

## FOREWORD

In the last years of this millennium, we are either at the beginning of - or in the middle of, depending on whom you listen to - a digital revolution. In future, we are told, we shall acquire all our information and entertainment in digital form from a screen. Throughout the developed world projects are under way to create 'virtual libraries', containing digital representations of museum objects, works of art, texts and moving images. Such developments are undoubtedly bringing to a far wider audience items and texts that are currently accessible to a relative few. Initiatives such as those set out in the Library and Information Commission's report *New library: the people's network* (1997) and new sources of finance through, for example, the New Opportunities Fund of the National Lottery represent enormously important developments in widening access to such materials.

The initiatives continue and build upon the massive investment made in recent years by public and private sector organisations in the creation and acquisition of digital resources, to the extent that such resources now form a significant and growing part of our cultural and intellectual heritage. The opportunities presented by digital technology have also come to be seen by some as a panacea, as not just a means of providing access to materials that they believe have lain unused and unusable in libraries, archives and museums all over the world, but also as a means of preserving them for ever. However, there is already evidence that electronic information created or acquired with public money is being lost through neglect or through a lack of awareness of the need to take active steps to ensure its preservation.

Traditional library and archive materials (paper, vellum, even papyrus) and the materials used to bind them together (cloth and leather) present many preservation problems, with which conservators and preservation administrators have been wrestling for years, but as far as paper-based information resources are concerned the solutions are generally well understood. As we move towards ever greater dependence on electronic sources of information, however, we encounter preservation problems of a completely different order of magnitude and a completely different type. The problems of brittle paper and decaying bindings are serious ones, but, nevertheless, the time-scale for most papers becoming brittle or their bindings degrading, if they are stored in reasonable conditions, is at least a good number of decades, if not several centuries. With electronic information the time-scale is very much shorter.

The main problem is that of obsolescence. Systems change rapidly and there is no guarantee that today's software will be readable by tomorrow's hardware.

There has been concern about digital preservation in the library community for some years, but a serious and active interest is a relatively recent phenomenon. In 1996 the Commission on Preservation and Access (CPA) and the Research Libraries Group (RLG) in the USA published a joint report on *Preserving digital information* which identified problems, made recommendations and suggested areas for further research (Waters and Garrett, 1996). In the UK, in November 1995, the JISC (Joint Information Services Committee of the Higher Education Funding Councils) and the British Library addressed the question of the preservation of digital media by holding

a national conference in Warwick, where a number of action points were identified (Fresko, 1996).

The first of these was an analysis of the CPA/RLG report (then nearing completion) to identify those recommendations in it which were relevant to the UK situation and where research could usefully be undertaken to complement, but not duplicate, that in progress or planned elsewhere, particularly in the USA and Australia (Matthews *et al.*, 1997). A part of this analysis resulted in the holding of a seminar in December 1996 at the British Library, which was attended by representatives from the library and archive profession, data archives, and publishers, and where it was agreed that the JISC would fund a number of studies on digital archiving, in collaboration with the National Preservation Office (NPO), the library, archival and publishing communities.

Those studies form the basis of this book. A full list is given on page **XX**. The programme of 'JISC/NPO Studies on the Preservation of Electronic Materials' was guided by a specially established committee, the Digital Archiving Working Group, which contained experts in the field from higher education, data archives, the Public Record Office, the National Preservation Office, the British Library and the Publishers' Association (for a list of members see page **XX**). It reported to the Management Committee of the National Preservation Office. The programme was administered by the British Library Research and Innovation Centre and funded by JISC through the Electronic Libraries (eLib) Programme.

The first study (Bennett, 1997) developed a framework of data types and formats, in order to indicate the likely problems, requirements and responsibilities appropriate to each category, and to identify the most appropriate method of preservation. Closely related to this was a comparison of preservation methods and costing models (Hendley, 1998), which aimed, on the basis of a matrix of data types, to draw up a decision model to assess the most appropriate method of long term preservation and to produce a further model for comparing the costs of the preferred methods of preservation.

Two studies were concerned with the needs of data creators and the responsibility for archiving of such data. In the traditional area of publishing it is quite clear where the responsibility for maintaining an archive of published information lies: publishers do not regard it as residing with them, and if libraries wish to preserve the books or journals they have bought, then it is their responsibility to do so. In electronic publishing the issues are not nearly as clear. In many cases, for example, libraries do not hold the database - that resides with the publisher. What happens if the publisher goes out of business or loses interest in maintaining the database because the income stream from it has dried up? The report emanating from this study (Haynes *et al.*, 1997) recommended that a national body be established in the UK to coordinate such archiving and that it should be funded from the public sector, with an extension of legal deposit legislation to cover electronic publications. As far as unpublished data are concerned, universities and the funding agencies which support scholarly research are major sponsors of digital resource creation and, therefore, have a responsibility for ensuring that the research they help to create is preserved on a long-term basis. The Data Archive at the University of Essex (1998) sought to establish how much of these digital resources were being created, as well as the level of provision which is being

made for their preservation. The report also considered what the future needs of these bodies were with regard to digital preservation.

A further study (Beagrie and Greenstein, 1998) produced a strategic policy framework, which examined how different organisations are approaching the key stages in the life-cycle of digital resources, from creation, through access to preservation. Finally, the question of post-hoc rescue or 'digital archaeology' was addressed (Ross and Gow, 1998). We are already in a situation where some data appear to be inaccessible due to the obsolescence of the hardware or software required to read them. The study examined approaches to accessing digital materials where the media have become damaged, through disaster or age, or where the hardware or software is either no longer available or unknown. It illustrated some methods of recovery, showing that most data can be rescued, if there is enough time and money, but emphasising that the value of the data must be weighed against the cost of recovery.

These were all fairly short-term studies, but, given the relative paucity of research in this area, it was considered essential for this basic work to be undertaken before any commitment to funding for practical experiments was made available. The studies are also quite technical, and so it was decided to compress the results of this research into this short book which, it is hoped, will be read by policy makers and funders both inside and outside the public sector.

These preliminary studies have prepared the ground for more extensive work at a practical level to assess the problems involved in the long-term preservation of electronic information. That work is now being undertaken in several parts of the world. In the UK, JISC is funding the CEDARS project (CURL Exemplars in Digital ARchiveS), which will produce both recommendations and guidelines as well as practical and scaleable models for establishing digital archives.

A clear message that emerged from the studies was that a great deal of money can be wasted if digitisation projects are undertaken without due regard to the long-term preservation of the digital files. It is relatively easy to produce a digital version of a book, manuscript or museum object. Unfortunately it is also easy to do so in such a way that either the long-term preservation of the file becomes expensive, because of the way it was created, or with the result that the work will have to be repeated because no plan was in place for archiving the file.

But digital preservation is about much more than digitising to facilitate the preservation of items which were originally produced in a different medium. The preservation of digital materials which were *created* in the digital domain provides an even greater challenge, since there is no opportunity to return to the non-digital original. As our world moves increasingly towards dealing with this 'born-digital' information, the potentially devastating impact on the future of scholarship increases, and so must our sense of duty to solve these problems.

It is the hope of the Digital Archiving Working Group, which commissioned this book, that it will cause those who are about to embark on a digitisation project to consider the long-term archiving of the files they are about to create. The Group also

hopes that it will encourage bodies that are about to fund digitisation programmes to ensure that all proposals include a workable archiving strategy.

*Peter Fox*  
*Chairman*  
*Digital Archiving Working Group*

*February 1999*