LIBER QUARTERLY, ISSN 1435-5205

K.G. Saur, Munich, Printed in Germany

Preservation of Access? Developing Strategies for Microfilming and Digitisation

by GRAHAM JEFCOATE

INTRODUCTION

First of all, I should like to thank Maria Luisa Cabral and the organisers of this LIBER workshop for inviting me to give this keynote address. I hardly need an excuse to come to the Hague and to the KB, the National Library of the Netherlands, where, through working with KB colleagues, I have learnt such a lot about the future potential of European libraries over the years. When the invitation was issued last year, I had myself only recently taken charge of a large international research library. As many of you know, I have now left that post at Berlin, so that what I shall be saying today is based on my thinking about some of the issues from the point of view of library decision makers, but I won't be constrained by current institutional responsibility. Needless to say, I shall also be drawing on my personal experience of preservation and digitisation issues not only at Berlin State Library but also at the British Library, latterly as the Head of Early Printed Collections, and especially as it relates to surrogacy or reformatting.

I should stress I shall be drawing on my own professional experience, and not on any specialist expertise in the preservation field. That experience centres on historic collections, and particularly early printed materials, but also extends to digital library issues and of course to library management. As is appropriate, I hope, for a keynote address, I shall therefore be concentrating on some policy and strategic aspects of the topic, rather than providing any new insights into technical issues. I shall be posing some rather difficult questions, rather than offering any startling solutions. If I succeed in provoking debate among the experts assembled here about those wider issues, I shall have achieved what I set out today, but I apologise in advance to any who find my ruminations too vague or philosophical.

I want to begin by reviewing briefly how we got where we are today – the developing significance of microform and digital surrogacy in libraries, and especially in large research libraries. I shall then turn to the policy and strategic challenges that the current state-of-the-art and future developments present for library managers. I shall argue that they are creating a new complexity in the decision-making process. This complexity means that decisions about surrogacy can no longer be made purely on conservation or preservation grounds. Rather they need to be seen holistically in relation to the library's overall policy on acquisitions, retention and access. In other words, they have ceased to

be rather marginal in terms of library policy; the formulation and implementation of strategies on surrogacy will become increasing central to the library's whole development.

MICROFILM

Ten years ago, it all seemed so simple. Microform had established itself since the Second World War as the preferred surrogacy medium for both conservation and preservation purposes. In addition, its benefits as a useful medium for the production of facsimiles were apparent, a facsimile that was relatively simple and cheap to produce and could be successfully reproduced and, where feasible, marketed. A classic example for the half century of success for microform is the series Early English Books, produced by University Microfilms, now part of ProQuest. Eugene Power, the founder of UMI, began his work at the British Museum Library in the late 1930s, collecting and filming the canon of printed works in English from Caxton to 1700. The series was based on the *Short Title Catalogue*, the national bibliography of English works printed before 1701, so that Power could be sure his selection of texts was rational and appropriate. The urgency of the task in the shadow of war was also abundantly clear. The resulting film collection had an obvious preservation benefit, in that the texts of many rare or unique copies could be stored as durable facsimiles at a safe location away from the originals.

Much of the success of the early microfilming business was based on the record of intelligence programmes during the Second World War, when, for example, printed materials and other enemy documents where surreptitiously filmed in neutral countries and shipped for evaluation to allied capitals. Government agencies took note of this success and were willing to support filming programmes, but the commercial publishing potential was obvious too: libraries could, for example, acquire from UMI facsimile copies of early English texts of reasonable quality. This allowed them to build a critical mass of often rare or unique research materials at acceptable cost and to store them conveniently in very little space. Research libraries across Britain and the United States, as well as some on the continent, took advantage of this new accessibility. The users grumbled about the relatively cumbersome microfilm reading and copying equipment, but the access benefit was scarcely in dispute. Some information and library theorists predicted, not for the first or indeed the last time, the end of the paper era.

The British Museum Library, and later the British Library, saw multiple benefits in microfilming programmes: it obtained preservation copies of key collection works; it achieved a conservation benefit by being able to restrict access to many originals; it reduced pressure on its facilities (at least in theory) by making its collections available as photographic facsimiles at libraries abroad; and it received royalties on the sales of its

material sold by UMI. It is hardly surprising that the British Library has pursued an active policy of the preservation microfilming of key collections ever since, both through its own internally funded programmes and in close collaboration with publishing partners such as ProQuest and the Gale Group.

The proven success of microform as a medium has ensured that the British Library and other agencies across the world have accepted it as the preferred surrogacy form for preservation purposes. Preservation projects would be most unlikely to receive national funding for surrogacy, for example from the British Heritage Lottery Fund or the National Endowment for the Humanities in the US, unless their applications expressed an intention to film collections to the appropriate national or international standards. Nevertheless, microfilming has quite suddenly become controversial, at least among the general public and especially in the United States where, as the writer Nicholson Baker lamented **[1]**, newspapers and other originals, the preservation of which in original formats was regarded as uneconomic, were discarded after filming. Above all, Baker deplored the limitations of microfilm as a medium of surrogacy.

DIGITAL SURROGACY

Despite Nicholson Baker's objections, after 50 years, the history of microfilming programmes in libraries could be characterised as a record of success. By the mid-1990s, however, the shortcomings of microfilm as a medium were becoming more apparent as new surrogacy technologies began to emerge. Digital photography could provide the same or even superior conservation benefits to microfilm: the accuracy of - and level of detail in - the photographic image would mean that even fewer scholars would be able to claim the need to access fragile originals. Digital photography was achieving not only a better facsimile; linked with networking technologies, it could provide much more convenient access. Ubiquitous PCs connected to the Internet were far more popular with users than cumbersome microfilm reader-printers in libraries' reading rooms.

In common with other libraries, the British Library accordingly began to investigate the potential of the new medium during the mid-1990s in a programme called, not insignificantly, *Initiatives for Access*. **[2]** The project to digitise the Anglo-Saxon manuscript *Beowulf*, for example, showed that high-definition digital images of unique material could not only improve access in the sense of the making an accurate facsimile available at a distance from the original; digital photography was also opening up access to parts of the text invisible to the naked eye, enabling new ways of examining and researching the manuscript.

One key project under the *Initiatives for Access* programme actually addressed the potential shift from microfilming to digitisation: DAMP, the Digitisation of Ageing Microfilm project, set out to digitise the films of the *Burney Collection of Early English Newspapers*. Ironically, the relatively old and much used microfilm of the national collection of mainly 18th century London newspaper titles itself presented a preservation and access problem, its much deteriorated state hindering use of a key collection (the originals themselves had been too fragile for use by readers for several decades and were certainly regarded as too fragile to be re-filmed). DAMP sought to address this by converting one surrogate form into another in order to make the texts available in a more convenient and machine-readable form. Again, a key aspect of this project was to investigate the potential of digitisation in widening access to the content, in this case by converting and indexing the printed texts through an early version of OCR. Microfilm was itself beginning to look like a legacy medium.

But if digital facsimiles easily outclassed their microfilm equivalents in terms of conservation and access, what about that other test of the success of a surrogacy medium, preservation? Here real doubts about the long-term archiving of digital facsimiles put them at an obvious disadvantage to their analogue, microfilm equivalents. We felt we understood how to store and preserve analogue materials, even if we rarely had the resources to preserve them comprehensively; no one was quite sure about what to do with data stored on CD-Roms or hard disks. Some solutions involved conversion of digital data back to analogue media, for example, by storing images on microfilm, although this is surely an unsatisfactory solution in the longer term. Critically, few funding programmes, at least in Britain, would support preservation projects that proposed digitisation as the only surrogacy medium. A new orthodoxy appeared to be emerging by the end of the 1990s: microfilm remained the preferred preservation medium; digitisation was for access. Indeed, our own programme in Early Printed Collections at the British Library was called Digitisation for Access. We emphasised the unprecedented visibility of rare works and collections through high-definition, networked images. The Library's microfilming programmes, however, continued to receive the lion's share of budgetary allocations for collection security and preservation purposes. An analogy with retrospective cataloguing was made: the production of digital facsimiles available online was regarded as an additional access tool, equivalent – and often closely related to - the production of machine-readable records in online catalogues. Digitisation projects often needed to be funded not from preservation budgets, but from other, mostly external, sources.

THE NEW COMPLEXITY

I need hardly tell this audience today, what I might call this "interim orthodoxy" quickly began to break down, and essentially under the pressure of the spread of "born digital" materials, digital materials which are not themselves facsimiles or surrogates for analogue materials. Major research libraries and national libraries, with their responsibility for national collections and legal deposit, have been especially concerned about creating the digital equivalents of physical stores for long-term preservation. The new technologies of migration and emulation have appeared to suggest long-term solutions to the problem of rapidly developing technical platforms. We have perhaps become more confident about understanding the issues involved in the long-term preservation of digital collections. It is no longer possible simply to say, as one American commentator did in an article entitled "The Digital Dark Ages?" in the late 1990s: "Being digital means being ephemeral" (Kuny, 1998). The National Library of Australia was one of the first apparently to break with the former consensus. While acknowledging the reliability of microfilm, the NLA tells us in its Policy statement on preservation copying that it is:

"committed to producing digital copies that can be preserved, and investing in preserving both their data integrity and means of accessing them. For this reason the Library accepts its digital copies are preservation copies."

Quite understandably, some funding agencies, for example the New Opportunities Fund in its NOF-Digitise programme, would now not accept models that called for the parallel, preservation microfilming of everything that was digitised. The digital collections now being created will nevertheless need to be preserved even when resources are not available for retroconversion of the digital image to analogue microfilm. The digital images will become *de facto* the preservation masters.

If the "interim orthodoxy" is breaking down, then clearly we are now dealing with a new complexity, and one in my view that takes much of the decision-making on these issues out of preservation departments and places it at the heart of a library's strategic process. The decision-making matrix of a cost/benefit analysis would still be based on the criteria mentioned above: conservation, long-term preservation and access. Libraries will need to reach holistic decisions on surrogacy that make sense for them globally.

To get a better sense of the complexity of the decision-making process, let us return to the example for UMI's Early English Books series. By coincidence perhaps, the virtual completion of the microfilming programme in the 1990s fell at roughly the same time as the emergence of the networked digital facsimile as a viable new medium. Internet technologies had developed to enable commercial publishers to offer commercial products to online subscribers. It is hardly surprising that ProQuest and other companies

decided to bank in on the contents of their microfilm vaults by converting huge quantities of text to digital form. Essentially, this is what the music industry had already done with its legacy material on analogue data-carriers since the 1980s. Now, digitised images of microfilm facsimiles are being offered over the Internet. In a parallel development, a selection of the *Early English Texts* series is being converted into fully machine-readable and searchable form. Clearly these developments raise a wide range of issues for research libraries with an interest in this field.

Let's summarise some of these issues and consider a few of the implications. Libraries that acquired the microfilm series must now consider whether they should invest in online access. Here they will take a view about the benefits to their users of the additional functionality in view of the not inconsiderable costs of subscription. But any library subscribing to the new service would need to consider a range of complex issues. The decision will be fundamental, and probably not one that could be taken by a single department. A subscription is clearly not a conventional acquisition. Are they going to attempt to preserve the digital facsimiles on their own store? If not, then will libraries rely on the commercial supplier to maintain the digital files in perpetuity? Will they also preserve through emulation the associated functionality? If not, access will end with the cancellation subscription and the Nicholson Bakers of the future may have a new campaign cause. And what research potential might the machine-readability of a critical mass of early texts have? How will its availability in the library – and for authorised users outside the library – affect services and user expectations? Will the library be cut out of the equation altogether if faculties and their members subscribe to the service independently? If not, how might they promote and develop the service in collaboration with user communities?

In other words, what were relatively simple decisions about the acquisition, access and preservation of analogue or even hand-held digital materials have been transformed into a rather complex matrix of issues relating quite fundamentally to the library's shifting role in the information process.

SOME CONCLUSIONS

In its advertising material for catalogue conversion equipment, a British company tells us: "It is no longer necessary to make a choice between microfilming and scanning. Both have their advantages. Microfilm is a low cost and reliable system that is well tested. Scanning gives almost immediate access and ease of use with the possibility of remote access." Well, I think we can now say that it is not quite as simple as that.

My first conclusion is that programmes based on the old, interim orthodoxy ("microfilm is for preservation; digitisation is for access") will surely need to be reconsidered. I note, for example, that Metamorfoze, the Dutch national preservation programme, was still maintaining that "digitisation can only take place after microfilming, using the microfilm as an intermediate" in 2002. This is also of course the underlying assumption of the influential report on Digital imaging and preservation microfilm drawn up by preservation experts at Harvard, Yale and Cornell who argued for a hybrid approach (Chapman et.al., 1999). As a non-expert, however, I must ask whether this view is still sustainable, not least for the very practical reason that libraries cannot reject external funding for digital surrogacy if funding agencies won't support parallel microfilming. I suspect that the Australians may have developed a more realistic policy within the hybrid environment: libraries will need to recognise that resources are simply not available to replicate the media as a matter of policy, that the issue of digital preservation is now better understood, and that both microform and digital surrogates will *de facto* become preservation masters.

My second conclusion is therefore that research libraries will need a comprehensive surrogacy strategy in line with their overall policy and institutional aims and objectives and closely linked with both preservation and access strategies. Priorities will need to be set and programmes developed that are cost effective and sustainable where resources are limited. One priority might well be the digital conversion of existing microfilm collections, in order to conserve microfilm masters, improve access, and potentially open up a new machine-readability of a critical mass of texts at a relatively low cost. Cooperation between libraries and the research community might help unlock that potential.

Limited resources will mean that such strategies can no longer be developed in isolation; libraries will need to decide what digital materials they have a responsibility to preserve and where it might need to rely on external agencies or consortial arrangements; registration tools and services such as European Register of Microform Masters (EROMM) might usefully be developed. To quote a policy statement of the Association of Research Libraries: "research libraries must carefully and collaboratively select and actively preserve the most informative and representative records of past intellectual achievement in order to ensure the continuing growth of knowledge" (Responsibility, 2002). Quite so.

Questions of surrogacy should therefore be regarded as central to the strategic development of research libraries and groupings of libraries, relating closely to most core activities and the allocation of limited resources. Here are some of them:

Collection building. Libraries will need to develop criteria to inform the decision-making process on acquisitions in the hybrid environment: are resources best deployed acquiring

physical media, licensing access to online content or retro-converting existing content to a surrogate form? Where can this particular library best add value in the distributed information environment?

Retention and preservation. Within the library network, for which material does a particular library have a special responsibility to retain an item or collection in a physical or digital store? Where should it rely on an external or consortial solution? What role do national deposit centres play?

Access. How can the library improve access to its own content or content to which it subscribes or otherwise makes available? Can it forge alliances with the publishing or research communities to promote new forms of access? How can traditional library skills, for example in the production of collection description and metadata, best be deployed to add value to the information process?

And finally I should like to endorse Deanna B. Marcum, the President of the American Council on Library and Information Resources, who, in her preface to a recent report on preservation programmes in American college libraries, argues that libraries need to "approach preservation in a new way" (Kenney, 2002). She continues: "It must be integrated into every aspect of the library's work. Preservation must be considered at the highest levels of the institution and reconceived in the digital environment". This must certainly be true for the issue of surrogacy and reformatting in the hybrid library.

NOTES

1. See his *Double Fold: Libraries and the Assault on Paper*. Random House, 2001. ISBN**:** 0375504443
2. See *Towards the Digital Library: The British Library's Initiatives for Access Programme* by Leona Carpenter (Editor), Simon Shaw (Editor), Andrew Prescott (Editor), Anthony Kenny. British Library Pubns. 1997. ISBN**:** 071234540X

REFERENCES

1. Chapman, S; P. Conway & A.R. Kenney: "Digital imaging and preservation microfilm: the future of the hybrid approach for the preservation of brittle books", *RLG DigiNews*, 3(1999)1. http://www.rlg.org/preserv/diginews/diginews3-1.html#feature1

2. Kenney, A.R. & D.C. Stam, The state of preservation programs in American college and research libraries: Building a common understanding and action agenda. CLIR, 2002. http://www.clir.org/pubs/reports/pub111/contents.html

3. Kuny, T.: "The Digital Dark Ages? Challenges in the preservation of electronic information". *International Preservation News*, No. 17, May 1998, p. 13.

4. The Responsibility of Research Libraries for Preservation. May 22, 2002. http://www.arl.org/preserv/responsibility.html

WEB SITES REFERRED TO IN THE TEXT

Association of Research Libraries. http://www.arl.org/

Berlin State Library. http://www.sbb.spk-berlin.de/english/allgemeines/intro/index.html

The British Library. http://www.bl.uk/index.shtml

Digitisation for Access. http://www.bl.uk/collections/early/digintro.html

Early English Books Online. http://wwwlib.umi.com/eebo

European Register of Microform Masters (EROMM). http://www.eromm.org/

Metamorfoze project. http://www.kb.nl/coop/metamorfoze/home.html.

National Library of Australia. http://www.nla.gov.au/

New Opportunities Fund (NOF). http://www.nof.org.uk/

Policy statement on preservation copying. http://www.nla.gov.au/policy/micro.html