Clearing Out Legacy Electronic Records

For whom is this guidance intended?

This guidance is intended for any member of University staff who has a sizeable collection of old electronic records, such as tens of thousands of e-mails or records stored on a shared or personal drive, on tapes, CDs, floppy disks, optical drives or the EUCS archive system. It is a practical step-by-step guide to how to clear out these records, including shortcuts and issues to consider when deciding whether or not to throw the records away.

Guidance on how to deal with paper legacy records is also available. Although the principles and considerations are the same, the technique is different because of the different physical format.

What are legacy electronic records?

Legacy electronic records are collections of 'old' records (including e-mails, databases or word processed documents) with no apparent organisation, created in folders or on machines that are no longer used, or inherited from predecessors or staff who no longer work for your area. They may still be on a shared or personal drive, but could also be stored on tape, CD, optical drive, floppy disk or the EUCS archive system. In many cases you may have little knowledge of their content or importance.

Whose responsibility is it to clear out legacy electronic records?

You are responsible for clearing out the legacy electronic records that you hold in your e-mail box, personal drive, or off-line storage (such as CD, floppy disk, or the EUCS archive system), and also for those that you have created on a shared drive (unless your area has records management policies that assign the responsibility elsewhere). In the case of unattributed records on a shared drive, for example, records whose creator has since left, legacy electronic records are the responsibility of the business area that is holding them and the freedom of information practitioner would ordinarily take the lead in clearing these out.

Why should I clear out my legacy electronic records?

The only reason for keeping records is so that the information contained in them can be retrieved, whether for operational reasons or because the records have a long-term research value. If you have a collection of records, but do not know what information they contain, then there is little purpose to keeping the records. Although it is possible to retrieve information using the search facilities in e-mail and file directories, the more records you have, the more unhelpful this facility will become. As the number of records you hold increases, the search time and the number of hits will also increase until you reach a point that a search generates so many hits that it is impossible to pinpoint the required information. Storing large amounts of unnecessary information on the IT network also leads to longer IT back up and restoration time.

Keeping unnecessary information can also incur costs under freedom of information and data protection legislation. Under the Freedom of Information (Scotland) Act

2002 and data protection legislation, people have the right to ask for copies of any information held by the University. The cost of locating information cannot be passed on to the enquirer, and a maximum of 10% of other costs can be recovered. Therefore, it is in our interest to improve our management of legacy records so that we do not have to spend time searching through old information to respond to requests for information.

How do I go about clearing out a collection of legacy records?

Examining individual electronic documents or e-mails is an expensive and time-consuming process. Therefore, this should only be done as a last resort, after a number of other techniques have been used to reduce the volume of records concerned. You may find the technique identified in the third paragraph of stage 7 of this question particularly helpful.

When clearing out these records, the aim is to identify those records that should be preserved because they have a historical research value, to delete of those records that are no longer required, and to identify those records still required for business reasons and to integrate them into the current recordkeeping systems.

This section provides step-by-step instructions for clearing out a collection of legacy electronic records from e-mail accounts or personal or shared drives. Please note that these techniques are not always appropriate for databases: the next question will discuss databases in more detail.

1. Establish electronic records management systems and procedures

It is advisable to set up systems and procedures for the future management of electronic records before starting to clear out backlogs. Although it is possible to store electronic records in an unmanaged environment and clear them out periodically, it is more efficient to avoid the need to do this by setting up and following new systems, and only carrying out the one clear out as part of the transfer to the new systems.

2. Background research

Find out the background to the records. Try to answer the following questions:

If you did not do so, who created these records?
- What were their responsibilities?

When were they created?
What are the records about?
Are the records still used?
Is the version history of the records clear?
- If not, it is doubtful that there is any value in keeping an electronic copy as you don't know whether you have the current version or a draft.

Do these records have a relationship with any other records held by the University, possibly in another business area or the University Archives?
- Are they the predecessor of a current set of records?

If yes, perhaps the current records have a retention schedule that you can also apply to these records.

- Are they used in conjunction with another set of records?
  If yes, consider the two sets of records together.
- Are they the continuation of a series of records already held elsewhere?
  If yes, apply the same disposal criteria to both sets of records.
- Are they a duplicate of another set of records?
  If yes, consider which is the lead set of records. If there is no further business need for the extra set, they can be destroyed immediately. If your area has a policy that its official record will be kept on paper, it can sometimes be useful to keep an electronic copy of the information for ease of dissemination, copying and amending.

3. Identify any relevant documentation

Consider whether there any existing policies, procedures or documentation that would be relevant to these records, such as a retention schedule, either for these records or for their successors or predecessors. Look at the University's archive selection criteria within the Collections Management Policy to familiarise yourself with the sort of material that should be offered to the University Archive.

4. Conduct an initial survey of the records

Do not look at every record, but try to form an overall impression of the record collection. Look at file/e-mail titles, dates, folder titles and, for e-mails, to and from information. Open up a small, unscientific sample of the files.

Consider the following questions:

Do the records conform to your expectations given the information you have already collected?
Do the contents of the records match up with their titles?

- If the answer to either of these questions is no, you will need to be more cautious in your deletion decisions.

5. Develop selection criteria

Based on the information you have collected so far, develop some basic selection criteria for the record collection. These will include identifying types of records that can be destroyed, either now or at a date in the future, types of records that are still required for everyday business and should be integrated into current filing systems, and types of records that should be offered to the University Archive. Developing the criteria is an iterative process, as you will want to go back to them to add new decisions as your experience of working with the records grows.

In developing the criteria, refer to any existing retention schedules for related records, to the University Archive selection policy and selection criteria, and to the 'Deciding what to keep' section of the retention scheduling guidance.

6. Use short cut measures to clear out obviously unnecessary records

The precise techniques for doing this will vary depending on the type of electronic record.

For information in shared and personal drives, the most efficient technique is to use the directory management system. In Microsoft systems this is called Windows Explorer. It provides a view of the directories, folders and records in folders that you have. You can set it so that it provides details of each document, including title, size, document type, date created, date last modified, date last accessed, document owner and document author. Browse this information to identify records that are out-of-date, have been superseded by later versions or have not been used for some time. It is probable that these records can be deleted.

For e-mails, the following techniques can help to identify messages that are no longer needed:

- sort e-mails by date and delete all those over a certain age
- sort e-mails by addressee and delete all those sent to/received from certain individuals
- sort e-mails by subject and delete those relating to completed business
- sort e-mails by size and delete large e-mails that are no longer required.

Some e-mail systems save deleted items in a folder rather than actually deleting them. Microsoft Outlook is one system that does this. It is important to ensure that e-mails you meant to delete are actually deleted.

In Outlook, you can do this by setting your deleted items folder to empty whenever you exit Outlook. To do this take the following steps:
a. on the Tools menu, click Options, and then click the Other tab.
b. select the 'Empty the deleted items folder upon exiting' check box.

7. Review remaining records

There are two possible approaches to this.

The first involves looking at each file individually to assess its value in accordance with your already-determined selection criteria. File-by-file review should only be used as a last resort, once you have used the methods above to winnow down the number of files for examination. The aim should be to spend the minimum amount of time looking at each document or e-mail.

The second is far less resource intensive, but will only succeed if there are new records management systems and procedures in place for managing electronic records. Set up a dedicated folder within your e-mail account and on your personal or shared drives for the legacy records. Move all remaining legacy records into it and keep the folder and its contents for one year. During that time, whenever you or your colleagues retrieve an item from the legacy folders, you should consider at that time, whether the item is still required. If it is, it should immediately be moved to an appropriate folder in the new records management system. At the end of the year, any items remaining in the legacy folders can probably be deleted.

How should I deal with databases?

For records management purposes databases can be classified into two types: static and dynamic. Static databases are databases where the data is never amended, added to or deleted once it has been added to the database. For example, a database of survey results. Dynamic databases are those where the data is liable to change over time. For example, a database of contacts will change regularly as addresses change, new contacts are added and old ones are deleted.

It is possible to take retention decisions about static databases once the database is complete using the background information that you have about database.

For dynamic databases, arrangements must be put in place to delete information from the database when it is no longer required. Retention decisions need to be taken by considering how long individual data elements or groups of elements should be kept and measures put in place to identify unnecessary data and delete it. If a dynamic database has historical or long-term research significance, arrangements should also be put in place to take regular snap shots of the data for preservation for research purposes. Sometimes the cost or logistical implications of amending an existing system are so great that the best option is to wait until the present system is replaced, and then introduce measures to implement retention decisions.

What should I do about back up copies of electronic records?

You should confirm with your IT support service that the back up copies of your electronic records are subject to an appropriate system of tape recycling that means back up copies of electronic records will eventually be overwritten. The exact nature of this cycle will vary according to the business needs associated with particular servers, but an example would be that daily back up tapes are overwritten on a monthly cycle. An electronic copy of a deleted record may survive on a back up tape for some time after its official deletion. Provided that the back up is subject to an appropriate retention period, this delay is acceptable - it is not practical to attempt to delete individual records from back ups. We would not normally be expected to retrieve information from back up tapes to respond to freedom of information or data protection requests.

How do I decide which records I need to keep and develop my selection criteria?

In deciding how long to keep the records, you should consider their relationship with other records, the business (including legal and accountability) need for the records, the costs associated with keeping the records and their long-term research value. Figure 1 shows this process as a flow chart.

1. Consider the relationship with other records

You should ask the following questions:
Do these records support the interpretation and use of other records?
If yes, you must consider both sets of records together to ensure that their retention periods are co-ordinated.
Do these records duplicate other records, held within the business unit or elsewhere?
If yes, you should decide which set is the official record. (For example, the chair or secretary of a committee will hold the official record set of its electronics). The University only has a business need for one official record set. All others should be destroyed. (In some cases, business sections may decide to keep a duplicate set for a short time, for reasons of convenience. This might arise if you need to consult the records regularly and you do not have quick and easy access to the record set).
If these records are derived from a wider body of information, how much value do they add to the original information?
Do we need to keep the wider information and these records or will one of the sets of information suffice?

2. Consider the business need for, and long-term research value of, the records

You should ask the following questions:

For how long will there be a continuing need for this information for current business processes?
For how long are these records needed to document the business process/ decisions taken/ actions carried out for future reference use?
For how long are these records needed to fulfill legislative, regulatory or financial requirements?
For how long are these records needed for accountability purposes?

3. Consider the possible implications of not having the records, balanced against the cost of keeping the records.

Every document retained represents a resource burden for the university, in terms of storage costs, administration, and freedom of information and data protection obligations.

4. Consider if the records have a long-term research value for historical or other purposes.

These stages can be represented in the following matrix:

	High	Medium	Low	Very low
How likely is it that we will need the records again for business purposes?				
How serious would the consequences be if we did not have the records?				
How expensive is it to keep the records?				
What long-term research value do these records have?				

This matrix will not provide a simple answer, but is intended to help you to balance the issues involved in deciding whether to keep the records. A medium or high response to the first two questions is a strong indication that the records should be kept, provided that the resource implications are not too high. If a set of records is critical to the business of the organisation but the resource implications of keeping them are very high this collection will require careful consideration before fixing a retention period. Records consistently rated as low/very low are either not required for continuing business purposes or are likely to merit preservation for a short period, such as 1-2 years.

If you give a medium or high response to the last question, then you should confirm your decision with the University Archivist and take appropriate digital preservation measures. The University Archive does not presently have facilities for the archival storage of digital records, but advice on their long-term preservation is available at http://www.lib.ed.ac.uk/digpres/.

You can use this matrix to help you decide the retention period for records by considering how the answers to the questions would change over time. Possible retention periods to consider are, immediately after creation, after 6 months, after 1 year, after 2 years, after 5 years, after 10 years etc. You are unlikely to need to consider each of these periods for every record set.

You must not destroy any records which are currently the subject of a data protection or a freedom of information request until the response to the request is complete.

How do I integrate my legacy records into existing systems?

The records should be added to folders at the appropriate place in the new filing scheme and, if necessary, added to the records inventory.

I have a batch of case records of varying dates. How do I apply retention periods to them?

Frequently, the retention period for a case record is expressed in terms of a number of years after a trigger event, such as ‘destroy four years after graduation.’ With legacy records of mixed dates, it would be very time consuming to examine each file to establish when the trigger even occurred and, therefore, exactly when that record should be destroyed. It is usually more cost effective to make a conservative estimate of the likely latest retention period of any record in the set, and to assign this period to all the records.

This approach can also work with other types of files, but is most effective when dealing with case records.

What help is available?

The University’s Records Manager provides advice, guidance and training on records management issues. Although it is the local area’s responsibility to review legacy records, training and advice can be provided on deciding what records you need to keep, and tracking work in progress.

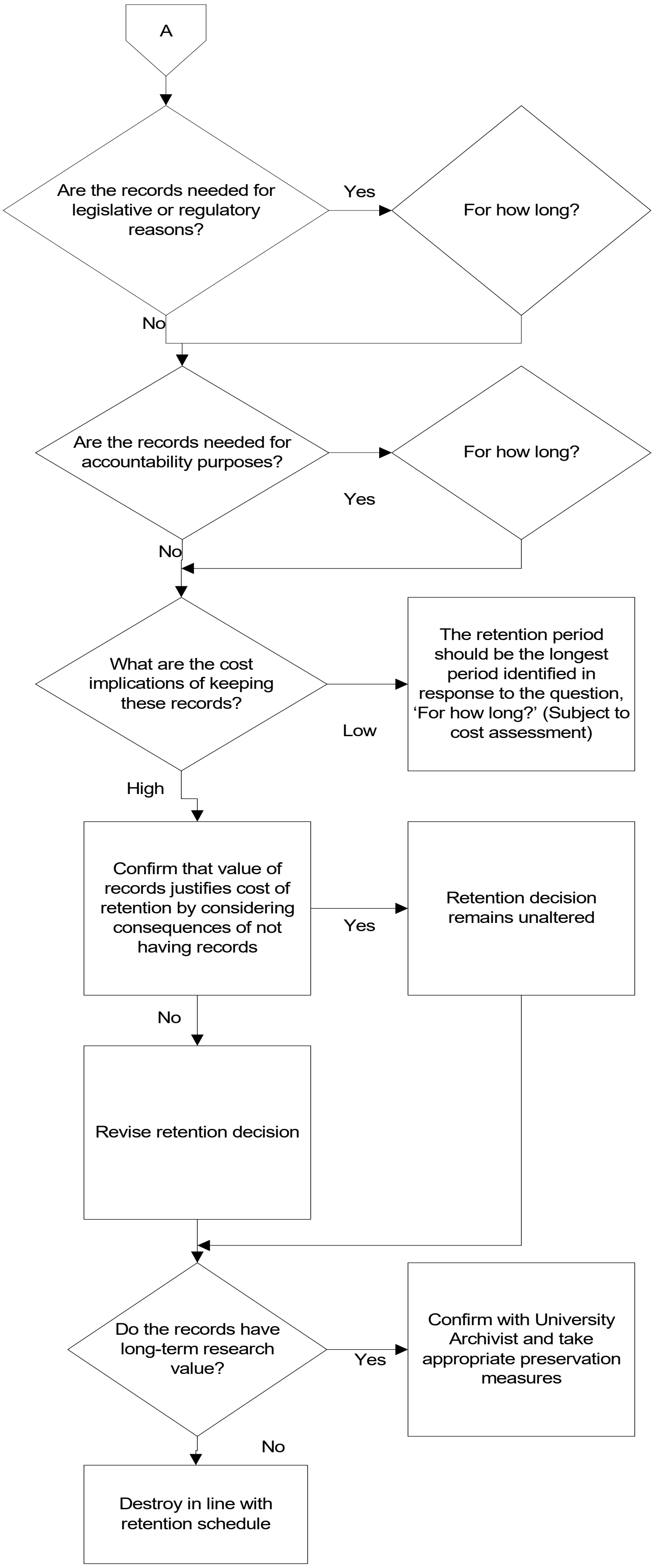
The Joint Information Systems Committee (JISC) has prepared a high-level retention schedule for the functions and activities of universities. This can use as the starting point for deciding the retention periods of our records but requires considerable customisation. It can be found at:

Records retention management - Jisc

The National Archives has prepared model retention schedules for common types of government records, including buildings, personnel, finance, health and safety, projects, press and publicity, complaints, contracts, internal audit and information management records. Although the recommended retention periods are for government records rather than University ones, these schedules are a useful source of precedents. They can be found at:

Advice on retention - The National Archives

Susan Graham
November 2004


A
Are the records needed for legislative or regulatory reasons?
Yes
For how long?
No
Are the records needed for accountability purposes?
For how long?
Yes
No
What are the cost implications of keeping these records?
The retention period should be the longest period identified in response to the question, 'For how long?' (Subject to cost assessment)
Low
High
Confirm that value of records justifies cost of retention by considering consequences of not having records
Yes
Retention decision remains unaltered
No
Revise retention decision
Do the records have long-term research value?
Yes
Confirm with University Archivist and take appropriate preservation measures
No
Destroy in line with retention schedule