Developing an Electronic Records Preservation Strategy

1. For whom is this guidance intended?

1.1 This document is intended for all business units at the University of Edinburgh and in particular for freedom of information practitioners and IT staff who have responsibility for ensuring that the electronic records in their area are preserved for as long as they are needed.

1.2 This guidance will enable relevant staff to decide upon the best means of ensuring that their area’s electronic records remain accessible for as long as they are needed by providing them with the tools to develop a strategy for the immediate short to medium term preservation of their area’s electronic records. This is mainly an issue for those areas not using a centrally supported IT service as areas with a centrally supported IT service are generally using onward migration.

2. What are electronic records?

2.1 Electronic records are records stored in digital format, such as e-mail, word processed documents, databases and scanned images.

2.2 You will not need a preservation strategy for all electronic records. Only electronic records that the University needs for longer than 5 years for business purposes or for their research value will normally need to be included in the preservation strategy.

3. What is an electronic records preservation strategy?

3.1 An electronic records preservation strategy will enable your area to ensure that its electronic records are maintained and sustained in an accessible format as long as a need for those records exists. The strategy will provide the mechanisms to ensure that those electronic records of current and continuing future value to the University are maintained in a way that will both preserve their value and include appropriate access arrangements to them.

3.2 This need may be business related, as certain records are crucial to the running of the University such as corporate records, research contracts, student records, human resources records. However, other records will be kept despite there no longer being a current business need for them because of their future historical importance for the University.

4. Why does the preservation of electronic records matter?

4.1 An effective electronic records preservation strategy will allow the University to meet its business needs and as a result will help it to run more efficiently. It is more costly to reproduce records that were not preserved and then subsequently needed than it is to preserve records that are never required and can be disposed of after a set period of time. For example, to retain electronic engineering records
in the US it costs $5-7 per megabyte but around $1250 to re-create them. (The Digital Preservation Handbook by Maggie Jones and Neil Beagrie, p22)

4.2 In terms of the legislative obligations that the University faces, both the Data Protection Act and the Freedom of Information (Scotland) Act require the University to be able to produce accurate and reliable information when it is asked for it. Preservation of electronic records can help by ensuring that the electronic records the University keeps are retrievable, accurate and reliable.

If the University has kept a record that is then requested and found to be in an obsolete format then it will be the responsibility of the University to convert this record and bear any costs of doing this.

5. Why do we need an electronic records preservation strategy?

5.1 The need to take special measures to preserve electronic records arises from the challenges presented when trying to ensure their long-term preservation and reliability. When a record is stored digitally it becomes significantly easier to alter it and there is a greater risk of changes to the content and therefore to the authenticity and reliability of the record.

5.2 The authenticity of the record is important; it is necessary for any organisation to be confident that the records it preserves are the same as when they were created. Without confidence in the authenticity of the record, its reliability can be questioned. Therefore electronic records must be preserved in a form and format that will maintain the record in a form as close to its original as possible and an understanding of the record and its context. As a result preservation measures should be taken as close to the original creation of the record as is possible to ensure that it is done successfully.

5.3 Further problems in preserving electronic records arise from the nature of electronic records themselves and the requirement for specific hardware and software to read them. This means that they are at risk from changing technology in terms of new software and hardware that could have an impact on the accessibility of the record. Furthermore the speed of changes in the digital world is such that the timeframes for action are short. Prompt action must be taken to guard against technological obsolescence and electronic records will generally require active preservation at an early stage, as it is more difficult to preserve an electronic record in retrospect than it would be to do so for a paper record.

5.4 Finally, the fragility of the media means that when computer disks, tapes and CDs are used they can be easily damaged unless they are kept in suitable conditions.

6. Who is responsible for the preservation of electronic records?

6.1 It is the responsibility of each business unit to develop and implement preservation strategies for the records it holds. This might be done by those responsible for carrying out the functions covered by the record, or by the nominated records officer for the area concerned.
6.2 The University Secretary has a particular responsibility in ensuring that the University corporately meets its legal responsibilities, and internal and external governance and accountability requirements. Fulfilling this responsibility requires good records management; the day-to-day responsibility for this is delegated to the University Records Management Section.

6.3 The Records Management Section will provide advice and guidance on records management issues.

6.4 Your IT support service can assist with the technological issues involved in preserving records.

6.5 The University archives service will also be involved should a particular electronic record be deemed to have future historical value. The selection criteria for records is outlined here [www.lib.ed.ac.uk/resources/collections/specdivision/criteria.pdf](http://www.lib.ed.ac.uk/resources/collections/specdivision/criteria.pdf) and should be consulted before contacting the University archives.

7. Practical steps towards the preservation of electronic records

7.1 The first step towards electronic records preservation is to find out what records you have in your care. By drawing up a comprehensive list you will know what records you have, their purpose and you may then start to assess how long they need to be kept for. If you require assistance with this, guidance is available from the Records Management Section on compiling an inventory. [www.recordsmanagement.ed.ac.uk/InfoStaff/RMstaff/Inventory/inventory.htm](http://www.recordsmanagement.ed.ac.uk/InfoStaff/RMstaff/Inventory/inventory.htm)

7.2 The next step is to know how long the records are needed for and then create a retention schedule to record this. Guidance on creating a retention schedule is available from the Records Management Section [www.recordsmanagement.ed.ac.uk/InfoStaff/RMstaff/Retention/Retention.htm](http://www.recordsmanagement.ed.ac.uk/InfoStaff/RMstaff/Retention/Retention.htm)

Only those which need to be kept for more than 5 years should be included in the preservation strategy.

7.3 A risk assessment of the records that are to be preserved for more than 5 years should then be undertaken. The issues to consider are what the impact of losing these records would be for the University, how likely is it that this would happen and the likely cost of digital preservation measures. These considerations can be represented in the following matrix:

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
<th>Very low</th>
</tr>
</thead>
<tbody>
<tr>
<td>How serious would the consequences be if we did not have the records?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How high are the risks to the records in their current format?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How expensive will it be to preserve the electronic records?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The matrix will not provide a simple answer, but is intended to help you consider the issues involved. A medium or high response to the first two questions is a strong indication that measures should be taken to preserve the records. The costs of preserving them will depend on the method chosen. The costs should be balanced against how vital the records are for business purposes.

7.4 Once this has been done, the final step is to decide which method of preservation would be most suitable; taking into account the access arrangements and preservation format required.

8. What access arrangements will be needed?

8.1 Preparing an electronic record preservation strategy will involve considering two main issues, the accessibility of the record and in what format it will be preserved.

8.2 The accessibility of a record is related to how it is stored. The two storage options that are commonly used are:

- **On-line**
  This is where records are stored on a network server. This method gives the user quick access and allows for some central control and security measures to be introduced. It also means that it is not necessary to worry about the physical degradation of disks, tapes and CDs.

- **Near-line**
  The records are stored adjacent to the server on an automatic loading device such as a DC juke-box or robotic tape storage. This device is quicker than off-line and means that physical retrieval of the record is not necessary.

Other storage options have traditionally been available but as disk space has become more affordable they no longer offer the advantages they used to. They are:

- **Far-line**
  This is where an on-line index of the records is kept, with the records themselves held on disk or tape in off-line storage.

- **Off-line**
  In this case, records are stored on media externally to the network but physical retrieval is possible upon request. This means that access to the material is slower but it frees space on the server for other more commonly used records.

The preferred method of storage is currently on-line, as this offers secure storage, back up, and also ease of access for users. However, it should be noted that electronic records stored on-line must be stored on the network server rather than the hard drive of a local computer.

8.3 There are two additional services that staff can use to store material on-line at the University. These are:
• EUCS Archive Service
This is run by Edinburgh University Computing Services and is an on-line archive service which provides a permanent store for important data files for all staff and postgraduate students. For more information please see http://www.ed.ac.uk/schools-departments/information-services/services/computing/computing-infrastructure/storage/archive/archive-service.

• Storage Area Network (SAN)
San provides fast access to on-line data; there are two versions of SAN available. One is for research purposes and is under the direction of the Edinburgh Parallel Computing Centre (EPCC). The other is a general SAN available to anyone who wishes to use its facilities. For more information on this please contact EPCC at info@epcc.ed.ac.uk or your local EUCS support team.

8.4 In addition to considering the storage options, it is also necessary to look at the file format and technological requirements. The correct file format will suit the purpose of the record, the storage facilities and aid long-term preservation. In many cases the native format will be suitable.

8.5 When considering what format to choose, you should consider for how long the record will be needed and how it will be used. For short-term preservation of records the native format will suffice. However for long-term preservation proprietary formats such as MS Word and MS Excel should be avoided and instead documents should be stored in non-proprietary, standard formats such as XML or plain text. It is not possible to preserve a database in its original format in the long-term and it will be necessary to extract the data and save it in a different format.

8.6 By saving an electronic record in the correct file format, it reduces the risks posed by the technological obsolescence of computer software and hardware. For example if a Word document is saved in rich text format it will be easier to read it on future versions of Word and other word processors than if it had been saved in a Word 2000 format.

8.7 However, when choosing a format it will sometimes be necessary to compromise between the original use of the record and preserving the data. For example, an Excel spreadsheet cannot be preserved for the long-term in the Excel format. At some point the data will have to be exported from Excel into a format that is suitable for long-term preservation, for example ASCII text or CSV (comma separated value) format. The main issue to consider here is to try and maintain a balance between long-term preservation and format that best suits the purpose of the record.

9. What are the different methods of electronic record preservation and how should I choose one?

9.1 There are several methods of electronic record preservation. Emulation and migration are the most common but technological preservation is also used, particularly in the short-term.
9.2 Emulation is software that is used to mimic a piece of hardware or software which preserves access to the electronic records. This method also preserves the original feel of the record by maintaining its appearance and style. It will require IT specialists to create this software. An example of emulation software is the creation of ZX Spectrum emulation software to play computer games.

9.3 Migration transfers records from one generation of hardware/software to the next, for example converting a document from Word 6 to Word 7. In this way the content of the record is preserved, though not always the original look and feel. However it is a labour intensive option.

9.4 Technological preservation involves keeping the original hardware and software. It can be an option for records that will only need to be kept for a very short period of time, by simply maintaining them on their existing format and equipment. You may need to liaise annually with your local IT support to confirm that this continues to be possible, as it may be dependant on continuing maintenance contracts for the equipment and software. This option carries a risk of breakdowns in the technology used.

9.5 In most cases, onward migration will be the preferred option.

9.6 The following table explains the advantages and disadvantages of each method.
<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>When should it be used?</th>
</tr>
</thead>
</table>
| **Technological Preservation**     | • Provides continued access to the original hardware/software used to create and view the record  
• The material is accessible if all the hardware, software and configurations are conserved | • By storing original equipment it retains the functionality, look and feel of the original as well as the intellectual content  
• It is a practical short-term solution to preservation  
• Delays the need to implement other preservation strategies | • It can only be a short-term strategy  
• There is a risk that the equipment will breakdown.  
• The technical support will disappear over time  
• Maintaining full access may become problematic in time  
• The use of the resource is restricted to physical access to the hardware system | • This is a good strategy to ensure limited short-term access to a valued resource  
• It will only be reliable for about 3 years after the technology concerned has gone out of common use |
| **Emulation**                      | • Develops techniques where software is used to mimic a piece of hardware or software so that access to electronic records can be preserved | • The integrity of the record can be better preserved as the original data is not altered, only the means by which to view it  
• Emulation allows the original functionality, appearance and feel of the record to be recreated | • There has been little practical use of emulation and it is still considered to be in the research stage  
• Emulation software is expensive  
• Emulation software faces the same risks from ageing technologies as other software | • Emulation could be considered if you have highly complex and valuable electronic records as it will offer the best way of preserving them with minimal alterations |
| **Migration**                      | • Records are transferred from one generation of hardware/software to the next | • The intellectual content of the electronic record is preserved  
• By using current technology maximum accessibility is guaranteed  
• Evidence suggests that it may become simpler as technology advances | • It is a process that will continually need to be done as technologies change  
• It can be costly depending on the pace of change  
• It can result in the loss of record’s original style and integrity  
• It may result in a loss of information and if used must be carefully documented | • Migration is probably the method most commonly used at present  
• It can be used for the majority of records, particularly when access to content is a priority |
10. What help is available?

10.1 The University Records Management Section provides advice, guidance and training on data protection, records management and freedom of information issues. We can be contacted at recordsmanagement@ed.ac.uk.

10.2 You should contact your IT support service for advice on IT issues, such as software and hardware requirements.

10.3 Contact the University Archives for advice on formats for long-term digital preservation.

10.4 For further information on the preservation of electronic records, see

- *The Digital Preservation Handbook* by Maggie Jones and Neil Beagrie which can be found at www.dpconline.org/graphics/handbook/.

Kiara King
August 2005