preservation can be endangered by **carrier deterioration** or by the impossibility of **decoding information** due to the continuous development of hardware, software and media solutions.

To preserve your material, you can:

- use **readable metadata** and the **most frequently used data formats**,
- **monitor the quality** of the media,
- **refresh** digital records,
- **migrate** the digital objects into newer supported formats.



Exercises

The following five exercises will help you test your understanding of the concepts covered in the lesson and will provide you with feedback.

Good luck!



Exercise 1

Which of these information carriers is not suitable for recording digital information?

- DVD
- diskette
- Video 8 cassette

Please click on the answer of your choice

Exercise 2

What do the following problems relate to?

Acid paper has disintegrated into small pieces.

We can not read old shellac discs as the hardware devices to read them are about to disappear.

Carrier deterioration

Carrier deterioration

Etruscan inscriptions in stone are not understandable because we do not know the language in which they were written.

Decoding

Decoding

Stone inscriptions have been heavily damaged.

Click each option, drag it and drop it in the corresponding box.
When you have finished, click on the confirm button.

Exercise 3

During which kind of copying from one carrier to another do you lose a part of the information?

- From hard disk to compact disc.
- From video 8 cassette to VHS video cassette.
- From hard disk to diskette.

Please click on the answer of your choice

Exercise 4

Which one of the following is a migration operation?

- Copying a record from obsolete media to a new well-supported media.
- Modifying the metadata associated with your documents in order to follow new metadata standards.

Please click on the answer of your choice

Exercise 5

If we migrate digital information, can part of it be lost?

- Yes
- No

Please click on the answer of your choice

Resources on this CD

Moving Theory into Practice - Digital Imaging Tutorial

The Digital Imaging Tutorial offers base-level information on the use of digital imaging to convert and make accessible cultural heritage materials. It contains up-to-date technical information, formulas, and reality checks, designed to test your level of understanding. This tutorial is being provided by the Cornell University Library/Research Department, through the Mann Library, as a contribution to IMARK.



If you want to know more...

Online Resources:

Guidelines for the Preservation of Digital Heritage. Prepared by the National Library of Australia. Paris, UNESCO, 2003. 170 pp.
(<http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>)

Preservation management of digital materials: the handbook. Digital Preservation Coalition, London (United Kingdom), 2003 (<http://www.dpconline.org/graphics/handbook/>)

Handbook for digital projects: a management tool for preservation and access. Sitts, M.K. (ed.) Northeast Document Conservation Center, Andover, MA (USA), 2000
(<http://www.nedcc.org/digital/dighome.htm>)

Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation by Jeff Rothenberg, January 1998
(<http://www.clir.org/pubs/reports/rothenberg/contents.html>)

Preservation Metadata for Digital Collections. National Library of Australia
(<http://www.nla.gov.au/preserve/pm-meta.html>)

PADI - Preserving Access to Digital Information: Preservation Metadata
(<http://www.nla.gov.au/padi/topics/32.html>)

