

Archives, libraries and museums have always organised their collections to enable users to find the information they need more readily. This function is equally important in the digital environment where the speed of development and uptake of the Internet as a publishing medium has made the discovery of quality resources increasingly hazardous. Much work has been undertaken to develop standardised means of discovering online resources, most notably the fifteen elements represented in Dublin Core ([see note](#)). Increasingly, attention is being turned to the crucial role of documentation and metadata to facilitate the preservation of digital resources. Just as metadata to support resource discovery is not a new phenomenon, neither is the importance of documentation in preservation programmes:

"Documentation has always played a key role in preservation practice. This is not just a matter of academic interest: to manage collections or individual items one needs to know what one is dealing with. There are many instances where documentation provided the only information about processes that had been applied and might need to be corrected ([see note](#))."

While the concept is not new, there are factors which make documentation particularly critical for the continued viability of digital materials and they relate to fundamental differences between traditional and digital resources:

- **Technology** Digital resources are dependent on hardware and software to render them intelligible. There are many potential permutations of technical requirements which need to be documented so that decisions on appropriate preservation and access strategies may be made.
- **Change** The resource cannot be preserved as a single physical entity over time. The information it contains will need to be separated from its physical carrier and moved across different technological platforms if it is to remain accessible. This will inevitably produce changes which may or may not significantly affect the integrity and/or functionality of the resource.
Documentation of actions taken on a resource and changes occurring as a result will provide a key to future managers and users of the resource.
- **Rights management** While traditional resources may or may not be copied as part of their preservation programme, digital resources must be copied if they are to remain accessible. Managers need to know that they have the right to copy for the purposes of preservation, what (if any) devices to control rights management, such as encryption, have been used, and what (if any) implications there are for controlling access.
- **Continuity** There will be many different decision-makers and operators and quite possibly different institutions influencing the management of digital materials across time. While traditional materials may be preserved by predominantly passive preventive preservation programmes, digital materials will be subject to repeated actions over a prolonged period of time.

- **Accountability** Documentation provides an audit trail of decisions affecting the long-term viability of the material.
- **Authenticity** Documentation may be the major, if not the only, means of reliably establishing the authenticity of material following changes.
- **Cost** It will be more complex and therefore more costly to maintain access to digital materials without documentation describing its technical characteristics.
- **Feasibility** It may not be possible to recreate the material without adequate documentation or at least not cost-effective to undertake complex restoration required as a result.
- **Future** Re-use.

Additional issues needing to be resolved are:

- **Costs** Given the complexity of digital materials and their requirements for preservation, it can be assumed that only a relatively limited set of essential preservation metadata can be automatically generated. This leads to questions of to what extent there may be overlapping needs of creators/owners and those taking on responsibility for long-term preservation of the resource:
 - What metadata needs to be/can be provided by creators/owners?
 - What will need to be/can be provided by the repository accepting preservation responsibility?
 - What is the most efficient and cost-effective means of gathering all necessary metadata and documentation prior to or simultaneously with ingest/acquisition?
 - What are the most efficient and effective means of ensuring that all necessary documentation and metadata is preserved along with the digital resource itself?

What still needs to be done?

While much progress has been made in defining what is required, actually ensuring that the information is readily and cost effectively accessible remains problematic. The technical environment changes so rapidly that software can become outdated before the repository undertakes responsibility. If a third party is undertaking responsibility for preservation the issue can become even more urgent when not even corporate memory is available to help unravel the puzzle.

Two recent studies have both drawn attention to the major obstacle of locating relevant hardware, software, and format documentation.

"Documentation for hardware and software initially ubiquitous when products are first released become increasingly difficult (and in some cases impossible) to locate over time. A concerted effort should be undertaken to collect documentation, including designs." ([Ross and Gow 1999](#))

An investigation undertaken by Cornell ([Lawrence 2000](#)) found that successful migration programs were significantly hampered by the disparity between openly published file format specifications and the increasing use of modifications to the basic standard, the latter being rarely, if ever, publicly available. Their conclusion was that:

"There is a real and pressing need to establish reliable, sustained repositories of file format specifications, documentation, implementation guides, and related software. Cornell recommends the establishment of such repositories as a prerequisite to the development of an effective national strategy." ([see note](#))

Until these recommendations have been implemented, it will continue to be a hazardous and time consuming task successfully to preserve digital resources.

[See Exemplars and Further Reading](#)