Author: R. John Robertson
Contact details:
Centre for Digital Library Research
Department of Computer and Information Sciences
University of Strathclyde, Livingstone Tower
26 Richmond Street, Glasgow, G1 1XH, UK
Tel: +44 (0) 141 548 5854
Email: robert.robertson@cis.strath.ac.uk

Title: Digital Preservation in the Tertiary Education Sector: management implications

Paper type: Viewpoint

Structured abstract:

Purpose
This paper assesses the future of long-term curation and preservation of digital assets with particular reference to Further Education (FE) in the UK.

Methodology/Approach
Reviews current requirements of digital preservation and the efforts underway to support them. Drawing on other recent work and the author's experience in a recent development project it subsequently comments on these efforts in the context of FE.

Findings
Argues that the long-term curation and preservation of digital assets produced by further education colleges should not be the responsibility of those colleges.

Research limitations/implications
Written with direct reference to the UK; how the suggested solution would be applied elsewhere remains undeveloped.

Practical implications
National memory institutions should strive to establish collaborative curatorial practices with the FE sector.

Originality/value of paper
Suggests that the preservation of digital assets in the FE sector requires a different approach than the HE sector.

Keywords: Digital Preservation; Digital Curation; United Kingdom; Further Education; National Libraries and Archives

Digital Preservation in the Tertiary Education Sector: management implications

Introduction

Preservation of digital assets across sectors requires a co-ordinated approach at a functional as well as a strategic level. The development of effective preservation of digital assets within some communities requires not only strategic agreements, or shared tools, but also a transfer of curatorial responsibility. This paper will make specific reference to the further education sector and suggest that for the preservation of digital assets some of the responsibility for their long-term curation has to pass, in a structured manner, to national memory institutions or their equivalent.

The selection and preservation of digital assets, such as documents, images, datasets, and learning objects presents a more complex challenge. This is, not least

because such assets require not only the physical preservation of the asset but the preservation, emulation, or creation of appropriate technology to make the asset accessible, and, in some cases, enough ancillary information to make the asset comprehensible. This challenge of maintaining access to digital assets over time has been discussed extensively (see, for example: Cordeiro, 2004; Muir, 2001) and is the subject of both substantial funding initiatives and growing support and advisory networks. In the UK this has been exemplified by the Joint Information Systems Committee's (JISC's) Supporting Digital Preservation and Asset Management in Institutions programme (http://www.jisc.ac.uk/index.cfm?name=programme_404) and the Digital Curation Centre (DCC) (http://www.dcc.ac.uk). The rate of technological change implies that such long-term preservation requires a degree of ongoing curation of assets. Unlike physical assets, digital assets do not benefit from benign neglect.

Briefly stated, the long-term curation of digital assets includes an ongoing response to file format and software changes, ongoing monitoring of access to storage media, ongoing transfer of assets to new media, and an ongoing response to changes in metadata standards.

Digital curation and preservation in tertiary education

In an attempt to address the aforementioned challenges within the tertiary education sector efforts are underway to develop national policies to support digital curation and preservation. Within the UK much of this work has been carried out by co-ordinating bodies such as the Digital Preservation Coalition (DPC) and the DCC. With the establishment of such organisations and bodies, the co-ordination of national policies, initiatives and organisations appears to making significant progress (on the need for such co-ordination see: Chilvers, 2002). Alongside this co-ordinated approach, research councils are establishing data centres and mandating that funded projects deposit their data in them; thus ensuring that annotated data is available through a central location for reference and preservation.

From the perspective of the institution engaging in digital preservation, complying with these efforts is not cost-free. One group engaging with the issue of cost as part of the challenge of sustainable preservation is the *espida* project (http://www.gla.ac.uk/espida/) based at the University of Glasgow, which argues that successful preservation must make "a business case and […] engage with the 'real' key players (i.e. both Senior Management and the creators of digital information resources)" (McKinney, 2005). This approach provides a key insight when thinking about preservation – unless digital preservation is directly related to the core business of an institution it will not be sustainable. To support such an approach their project is developing business models (addressing the issues of cost, value, and risk) that engage with the relationship between preservation and institutional priorities.

For any given institution some of the costs incurred by supporting digital preservation are negligible, but others represent a significant investment. Negligible institutional costs to digital preservation are those which are incurred as part of the core business, costs which are incurred in the course of core business functions include the provision of disk space, computing infrastructure, and ensuring ongoing access to assets over their 'useful' lifetime (i.e. as long as they are required in direct connection with the core business of the institution). From a local perspective, costs incurred to provide service monitoring and to advise about file-format changes, or those incurred to create migration software capable of refreshing common file formats can also be considered negligible as they will – in most cases - be subsumed by the wider community (e.g. it is reasonable to assume that someone will ensure that an

Adobe Acrobat or Word document created today will be accessible in five years, simply because the formats are so ubiquitous).

Costs incurred in a local setting become much more significant when staff time is specifically assigned to curation, when custom file formats are in use, or when locally-developed software is involved. The assignation of staff time to specifically support digital curation is required in two areas: the ongoing monitoring outlined above and the staff time involved in attempting to provide adequate metadata about an asset. Such metadata requires a greater level of detail as not only should it describe what the object is, but it should include information about the structure, provenance, and technical requirements of the asset. Maintaining locally produced file formats or software can also incur significant expense in the longer-term as they may need to be migrated to new systems or re-developed as other systems on which they depend, or interact with, change.

Within Higher Education

The case for digital curation in the university setting is bolstered by the fact that universities are used to the idea of preservation; some of the oldest collections of books and archives exist in, and because of, universities. More importantly, many of the digital assets to be preserved in a university setting are the product of the university's core business: research. If not the product of research, other digital assets in a university such as learning objects or websites are arguably the subject of future research. There is an obvious connection, but, even with this direct relation to research and the preservation of the scholarly record, the investment involved is encouraging universities to engage in collaborative ventures to share the costs and responsibilities inherent in such endeavours (such as the aforementioned DPC or shared data centres). Needless to say, no university will be able to preserve and curate every digital asset it produces and it will primarily be assets related to the local specialities and foci of research that receive curatorial attention – just as it is for physical assets. As an overall approach, the facilitation of long-term local curation of digital assets fits with the outlook and purpose of universities.

Within Further Education

The setting of FE, however, presents a different problem for digital curation and preservation. If it can be argued that, as a result of digital preservation's direct relation to research, universities should care about curation, the inverse argument can be made that FE colleges should not.[1] The long-term preservation of digital assets is neither part of, nor related to, the core business of a FE college. Consequently there is little motivation to curate an asset beyond its immediate utility; more than this, the efforts required to support curation are likely to be in direct conflict with a focus on the core business.

In the recent report of the UK Department for Education and Skills (dfes), *Realising the Potential: A review of the future role of further education colleges* (2005), Sir Norman Foster sets out an understanding of the current and future role of FE. In this report he states that an FE college must: "Be absolutely clear about its primary purpose: to improve employability and skills in its local area contributing to economic growth and social inclusion".

The report elucidates on this primary purpose by making strategic recommendations. Two of these recommendations are that FE colleges should support a greater use of e-learning and upgrade their information technology infrastructure (Foster, 2005). It is suggested that the report's recommendations are compatible with

digital curation in the short-term but that the effort required for longer term preservation is incompatible with the remit of FE. In the short-term, acting on the report's recommendations supplies the basic storage facilities needed for the curation of digital assets and the metadata required to support the functionality of e-learning also fulfils at least some of the information required for curation. In a sense these are the negligible costs of curation.

Committing to longer-term preservation requires further investment in personnel, not only in assigning staff to monitor format developments and continue to check media survival or asset functionality, but also in funding the staff time required to support the capture of additional contextual information necessary to allow the preservation of an object in the longer-term. Such investment in longer-term curation is incongruous with the purpose of FE colleges and it is suggested that after a digital asset has passed from active use the best curation that can be expected is benign neglect – an approach inimical to the future accessibility of assets. Tools and services are under development that will significantly improve the creation of the aforementioned necessary contextual information (such as the DCC representation information registry) but using and supporting them will continue to carry a degree of overhead.

In line with experience gleaned via the Mandate project - a project examining the management of digital assets in tertiary education (http://mandate.cdlr.strath.ac.uk) - this article suggests that within FE it is not feasible to assume that any curation of digital assets will or should take place after the asset has passed from active use within the college. In particular, assets associated with teaching and learning are subject to rapidly changing course syllabi and are both volatile and disposable – a new course requires either a new asset or a new version.

In light of this argument that FE colleges can not be expected to take responsibility for the long-term care (i.e. curation) of their digital assets, the challenge remains: who should care for these assets?

The role of national memory institutions

It is suggested that just as copyright libraries and national archives preserve a national record of physical information assets, they should also curate digital assets for future generations. The idea that such institutions should curate digital assets, in itself is not ground-breaking; indeed many such institutions have advocated this for some time and in the UK, having secured legislation to extend the provisions of the legal deposit act to cover electronic materials, are now in a position to begin to engage in this area more extensively. This legislation provides national memory institutions with the right to ask for copies of electronically published assets. Its current application is, however, to preserve electronic journals and key websites. Taking on responsibility for preserving the digital assets of the FE sector is not yet envisioned.

National memory institutions would, however, be capable of supporting such a role in so far as they have, or are developing, as part of their core purpose the required expertise, knowledge, and infrastructure to address the curation and preservation of all sorts of digital assets. Their position as national institutions also implies that they are already in a position to control access to materials such as learning objects which, even if expired, hold some degree of commercial value.

Obstacles and partnerships

For this approach to work there are several significant obstacles, aside from those associated with digital curation in general, that would have to be overcome:

these are the challenges of scale, of supporting key points of access, and of metadata transfer.

Even with a selective approach to ingesting digital assets, any attempt to centrally curate the FE sector's digital heritage would involve processing a very large volume of assets and have to address the practical issues of how such an approach could be managed and resourced.

The resourcing of such an effort would also have to address the issue that digital assets associated with education are ideally searchable through non-bibliographic metadata as well. Metadata describing the educational use of an asset is becoming increasingly important, as it provides additional points of access and is changing how such digital assets are selected for use.

Even with unlimited funding, to suggest that any national memory institution is going to be capable of ingesting and cataloguing the digital assets of the FE sector is a tall order. To further suggest that such assets should receive educational as well as bibliographic cataloguing from scratch turns a tall order into an impossible one. Few national libraries or archives would be capable of creating such metadata from scratch. It is not merely a question of resources, but rather that some types of metadata cannot be derived from the asset itself. If, however, FE colleges have already created some of this metadata for their own asset management purposes it could be supplied with the asset. Any additional metadata required for the purposes of preservation not held by the college system could be added at the point of export or ingest by drawing on a college profile (e.g. all this metadata is created and published by college 'A') or by drawing on services in use by the national institutions (e.g. this is a .tiff file and so it has this representation information).

Such an approach involving the transfer of both assets and metadata would require a collaborative approach via the establishment of partnerships between colleges and national memory institutions. Such institutions would have to move from a pull-based model of acquiring digital assets (wherein they take the online assets as they wish) to either a push-based model (wherein they receive the assets packaged with metadata) or a modified pull-based model in which, after establishing a relationship with the college they harvest known assets and metadata (cf Muir, 2001).

Conclusions

National memory institutions are faced with three significant obstacles before they can assume custodianship of FE's digital assets. These obstacles may be significant, but they are all known obstacles, with known solutions. This paper suggests that attempting to enable FE colleges to support the long-term curation and preservation of their assets internally is always going to conflict with their primary purpose. Some other institution has to assume responsibility for the longer term care of their assets and national libraries and archives are the logical choice.

[1] Comments made about the FE sector in the United Kingdom may be directly transferable to parallel settings in other countries, such as community colleges (in North America). The outworking of the proposed solution may not be as directly transferable but it is suggested that an approach of this sort is necessary.

References

Chilvers, A. (2002) "The super-metadata framework for managing long-term access to digital data objects: a possible way forward with specific reference to the UK," Journal of Documentation Vol. 58 No 2, pp146-174.

Cordeiro, M. I. (2004) "From rescue to long-term maintenance: preservation as a core function in the management of digital assets," VINE: The Journal of Information and Knowledge Management Systems Vol. 34 No 1, pp6-16.

Foster, N. (2005) Realising the Potential: A review of the future role of further education colleges. http://www.dfes.gov.uk/furthereducation/fereview/finalreport.shtml. Last accessed 14th December 2005.

JISC (2002) A Continuing Access and Digital Preservation Strategy for the Joint Information Systems Committee. http://www.jisc.ac.uk/index.cfm?name=pres_continuing. Last accessed 14th December 2005.

McKinney, P. (2005) "espida and Sustainable Digital Preservation," Archives, Records Management & Conservation Vol. 193. Pre-print available at http://hdl.handle.net/1905/443. Last accessed 14th December 2005.

Muir, A. (2001) "Legal Deposit and Preservation of Digital Publications: A review of research and development activity," Journal of Documentation Vol. 57, No. 5, pp652-682.

Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D., Lyon, L. and Atkinson, M. "The Digital Curation Centre: a vision for digital curation" Paper presented at From Local to Global: Data Interoperability--Challenges and Technologies, Mass Storage and Systems Technology Committee of the IEEE Computer Society, 20-24 June 2005, Sardinia, Italy. http://www.dcc.ac.uk/docs/DCC_Sardinia_paper_final.pdf. Last accessed 14th December 2005.