Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information

Maureen Pennock
Research Officer
Digital Curation Centre
UKOLN, University of Bath

Email: m.pennock@ukoln.ac.uk
Web site: http://www.dcc.ac.uk http://www.ukoln.ac.uk/

Digital data and technologies have fast become an integral aspect of 21st century life. Looking up information, taking digital photos, shopping online, accessing digital entertainment or online government services, and even electronic socialising and communications: all are commonplace activities for people in the 21st century. It is understandable therefore, that public organisations - particularly libraries, archives and educational establishments – face increasing demands for digital services from users who routinely and unthinkingly use or depend upon digital information in other walks of life. Many such organisations already provide access to digital information through their websites, or on physical storage media such as CD-ROM. Yet the provision of digital information online is only the first layer of the digital information challenge: content and services must not only be collected or created, but also properly managed, stored, and preserved in order to maximise the initial investment and ensure that objects and information remains reliable and available for users for as long as is deemed necessary. Digital curation is fast becoming recognised as the most viable way to meet this challenge and keep digital resources authentic and re-usable for future users.

What is Digital Curation?

Digital curation, broadly interpreted, is about maintaining and adding value to a trusted body of digital information for both current and future use: in other words, it is the active management and appraisal of digital information over its entire life cycle.
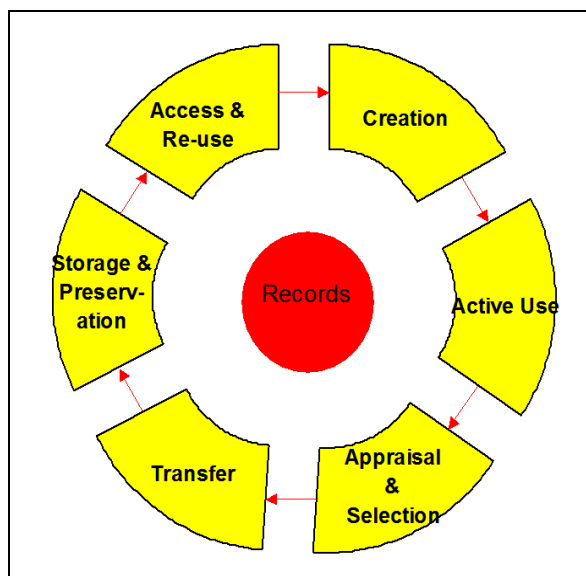
Figure 1: Sample life cycle model

The life cycle approach is necessary because:

- Digital materials are fragile and susceptible to change from technological advances throughout their life cycle, i.e. from creation onwards;

- Activities (or lack of) at each stage in the life cycle directly influence our ability to manage and preserve digital materials in subsequent stages;

- Reliable re-use of digital materials is only possible if materials are curated in such a way that their authenticity and integrity are retained.

Approaching digital information management from a life cycle perspective facilitates continuity of service; this in turn supports verification of the provenance of digital data despite technological and organisational changes in their context, and helps to maximise the initial investment made in creating or gathering them. It does however, require significant input and buy-in from the range of stakeholders with responsibilities for the materials at different stages of the life cycle. These range from creators of digital material to curators such as librarians and archivists, IT staff, and other stakeholders, including management.

Complete life cycle control is an ideal scenario. However, it is not always possible. Many organisations, particularly libraries and archives, manage and host resources created by external parties. It is not always possible in such instances to control external creation and resource management practices. Communication between these stakeholders and documentation of activities is key to achieving life cycle continuity in such a scenario.

In the UK, research and development into digital curation requirements and activities are the core activities of the Digital Curation Centre (DCC). Founded to support UK institutions which store, manage and preserve digital information, the DCC helps address the enhancement and continuing availability of digital information for long-term use. As a centre of expertise in digital curation and associated digital preservation, the DCC provides a national focus for research and development into curation issues and promotes expertise and good practice, both nationally and internationally, for the management of digital outputs.

Useful projects and initiatives

Although institutions across the cultural heritage sector, particularly museums, libraries and archives, can have very different requirements and approaches to curating their materials, the digital curation challenge is one which all types of institution face. The way in which it is tackled can vary significantly and may be affected by many factors including type of digital information, resource availability, the stage at which an institution takes responsibility for resources, institutional and object requirements, and even institutional policy. Despite this variance, the challenge remains a shared challenge and so initiatives from different domains can prove useful to others, not only in terms of techniques and lessons learned but also in shared services and avoiding unnecessary duplication of work. There is insufficient space to go into detail here, but some useful initiatives (from a range of domains and sectors) include:

- The UK Web Archiving Consortium (UKWAC), a consortium project between partners from different sectors to create a publicly accessible archive of social, historic and culturally significant web based materials;

- The International Internet Preservation Consortium (IIPC), which is developing an open source web archiving toolkit to support acquisition of websites, collection management, storage & maintenance, and access & finding aids;

- The Lots Of Copies Keeps Stuff Safe (LOCKSS) collaborative initiative for preserving subscription journals;

- Activities from each of the National Libraries and Archives, particularly those in the UK;

- The Digital Preservation Coalition, fostering joint action across UK organizations to address the digital preservation challenge in the UK.

The Digital Curation Centre is contributing to the curation challenge with a number of tools, research, services and outreach programmes that can assist institutions in addressing their curation responsibilities. As the UK's premiere source of information on curation and preservation, the DCC not only develops its own guidance and advice, but is a one-stop-shop for preservation and curation information and resources from around the world.

The DCC is jointly funded by the Joint Information Systems Committee (JISC) and the e-Science Core Programme. Further information about the Digital Curation Centre, including its tools, research, user services and community engagement activities, can be found on the DCC website. http://www.dcc.ac.uk/