

## Novice to Know-How Module Text

### Course 6: Preserving Digital Content

# Module 1: Storage for Digital Preservation

The development of this course was funded by The National Archives (UK) as part of the "Plugged In, Powered Up" digital capacity building strategy.

## 1. Introduction.

In this module we'll examine digital preservation storage. We'll consider what makes storage a vital topic and look at what the key considerations are that affect our storage choices.

## 2. The Hardware behind Bitstream Preservation.

The Bitstream Preservation module introduced how our digital data (streams of 0s and 1s) need to be stored without loss or damage, ensuring our data can be preserved over time. At a conceptual level we apply an approach that will maintain our data and allow us to mitigate potential storage risks. But what does this mean at the technology level? What hardware do we use to store our data? And how do we make sensible storage choices?

## 3. Why is Storage Important for Digital Preservation?

Without storing the 0s and 1s that make up our digital files, we will fail in our most fundamental digital preservation challenge. Just as non-digital storage mediums offer different advantages and disadvantages to their user, so does digital storage. Writing on a piece of paper is quick and practical but the paper may decay over the long term. Clay tablets are fragile, and not particularly practical to inscribe, but history has demonstrated that if looked after they can last for long periods of time. Different kinds of digital storage equally have their own attractive (and often less attractive) characteristics.

## 4. What are the Important Considerations for Choosing Storage?

The primary requirements for storage hardware are practicality and cost, but mitigation of preservation risks should also be considered.

Our storage needs to meet our requirements for access to the data – both to serve the data to users but also so that we can periodically check its integrity. The usage profile of our data might influence our choices:

- A Library might keep the majority of its data without frequent access for a long period before an event or new piece of research brings a particular collection to the fore. There may only be 5 or 10% of its data that is frequently accessed by users.
- A commercial TV station might, however, need to provide near instant access to almost all of its programmes. The majority of its data may be frequently accessed by users.

Our storage should be cost effective, particularly if large amounts of data are to be stored. So, we will need to find the right balance between quality and cost.

Meeting these requirements of practicality (and speed) of use, and cost effectiveness makes cheap, commodity storage an ideal choice. But commodity storage in turn presents a number of preservation risks. Typically it will be unreliable and have a short lifespan. Digital preservationists mitigate these shortcomings via the implementation of methodical preservation processes. Namely: keeping multiple copies, monitoring and replacing defective storage units and implementing frequent integrity checks

## 5. Choosing the Right Mix of Storage Technologies.

Digital preservationists often view a mix of storage types to be advantageous, where each copy of their data is stored using a different storage technology. In short, there is an obvious benefit to not placing all of your eggs in one basket.

If one type of storage fails, copies stored on the same hardware may also fail. This is often referred to as common mode failure. A mix of storage technologies will help to mitigate against this possibility.

Equally, if data is deleted accidentally (or maliciously) on one storage device, having other copies of data stored within in a different geographic location or under separate management could prevent the same unwanted deletion spreading (at least immediately) to other copies. Diversity of storage media can offer preservation benefit, albeit at the expense of complexity.

## 6. Environmental Impact.

It is also important to remember that the carbon footprint of digital storage can be considerable. Thinking about the environmental impact of your digital preservation activities is now more important than ever.

## 7. Wrap-Up.

So, to summarise:

- A robust approach to storage is essential to successful digital preservation.
- Our storage choices need to balance usage requirements, costs, and risks.
- We should aim for diversity amongst our storage media to mitigate risks.
- And our storage planning and processes should include:
  - Keeping multiple copies
  - Having storage in different locations

- Planning for the replacement of defective storage
- And, regular integrity checking