

## 2. Digital Preservation

### 2. Outline

#### Intended primary audience

Senior administrators, funding agencies, creators and publishers, anyone requiring an introduction to the subject.

#### Assumed level of knowledge of digital preservation

Novice.

#### Purpose

- To provide a strategic overview and senior management briefing, outlining the broad issues and the rationale for funding to be allocated to the tasks involved in preserving digital resources.
- To provide a synthesis of current thinking on digital preservation issues.
- To distinguish between the three major categories of issues.
- To help clarify how various issues will impact on decisions at various stages of the life-cycle of digital materials.
- To provide a focus for further debate and discussion within organizations.

### 2.1 Strategic Overview

- Why is it necessary to take action?
- How are digital materials different?
- What digital materials are being produced?
- Who needs to be involved?
- How much does it cost?

#### Why is it necessary to take action?

More and more information is being created in digital form, either through converting existing materials to digital form or, increasingly, "born digital", where there is no other format but the digital original. There are increasing expectations in all spheres of life that the information we all need will be available on the Internet or at least in an offline digital format, such as CD-ROM. Digital access has many advantages over paper-based or microform access in terms of convenience and functionality. The increasing proliferation of digital information, combined with the considerable challenges, detailed elsewhere in this handbook, associated with ensuring continued access to digital information, means that it is imperative that there be

31

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC.

concerted action to overcome these challenges. While there is as yet only largely anecdotal evidence, it is certain that many potentially valuable digital materials have already been lost. Some of these may have disappeared without ever having reached a wider audience than the original creators (**see Reference 1**). At the very least, this constitutes failure fully to maximise the potential benefits of the investment expended in creating these digital materials.

In 1996, a specially commissioned US Taskforce on Digital Archiving published the final report of its work (**Waters and Garrett 1996**). The impact of the work of the Taskforce has been felt world-wide. In the UK, it was a key influence in a workshop sponsored by the Joint Information Systems Committee (**JISC**) and the British Library (**Fresko 1995**).

The implications for preserving continued access to important digital materials is already being felt by libraries and archives, many of which have begun to consider and take initial steps to meet their responsibility effectively. As business records are increasingly being created digitally records managers in the commercial sector and government also need to consider how they will implement records management practices which will ensure continued access to important digital records. In addition, the museums and cultural heritage sectors are increasingly utilising digital technology to create digital surrogates of rare, unique and valuable collections. The primary objective of these projects is invariably access. Preservation considerations, if they are stated at all, tend to be primarily related to the preservation of the object being digitised, not to the digital surrogate. However, a logical consequence which very quickly becomes apparent, is the question of how long access to the digital surrogate can be maintained. If access cannot be maintained beyond the short-term, then how can the initial (and often substantial) investment in creating the digital resources be justified?

### **How are digital materials different?**

As a recent Research Libraries Group (RLG) Survey noted:

"Digital materials, regardless of whether they are created initially in digital form or converted to digital form, are threatened by technology obsolescence and physical deterioration." (**Hedstrom and Montgomery 1998**)

The challenges in maintaining access to digital resources over time are related to notable differences between digital and paper-based material:

- Machine Dependency. Digital materials all require specific hardware and software in order to access them.
- The speed of changes in technology means that the timeframe during which action must be taken is very much shorter than for paper. Timeframes during which action needs to be taken is measured in a few years, perhaps only 2-5, as opposed to decades or even centuries we associate with the preservation of traditional materials. Technology obsolescence is generally regarded as the greatest technical threat to ensuring continued access to digital material.

- Fragility of the media. The media digital materials are stored on is inherently unstable and without suitable storage conditions and management can deteriorate very quickly even though it may not appear to be damaged
- The ease with which changes can be made and the need to make some changes in order to manage the material means that there are challenges associated with ensuring the continued integrity, authenticity, and history of digital materials.
- The implications of allocating priorities are much more severe than for paper. A digital resource which is not selected for active preservation treatment at an early stage will very likely be lost or unusable in the near future.
- The nature of the technology requires a life-cycle management approach to be taken to its maintenance. A continual programme of active management is needed from the design and creation stage if preservation is to be successful. This in turn leads to much more involvement both within and between institutions and changing roles.

The above issues are all interconnected and mean that a radically different approach is required in managing digital materials than for paper-based materials, one in which action needs to be taken, and planned for, at regular intervals. Retrospective preservation of digital materials is at best costly, possibly prohibitively so for any but the most highly valued, and at worst impossible. While concrete cost examples are few, it is widely acknowledged that the most cost-effective means of ensuring continued access to important digital materials is to consider the preservation implications as early as possible, preferably at creation, and actively to plan for their management throughout their lifecycle.

### **What digital materials are being produced?**

Digital materials range from relatively simple, text-based files (e.g. word processing files), to highly sophisticated web-based resources which fully exploit the benefits of the technology (e.g. combining sound with images, the ability to link to other resources, the ability to interrogate the data). There have been numerous projects to digitise collections of texts and images, primarily to utilise digital technology to improve access to these materials, which would otherwise require a visit to the holding institution.

Increasingly, resources are being created for which there is no analogue equivalent. These "born digital" materials utilise the technology to provide a level of convenience and functionality which is not possible in the analogue environment. For example, dynamic databases which are constantly updated, to produce large scale mapping or on demand publications, are continuing to proliferate. These utilise the technology very effectively for current access but pose considerable challenges in terms of the ability to maintain access to them over time and also the ability to compare data at different points in time.

Both digital surrogates of analogue originals and "born digital" resources will ultimately pose similar challenges in terms of ensuring their continued survival, though the latter may be considered the most vulnerable as there is no analogue original if they are lost (PRO 1999). In general, the more complex the materials, the more challenging it will be to ensure that they remain accessible and retain the same functionality over time.

### **Who needs to be involved?**

Because of the nature of digital materials, as outlined above, the ability to preserve access to them well into the future depends upon the involvement of a wide range of stakeholders. Principal among these are the creators of digital content, whose involvement in their preservation might involve, for example, consideration of standards in terms of format and media, and ensuring enough documentation is available to enable their management by others. Another key stakeholder will be institutions which act as long-term repositories for digital materials. They must establish an ongoing dialogue with creators and a pro-active approach to potential future accessions.

The nature of digital technology dictates that it is not feasible simply to hand over stewardship of the resource at some point in the future, without having managed it sufficiently to facilitate management by whatever repository has accepted long-term preservation responsibility. Large institutions involved in creating digital materials may most sensibly be the ones which retain them over time, thus ensuring maximum return on the initial investment of

creation. Co-operative models for long-term preservation might include a number of organisations, some of which may have experience in ensuring the preservation of paper-based materials and seek logically to extend this remit to their digital counterparts, while others may specialise in particular subject areas and/or particular types of digital materials.

All public institutions such as archives, libraries, and museums need to be involved in applying their professional skills and expertise to the long-term preservation of digital materials, just as they have taken a role in the preservation of traditional materials. Throughout the world, some of these institutions have taken a strong leadership role in addressing the practical implications of digital preservation.

For some organisations, it may prove more cost-effective to contract all or part of their digital preservation responsibilities to a third party.

Nevertheless, staff will need to be sufficiently aware of digital preservation issues, particularly as they relate to legal, organisational and contractual problems, to manage these third party contracts effectively.

Whatever model is adopted for the long term, it will need to involve the cooperation and participation of all who have an interest in creating, acquiring and making accessible, digital materials.

### How much does it cost?

Digital preservation is essentially about preserving access over time. This makes it virtually impossible neatly to segregate costs which are only for digital preservation from costs which are only about access. Access costs are significant because both the technology and user expectations advance at a very rapid rate. The initial technical infrastructure costs required for creating and/or acquiring digital materials and providing access to them are substantial. It makes sense to consider means of protecting this investment from the outset.

Preservation costs are expected to be greater in the digital environment than for traditional paper collections based on four interrelated factors:

- The need actively to manage inevitable changes in technology at regular intervals and over a (potentially) infinite timeframe.
- The lack of standardisation in both the resources themselves and the licensing agreements with publishers and other data producers, making economies of scale difficult to achieve.
- The as yet unresolved means of reliably and accurately rendering certain digital objects so that they do not lose essential information after technology changes.
- That for some time to come digital preservation may be an additional cost on top of the costs for traditional collections unless cost savings can be realised. Institutions with responsibility for both digital and traditional collections, such as deposit libraries, face the most difficult challenge, as they need to balance resources equitably between two quite different requirements. These institutions are also more likely to have a higher priority on long-term preservation as opposed to short-term access. There is scope for shared cost models and these may prove to be the most cost-effective in the long term.

While there is understandable concern that the costs of preserving digital materials will be high, it is equally important to consider the costs and implications of not preserving them. The costs of recreating a digital resource may be much higher than that for preserving it; further, the opportunity to do so may no longer exist (**see Reference 6**). An increasing dependence on both digitally produced and accessed information means that there is a rapidly growing body of digital material for which there are legal, ethical, economic and/or cultural imperatives to retain, at least for a defined period of time and, in some cases, forever. If active steps are not taken to protect these digital materials, they will inevitably become inaccessible within a relatively brief timeframe.

Further information for various categories and calculation of costs, can be found in the **Costs and Business Modelling** Section. References on specific cost subjects from digitisation to calculating the costs of preserving digital objects over time can be explored in the **Exemplars and Further Reading for Costs and Business Modelling and Creating Digital Materials**.

The issue of cost is also discussed in Organisational Issues, where it is intended to be used as a basis for awareness raising and improved understanding of cost elements as they relate to organisations.

## 2.2 Preservation Issues

### 2.2.1 Technological Issues

#### Digital media

"Digital materials are especially vulnerable to loss and destruction because they are stored on fragile magnetic and optical media that deteriorate rapidly and that can fail suddenly from exposure to heat, humidity, airborne contaminants, or faulty reading and writing devices." (Hedstrom and Montgomery 1998)

Digital media are subject to destruction and deterioration in new ways, though unintended loss can be avoided if procedures are adapted to the needs of the technology. Precautions can be taken which will help significantly to reduce the danger of loss and include:

- Storing in a stable, controlled environment.
- Implementing regular refreshment cycles to copy onto newer media.
- Making preservation copies (assuming licensing/copyright permission).
- Implementing appropriate handling procedures.
- Transferring to "standard" storage media.

However, while the media on which the information are stored may or may not fail, what is certain is that technology will change rapidly so that even if the media is retained in pristine condition, it may still not be possible to access the information it contains. No matter how exemplary the care of the media is, it will not remove the requirement to deal with changes in technology, though responsible care should make it easier to manage technology changes.

#### Changes in technology

"Unlike the situation that applies to books, digital archiving requires relatively frequent investments to overcome rapid obsolescence introduced by galloping technological change." (Feeney 1999)

Because digital material is machine dependent, it is not possible to access the information unless there is appropriate hardware, and associated software which will make it intelligible. Technology advances even in the past decade illustrate this point:

- 5¼ inch floppy disks have been superseded by 3½ inch floppy disks;



- There have been several upgrades to Windows software since it was first introduced and it would now be very difficult to convert from earlier versions to the current versions;
- Thousands of software programs common in the early 1990s are now extinct and unavailable.

The certainty that there will be frequent technological change poses a major challenge and it is therefore not surprising that collection managers quoted in the RLG survey cited technological obsolescence as the greatest threat to successful digital preservation. Precautions can, and should be taken, which will greatly reduce the risk of inadvertently losing access to a resource because of changes in technology. These include:

- Using standard file and media formats, as recommended by reputable sources.
- Providing detailed documentation to enable both context to be determined and also to facilitate successful management. (Guides to good practice are available, some of these are provided in Metadata and Documentation.).

### Authenticity and context

"At each stage of the cycle, electronic records need to be actively managed according to established procedures, to ensure that they retain qualities of integrity, authenticity and reliability." (PRO 1999)

While it is technically feasible to alter records in a paper environment, the relative ease with which this can be achieved in the digital environment, either deliberately or inadvertently, has given this issue more pressing urgency. The Public Record Office (PRO) mandates mechanisms in accordance with BSI/PD0008 for its own records (BSI 1999). The PRO draft strategy suggests that the best way to achieve authenticity is through a combination of proper processes (as outlined in their guidance documents) with data integrity mechanisms such as MD5 signatures generated at the time of ingest and the establishment of audit trails for all actions. Duranti makes a useful distinction between authentication (the means used to prove that a record is what it purports to be at a given time) and authenticity (a concept already familiar in archival science and which refers to the quality of the record itself and its essential contextual information). Records management systems need to be able to link to essential contextual information regarding the business procedures of the creating agency. Authenticity and integrity of digital resources can be equally important in other sectors. For example, scholars will need to feel confident that references they cite will stay the same over time, courts of law will need to be assured that material can withstand legal evidential requirements, government departments may well have legally enforceable requirements regarding authenticity, and so on. This issue overlaps with both legal and organisational issues and it may be one which is best resolved within individual sectors rather than through generic procedures.

### Scale

37

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC.

Although computer storage is increasing in scale and its relative cost is decreasing constantly, the quantity of data and our ability to capture it with relative ease still matches or exceeds it in a number of areas. Some repositories still face significant challenges in developing and maintaining scaleable architectures and procedures to handle huge quantities of data generated from sources such as satellites or the web. The technical and managerial challenges in accessioning, managing and providing access to digital materials on this scale should not be underestimated.

### Strategies

"Three... approaches to digital preservation have been developed:

- Preserve the original software (and possible hardware) that was used to create and access the information. This is known as the technology preservation strategy. It also involves preserving both the original operating system and hardware on which to run it.
- Program future powerful computer systems to emulate older, obsolete computer platforms and operating systems as required. This is the technology emulation strategy.
- Ensure that the digital information is re-encoded in new formats before the old format becomes obsolete. This is the digital information migration strategy." (Feeney 1999)

Strategies for some formats are well established and tested over time. For example, migration has been used for electronic text, image, and database applications by the computing industry and a number of data archives and centres for decades.

However, all three of the current strategies have potential drawbacks in some circumstances.

The need for further research has been recognised and appropriate strategies are being tested but technology will continue to evolve and will continue to raise new issues. It may well be that there will never be a single definitive strategy and a range of strategies appropriate to different categories of digital materials may need to be employed. In this way a parallel can be drawn with the paper environment which also utilises a range of preservation strategies (deacidification, microfilming, appropriate storage and handling etc.). The major difference, and the major cause of concern in the digital environment, is that failure to address the long-term access requirements of digital materials at a very much earlier stage than for paper materials will almost inevitably result in their permanent loss.

See **Storage and Preservation** for more detailed discussion of digital preservation strategies.

### 2.2.2 Organisational Issues

Note: This section has been updated by Deborah Woodyard Robinson, [March 17 2006]

While technological issues are undeniably challenging, there are also numerous challenges which relate to the ability of organisations to integrate the management of digital materials into their organisational structure. In addition, there is an increasing need to go beyond the

38

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC.



confines of individual organisations, or even countries, to maximise the benefits of the technology, address issues such as copyright, and also to overcome the challenges cost-effectively. Most organisations readily acknowledge the benefits of increased collaboration but also indicate the difficulties of, what one case study interviewee described as "differing agendas and timescales", not to mention different funding mechanisms. The following issues are being faced, and in many cases, systematically addressed, by organisations world-wide.

### Costs

The cost of digital preservation cannot be easily isolated from other organisational expenses, nor should it be. As discussed in Strategic Overview digital preservation is essentially about preserving access over time and therefore the costs for all parts of the digital life cycle are relevant. In that context even the costs of creating digital materials are integral in so far as they need to include cost elements which will ultimately facilitate their long-term preservation.

Another example of the overlap of Organisational Issues is discussed in the Expertise section. The ability to employ and develop staff with appropriate skills is made more difficult by the speed of technological change and the range of skills needed. It is also limited by resource constraints on organisations which may well need to retain the same level of ongoing commitment to and management of traditional collections and may need to integrate commitment to digital collections without additional resources.

Nonetheless the exercise of calculating costs, however complex, is a valuable and necessary task to establish cost-effective practises and a reliable business model. The cost of the labour required for digital preservation will be the most significant by far and includes not only dedicated experts but varying proportions of many staff such as administration, management, IT support, legal advisers etc.

Other major issues to impact costs include organisational mission and goals including the type and size of collections, the level of preservation committed to and the quantity and level of access required, and time frame proposed for action. These are discussed in detail in the section on Costs and Business Modelling. The relationship of costs and institutional strategies such as collaboration, third party services, rights management, training and standards are also discussed in the previous sections.

An approach to developing a successful business model which builds incrementally on:

- experience within the institution;
- collaboration with others who are confronting the same challenges;
- development of shared tools and services

will all combine to reduce risk and help develop effective strategies and practices as well as contribute to driving down costs.

See also **Collaboration** for further discussion on the advantages of co-operation between creating and archiving organisations to reduce costs.

See also **Creating Digital Material** for references to models for specific aspects of costs, such as digitisation, and maintaining digital archives. Remember however that while there is a wide amount of project-related data on costs, they may or may not have any bearing on the costs of managing digital materials long-term. The further reading and exemplars all support the view that costs for maintaining the digital copy also need to be considered from the beginning whether those materials are produced as a result of digitising analogue materials or whether they are "born digital".

### Expertise

"The need for digital preservation expertise is high: asked to rate staff as expert, intermediate, or novice, only 8 of the 54 institutions considered their staff at the expert level." (**Hedstrom and Montgomery 1998**)

The dramatic speed of technological change means that few organisations have been able even fully to articulate what their needs are in this area, much less employ or develop staff with appropriate skills. In addition, there is little in the way of appropriate training and "learning by doing" can often be the most practical interim measure. The DLM Forum (see **Reference 17**) has been undertaking work in this area for records management and has made significant progress in terms of identifying a set of six core competencies which have been used as the basis for developing training programmes for records managers.

It will take time for these developments to filter through to the workplace, and in the meantime, organisations and professional organisations need to ensure their existing staff and members can develop, and continue to develop, the range of competencies they need to manage the digital materials in their care. In addition, continuous professional development will be at least as necessary for dealing with digital materials as it is for other developmental needs. Case study interviewers have stressed the need for focussed, tailor-made courses to provide them with their specific requirements. This handbook aims to help fill a gap between current needs and existing training courses by providing guidance and tools which can be used by individuals, institutions and trainers to meet current needs.

### Organisational structures

"In addition to redefining responsibilities of organisations, it may be necessary to redefine roles within organisations to ensure long-term access to digital information." (PADI)

The nature of the technology and dependencies in the preservation of digital materials are such that there are implications for organisational structures. Organisational structures tend to be segregated into discrete elements for the efficient processing of traditional collections, but will need to cross boundaries in order to draw on the full range of skills and expertise required for digital materials. Many of the activities converge, for example decisions about acquisition and preservation should sensibly be made at the same time. Even with clearly articulated policies in place, this is likely to place strains on resources which may be seen to be competing, at least in the interim. In the absence of strong policy development, it will be

40

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC.

impossible to develop effective strategies for managing digital materials. In a worst case scenario, it may even result in a situation in which the management of both traditional and digital materials is placed at risk.

### Roles

"Although there is continuity of purpose and value within cultural institutions, these exist alongside a fundamental examination of roles and practices." (**Dempsey 1999**)

There are some existing repositories which undertake responsibility for specific subject areas or specific formats. In the UK, for example, the Arts and Humanities Data Service and Data Archive are two examples of institutions undertaking responsibility for social science and humanities research data, while the National Sound Archive assumes responsibility for its collection of sound recordings. In addition, there is work going on in other countries to establish national co-operative models for digital preservation. Examples of these can be found in Collaboration. In time, it is expected that these efforts in individual countries will crystallise into clearly defined roles and responsibilities where it is as obvious which institution is likely to be the major preserver of specific digital materials as it is for non digital materials. Despite these encouraging developments, at the present time the question of who should be responsible for ensuring long-term preservation is by no means as established in the digital environment as it is in the analogue environment.

Even when it has been determined which organisations will undertake to act on their long-term digital preservation responsibilities the environment will demand far greater engagement with a much larger group of stakeholders than has previously been the case. Some will inevitably choose to contract out all or part of their digital preservation responsibilities to a third party provider. The lifecycle approach advocated by Beagrie and Greenstein has significant implications for the way organisations responsible for long-term preservation need to interact and collaborate with data producers and publishers and each other.

Roles are also changing within as well as between institutions. Assigning responsibility for preservation of digital materials acquired and/or created by an organisation will inevitably require involvement with personnel from different parts of the organisation working together. This can potentially present difficulties unless underpinned by a strong corporate vision which can be communicated to staff. Similarly, staff working in an increasingly electronic environment are needing to modify their role to reflect the different demands of the technology.

Finally, creators of digital materials need to be able to understand the implications of their actions in terms of the medium to long-term viability of the digital material they create. Whether it be a record created during the day-to-day business of the department, a digital copy of analogue collection material, or a "born digital" resource, guidance and support as well as an appropriate technical and organisational infrastructure will assist in facilitating greatly improved prospects for efficient management and preservation.

### Selection

"In the network environment, any individual with access to the Internet can be a publisher and the network publishing process does not always provide the initial screening and selection at the manuscript stage on which libraries have traditionally relied in the print environment."  
(National Library of Canada 1998)

The enormous quantity of information being produced digitally, its variable quality, and the resource constraints on those taking responsibility to preserve long-term access, makes selectivity inevitable if the objective is to preserve ongoing access. In the digital environment, it is possible to by-pass the traditional distribution channels, as well as filtering and quality control processes. While there are benefits for users in terms of swift access, there are also difficulties in terms of quality control. Selecting quality materials for long-term retention therefore places a burden on organisations in terms of resources and also in terms of the potential impact of selection.

With traditional collections, lack of selection for preservation may not necessarily mean that the item will be lost, allowing for a comfort zone of potential changes in criteria for selection at a later stage. No such comfort zone exists in the digital environment where non-selection for preservation will almost certainly mean loss of the item, even if it is subsequently considered to be worthwhile.

In cases where there may be multiple versions, decisions must be made in selecting which version is the best one for preservation, or whether more than one should be selected. Sampling dynamic resources as opposed to attempting to save each change, may be the only practical option but may have severe repercussions if the sampling is not undertaken within a well-defined framework and with due regard to the anticipated contemporary and future needs of the users.

Some consideration also needs to be given in the selection to the level of redundancy needed to ensure digital preservation. A level of redundancy with multiple copies held in different repositories is inherent in traditional print materials and has contributed to their preservation over centuries. Although in a digital environment a single institution can provide world-wide access and accept preservation responsibility, it remains an issue of concern to many that a level of redundancy should exist in the digital environment. Such concerns need to be balanced against the potential cost in duplication of effort. Either scenario points to a greater level of overt collaboration in selection between institutions to preserve electronic publications. In any scenario, it will be critical to establish sustainability and unequivocal acceptance of responsibility to avoid the danger of losing access over time. There still needs to be assurance that preservation responsibility will be undertaken, and a clear understanding of who will undertake that responsibility and for what period of time. Otherwise there can be no guarantee that, even if several copies are stored in various repositories, all of those repositories might, for a variety of reasons, cease maintenance of the digital object at some point.

Finally, in all successful preservation strategies it may well be necessary to repeat steps in the selection process, with appropriate documentation, as part of the long-term cycle of actions to maintain access in new technological environments.

See also **Appraisal and Selection**

### 2.2.3 Legal Issues

"Compounding the technical challenges of migrating digital information is the problem of managing the process in a legal and organizational environment that is in flux as it moves to accommodate rapidly changing digital technologies." (**Waters and Garrett 1996**)

This section provides an overview of legal issues involved in digital preservation. As such it does not attempt to provide guidance on general legal issues which impact on the operations of libraries, archives and other repositories, as these are covered in a number of other reference works. It is written from a UK perspective and legislation in this area will vary from country to country. It is also an area even in the UK where forthcoming legislation such as the draft EU Copyright Directive may have a substantial future impact. Please note this section does not constitute legal advice. This is a complex and rapidly changing area and readers must seek legal advice for their specific circumstances and national legal frameworks.

Further information on implementation and further reading is listed in Rights Management.

#### Intellectual property rights (IPR) and preservation:

Copyright and other intellectual property rights (IPR) such as moral rights have a substantial impact on digital preservation. As outlined in Technological Issues the preservation of digital materials is dependent on a range of strategies, which has implications for IPR in those materials. The IPR issues in digital materials are arguably more complex and significant than for traditional media and if not addressed can impede or even prevent preservation activities. Consideration may need to be given not only to content but to any associated software. Simply copying (refreshing) digital materials onto another medium, encapsulating content and software for emulation, or migrating content to new hardware and software, all involve activities which can infringe IPR unless statutory exemptions exist or specific permissions have been obtained from rights holders.

As both migration and emulation will involve manipulation and changing presentation and functionality to some degree (especially over any period of time) important issues of principle and practice are raised in negotiations. It is important to establish a dialogue with rights holders so that they are fully aware of these issues and the actions and rights required to ensure the preservation of selected items.

#### What is different about IPR and electronic materials?

Traditional materials are relatively stable and well established legal and organisational frameworks for preservation are in place. This is not the case for electronic materials. Digital materials need consideration of both content and also hardware and software, and require very different methods of preservation. In addition in the UK there are currently no similar legal

provisions for prescribed libraries and archives permitting preservation activities on electronic items in their permanent collections: the necessary permissions must be obtained from copyright holders.

The duration of IPR in electronic materials will often extend well beyond commercial interests in them and the technology which was used to generate them. Long-term preservation and access may require migration of the material into new forms or emulation of the original operating environment: all of which may be impossible without appropriate legal permissions from the original rights owners of the content and underlying software.

### Legal deposit of electronic publications

The position on legal deposit of electronic publications in the UK is different from that of print publications (the current Act refers only and specifically to print). Voluntary deposit arrangements were introduced in January 2000. Statutory provision for legal deposit of electronic publications may follow within two years. However, voluntary arrangements need to be negotiated on a case by case basis.

### Other statutory requirements

Other statutory requirements may also apply and influence preservation of digital resources. The requirements of the Public Records Act will apply to government records including those in electronic form. Statutory retention periods will apply to many electronic records (e.g. for accounting and tax purposes). Although these are often of limited duration, it is notable that requirements for retention of electronic records in some sectors (e.g. the pharmaceutical industry), are of increasingly long duration. In such cases long-term preservation strategies will apply as technological change will almost certainly have affected access to such records.

### Access and security

Some of the additional complexity in IPR issues relates to the fact that electronic materials are also easily copied and re-distributed. Rights holders are therefore particularly concerned with controlling access and potential infringements of copyright. Technology developed to address these concerns and provide copyright measures can also inhibit or prevent actions needed for preservation. These concerns over access and infringement and preservation need to be understood by organisations preserving digital materials and addressed by both parties in negotiating rights and procedures for preservation.

### Business models and licensing

Consideration of the business models for dissemination of electronic materials and the range of stakeholders also impacts on IPR and preservation. In most cases electronic publications (particularly electronic journals) are not physically owned by the subscribers, who license access from the publisher. Subscribers are therefore concerned that publishers consider the archiving and preservation of these works and include archiving and perpetual access to back issues in licensing of these works.

### Stakeholders, contract and grant conditions, and moral rights

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC.



Electronic materials are the result of substantial financial investment by public funds (e.g. research councils) and/or publishers and intellectual investment by individual scholars and authors. Each of these stakeholders may have an interest in preservation; the archiving organisation will need to acquire permissions from them to safeguard and maximise the financial investment or the intellectual and cultural value of the work for future generations. Such interests may be manifested through contract, licence, and grant conditions or through statutory provision such as "moral rights" for the authors.

### Privacy and confidentiality

Information held within the repository may be subject to the Data Protection Act or similar privacy legislation protecting information held on individuals. Information may also be subject to confidentiality agreements. Privacy and confidentiality concerns may impact on how digital materials can be managed within the repository or by third parties, and made accessible for use.

### Investment in deposited materials by the repository

Holders of the material over many decades will almost certainly need to invest resources to generate revised documentation and metadata and generate new forms of the material if access is to be maintained. Additional IPR issues in this new investment needs to be anticipated and future re-use of such materials considered.

Where a depositor or licensor retains the right to withdraw materials from the archive and significant investment could be anticipated in these materials over time by the holding institution, withdrawal fees to compensate for any investment may be built into deposit agreements.

### References

1. See for example the Catriona II project, which discovered that quality resources were being created in Scottish Universities,'However, since they are not being created with the aim of wider access and use, they are mostly not networked, difficult to find, or in difficult to access formats'. These findings might reasonably be expected to be extrapolated to other UK institutions and by no means confined to the higher education sector. <http://wp269.lib.strath.ac.uk:5050/Cat2/index.html>  
**Update 10 Oct 2006**  
Location no longer valid. Report available at <http://cdlr.strath.ac.uk/pubs/nicholsond/catriona2final.pdf>
2. Waters, D and Garrett, J. (1996). Preserving Digital Information: Report of the Task Force on Archiving of Digital Information commissioned by the Commission on Preservation and Access and the Research Libraries Group. Washington, DC: Commission on Preservation and Access. Also available online at: <http://www.rlg.org/ArchTF>
3. Fresko, M. (1995) Long Term Preservation of Electronic Materials (1995). A Report of a Workshop Organised by JISC/British Library, held at the University of Warwick

- on 27-28 November 1995. British Library R & D Report 6238.  
<http://www.ukoln.ac.uk/services/papers/bl/rdr6238>
4. Hedstrom, M. and Montgomery, S. (1998). Digital Preservation Needs and Requirements in RLG Member Institutions. Mountain View CA: RLG. p. 3.  
<http://www.rlg.org/preserv/digpres.html>  
**Update, 06 October 2006**  
From [rlg.ac.uk](http://rlg.ac.uk) to [rlg.org](http://rlg.org)
  5. For example, the PRO's strategy states that preservation systems must be capable of providing a guarantee of rates of data loss no greater than 1 in 100,000 for data surrogates and 1 in 100,000,000 for "born digital" objects where no originals exist from which to re-create the object. Source: A Digital Preservation Strategy for the PRO. 1999. 3.6.3.
  6. For example it has been reported that it costs about \$5-7 per megabyte per year to retain electronic engineering records in the US but about \$1250 per megabyte to re-create them. See <http://ssdoo.gsfc.nasa.gov/nost/isoas/us01/minutes.html>. Costs involved in creating other digital resources, e.g. satellite or archaeological data may be higher. In addition, the data may be unique and unrecoverable if destroyed.
  7. Hedstrom, M. and Montgomery, S. (1998). Digital Preservation Needs and Requirements in RLG Member Institutions. Mountain View CA: RLG. p. 1. Also available online at: <http://www.rlg.org/preserv/digpres.html>  
Update, 06 October 2006  
From [rlg.ac.uk](http://rlg.ac.uk) to [rlg.org](http://rlg.org)
  8. Feeney, M. (ed). (1999). Digital Culture: Maximising the Nation's Investment. London: The National Preservation Office. p.11. Also available online at: <http://www.bl.uk/npo/>
  9. Public Record Office. (1999). Guidelines on the Management, Appraisal and Preservation of Electronic Records. Kew: Public Record Office. Volume 1: Principles. 4.4. Available free from the PRO or online at: <http://www.pro.gov.uk/recordsmanagement/eros/guidelines>
  10. British Standards Institution (1999). Code of Practice for Legal Admissibility and Evidential Weight of Information Stored Electronically. DISC PD 0008:1999.
  11. Duranti, L. (2000). The Impact of Technological Changes on Archival Theory. International Congress of Archives, Seville, September 2000.  
[http://www.interpares.org/documents/ld\\_sevilla\\_2000.pdf](http://www.interpares.org/documents/ld_sevilla_2000.pdf)
  12. Feeney, M. (ed). (1999). Digital Culture: Maximising the Nation's Investment. London: The National Preservation Office. p.41. Also available online at: <http://www.pro.gov.uk/recordsmanagement/eros/guidelines>
  13. Ashley, K. (1999). Digital Archive Costs: Facts and Fallacies. DLM Forum '99. p. 1.  
[http://europa.eu.int/ISPO/dlm/fulltext/full\\_ashl\\_en.htm](http://europa.eu.int/ISPO/dlm/fulltext/full_ashl_en.htm)

- Update 04 April 2008  
No longer available - information at <http://ec.europa.eu/archives/ISPO/dlm/>
14. Beagrie, N and Greenstein, D. (1998). A Strategic Policy Framework for Creating and Preserving Digital Collections. British Library Research and Innovation Report 107. London: The British Library. Also available online at: <http://ahds.ac.uk/strategic.htm>
15. Hendley, T. (1998). Comparison of Methods and Costs of Digital Preservation. British Library Research and Innovation Report 67. London: The British Library. Also available online at:  
<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html>
16. Hedstrom, M. and Montgomery, S. (1998). Digital Preservation Needs and Requirements in RLG Member Institutions. Mountain View CA: RLG. p. 41. Also available online at: <http://www.rlg.org/preserv/digpres.html>  
Update, 06 October 2006  
From [rlg.ac.uk](http://rlg.ac.uk) to [rlg.org](http://rlg.org)
17. A European forum for archivists and records managers. Its stated aim is "to achieve multidisciplinary Europe-wide co-operation to produce guidelines and to provide support to solving the issues dealing with electronic records."  
<http://europa.eu.int/ISPO/dlm>  
Update 04 April 2008  
No longer available - information at  
<http://ec.europa.eu/archives/ISPO/dlm/>
18. PADI website. Issues: Roles and Responsibilities. Roles within organisations.  
<http://www.nla.gov.au/padi/topics/8.html>
19. Dempsey, L. (1999). Scientific, Industrial, and Cultural Heritage: a shared approach. Ariadne (22). December 1999. p. 4. <http://www.ariadne.ac.uk/issue22/dempsey>
20. National Library of Canada. (1998). Networked Electronic Publications: Policy and Guidelines. p.3. <http://www.nlc-bnc.ca/9/8/index-e.html>  
Update 17 September 2007  
Link new location [http://epe.lac-bac.gc.ca/100/200/301/nlc-bnc/networked\\_epubs-ef/9/8/index-e.html](http://epe.lac-bac.gc.ca/100/200/301/nlc-bnc/networked_epubs-ef/9/8/index-e.html)
21. Waters, D and Garrett, J. (1996). Preserving Digital Information: Report of the Task Force on Archiving of Digital Information commissioned by the Commission on Preservation and Access and the Research Libraries Group. Washington, DC: Commission on Preservation and Access. p.13. Also available online at:  
<http://www.rlg.org/ArchTF>