Preserving digital objects

Hilary Berthon

National Library of Australia

New digital information technologies have introduced new ways of creating, storing, manipulating, distributing, finding and using information. Through networks, documents may be accessed by remote users; multimedia objects which may be interacted with are now available; and there are powerful new ways of searching for and relating documents. Despite this apparently promising future, some have warned of the dangers of entrusting the ongoing survival of our heritage to 0s and 1s. At the National Preservation Office's *Capturing the rainbow* conference last year, Maggie Exon of the Curtin University of Technology succinctly described the predicament we face thus: There is a very real possibility that nothing created, stored and disseminated electronically will survive in the long run.

What are the characteristics of the digital object which make it so vulnerable? Over the millennia, there has been a progressive shift from recording information at low density on media of high stability to high density storage on media of low stability. A clay tablet, for example, holds approximately thirty-four characters per square inch; a newspaper holds around 174 characters per square inch; microfilm averages 10 000 characters per square inch; a floppy disk, 106 200 characters per square inch; and an optical disk can store 50 million characters per square inch. The life expectancy of a clay tablet is 10 000 years whereas that estimated for optical disks is five years. Interestingly, the short lifetime of the new storage media has less to do with their intrinsic stability than the rapid obselescence of the technologies used to create them.

The methods which libraries will adopt for preserving digital information will depend on how long the information is needed. Migration is the strategy currently recommended for material deemed to be of enduring value. Migration is defined in the recent report of the United States Task Force on Archiving of Digital Information as the periodic transfer of information from one hardware/software configuration to another or from one generation of computer technology to a subsequent generation.

Migration includes refreshing (copying information from one medium to another) but takes account of the fact that it is not always possible to copy a document exactly from one system to the next. As the Task Force report indicates, there are various levels at which the integrity of a document may be retained ranging from the bitstream (0s and 1s) level through to structure and intellectual content. Naturally, migration which is sensitive to the intellectual content of the document is to be desired.

Significant challenges exist in preserving digital information. Among the issues currently being discussed are: preserving contextual information (for example, the links in hypertext), metadata, standards and maintaining the authenticity of material capable of being interacted with and altered. Also, there is a growing recognition that preservation strategies need to be formed at the time of a document's creation. New organisational strategies are undoubtedly going to be crucial in the new digital environment and, according to a recent US Task Force report, the key that unlocks the path to the economies of the digital environment will not be technological but organisational.

There has been some convergence of interest in digital preservation both across disciplines and across sectors; indeed, forging cooperative links and developing cross-sectoral alliances will be crucial if digital material is to survive. This was the rationale behind the process of developing principles for the preservation and long-term access to Australian digital objects. This process began at a workshop convened by the National Preservation Office in Canberra last year in which thirty-five representatives from a range of sectors met. Since then, input has been sought from all with an interest in the preservation of information in digital form. A draft statement of these principles is available on the Internet at: http://www.nla.gov.au/3/npo/natco/draft.html.

Another cooperative initiative is the project to develop guidelines for the preservation of digital information which have relevance across sectors. This project is being undertaken by the Preservation of Australian Digital Information Working Party which was set up as a result of a Towards Federation 2001 resolution. The guidelines will be available on the Internet (http://www.nla.gov.au/dnc/tf2001/elecwp.html) and will take the form of a dynamic document encompassing new issues as they emerge.

The introduction of preservation codes for *Conspectus* will facilitate collaborative preservation activities within the Australian library community (see http://www.nla.gov.au/dnc/conspect/modpres.html). These codes are relevant to all library material including that stored in digital form. Australian libraries will be able to record decisions relating to preservation by the assignment of preservation priority codes and the inclusion of preservation notes in their *Conspectus* entries enabling a national picture of preservation strategies and responsibilities to be developed. The National Library of Australia will shortly release a position paper on a national stragegy for Australian electronic publications.

Will the digital information held in libraries today be readable in twenty years time? Our ability to cooperate to find creative ways of addressing the new issues which emerge will be crucial in ensuring that readers now and in the future can take advantage of the benefits which the new technologies offer. ■

...the key that unlocks the path to the economies of the digital environment will not be technological but organisational.