Introduction

1. This document has been drawn up to provide guidance to University IT staff who need to use real data about living identifiable people in their work to test, implement and maintain the University’s IT systems. Appendix one contains a short list of the things staff must do to ensure that their work with this data complies with the Data Protection Act.

Definitions

2. The Data Protection Act is concerned with ‘personal data’. This includes any information about an identifiable, living individual processed automatically. Data about people held in University systems would qualify as ‘personal data’, as would any copies of that data used as test data.

3. The Data Protection Act also identifies ‘sensitive personal data’ as a sub-set of ‘personal data’. The University is expected to take additional precautions when dealing with ‘sensitive personal data’ because of the nature of the information. The Act defines ‘sensitive personal data’ as personal data consisting of information as to:
   - racial or ethnic origins
   - political opinions
   - religious or similar beliefs
   - trade union membership
   - physical or mental health or condition
   - sexual life
   - commission or alleged commission or any offence, proceedings for any offence, their outcome and sentence.

4. In this document, ‘IT administration’ refers to the testing, training and demonstration activities necessary to select, install and maintain the University’s IT systems. ‘Live data’ means either the actual data about living, identifiable individuals held by the University or a duplicate dataset derived from actual data about living, identifiable individuals.

The University’s use of test data

5. The development of IT systems for the University passes through three stages:
   - Development environment
   - Test environment
   - Live environment

In the development environment developers usually use anonymised or fake data, while the live environment is where the processing to meet the purposes for collecting the data is carried out. This document is mostly concerned with the test environment.
6. IT systems must be tested thoroughly as part of their implementation. Failure to test systems thoroughly could lead to errors in the way that data is processed, possibly involving a breach of the Data Protection Act. To ensure that the systems are tested thoroughly the University uses copies of live data from existing systems. This ensures that the system can cope with comparable volumes of information, that a wide range of realistic scenarios is covered, and that the test will reflect all the possible combinations that occur in the real environment. It would not be possible to invent datasets for this purpose as it would be difficult to create a sufficiently large dataset which covers the vast number of possible combinations that occur in reality.

7. For many systems the University has a permanent ‘test’ environment which replicates the live environment. The live and the test environment run simultaneously, and the test environment is updated regularly with data from the live system. The environment is used to test new developments before they are implemented in the live system, and is also used as a diagnostic and repair tool when faults are discovered in the live environment. This ensures that the robustness of fixes can be tested thoroughly before they are introduced to the live system. Failure to do this could lead to errors being introduced into the live environment, possibly corrupting the real data in a way that would involve a breach of the Data Protection Act.

8. In most cases it would not be possible to anonymise the test data as the systems have built-in validation criteria which would reject false data. (For example, the Registry database has a validation requirement that a postcode must be provided, and it checks that this is a real postcode.) Moreover, in many cases, anonymising the information would involve removing so much information that the tests would no longer be meaningful. It is necessary to test how the system responds when presented with comparable volumes of data and with the wide diversity of combinations that can occur in the real environment.

9. Most testing is done by the business users or IT staff who would have access to the original live data for maintenance purposes, so the use of test data does not necessarily entail release of the data to individuals who would not ordinarily have access to it.

The University’s use of data for training purposes

10. Training might be done using either the live or the test environment. In either case, the people accessing the data will be limited to those who would have access to it in the real environment.

11. When the live environment is used for training purposes, safeguards are put in place to minimise any risk to the integrity of the live data, and to ensure that the system can be 'rolled back' to a very recent backup version of the system.
The University’s use of data for demonstration purposes

12. When demonstrating new systems, the University either uses ‘fake’ data or the data of an identifiable individual who has consented to its use for this purpose.

Data Protection Act Requirements

13. This section sets out the legal obligations imposed on us by the Data Protection Act 1998 and the options for meeting them when using data for IT administration.

14. The Data Protection Act requires organisations to process personal data ‘fairly and lawfully’.

15. To process data fairly, the individual must know what information you hold about them, what you will do with it, and to whom you will disclose it. You should ensure that a person is told this at the time that you collect the information, or, if you receive the information from a third party, as soon as is reasonably practicable after receiving the information.

16. Students are told that data is used for systems testing as part of the data protection information they receive on their annual matriculation forms. There is a need to ensure that alumni, lifelong learners, staff and visitors are also given this information.

17. To process data lawfully, you must meet certain conditions set out in the Act. There are two sets of conditions: one set that all processing must meet, and another that only applies to sensitive personal data. To process sensitive personal data lawfully you must meet at least one condition from each list.

18. The possible relevant conditions test data from the first list (personal data in general) are:
   • You have the consent of the data subject
   • Processing the information is necessary to pursue the University’s legitimate interests and does not cause unwarranted prejudice to the rights, freedoms or legitimate interests of the data subject.
   You must meet one of these conditions or you cannot process the information lawfully. In most cases it will be the second of these conditions that we rely on.

19. The conditions for processing sensitive personal data are much more difficult to meet. For test data you must have the explicit consent of the person in almost all cases. Explicit consent requires that the person understands what you are going to do and that they confirm that they are content for you to do it. People usually record explicit consent by asking someone to sign something but this is not essential.

20. We obtain consent for processing sensitive personal data about students for test purposes as part of the matriculation process. We do not process sensitive personal data about visitors and alumni. However, with the exception of the
HR database, there is a need to obtain consent from staff to process sensitive personal data about them for these purposes. Maintaining and developing the HR database is covered by the condition that, “the processing is necessary for the purposes of exercising or performing any right or obligation which is conferred or imposed by law on the data controller in connection with employment”.

21. The Act also requires that you ensure that any information you hold is adequate for the purpose, relevant and not excessive. A full copy of the data is required for testing purposes, but for training and demonstrations it may be more appropriate to use more limited datasets.

22. The Act requires that any personal data the University holds is accurate. This will not be the case with test, training and demonstration data, as part of the purpose of the test environment is to make alterations to the dataset without affecting the real data. However, using test data in this way helps to ensure that the real data will be accurate. No decisions about individuals are taken using test data, and there is no risk that the inaccurate test data will be confused with the real data as the two are strictly separated.

23. You must not keep personal data for any longer than you need it. Any datasets created for test, training or demonstration purposes should be deleted as soon as they are no longer used for this purpose.

24. In some circumstances the individual has the right to require you to stop using data about them for this purpose. If anyone asks you not to use test data about them, then you must respect their wishes. Delete their details from any existing test datasets and do not include them in any future test datasets. Some University systems have no facility for deleting an entry from the dataset, whether in the test or the real environment. In these cases, steps should be taken to change the data so as to render the individual unrecognisable; for example, by changing the name and some of the details. If the personal data includes any sensitive personal data, this must either be deleted or changed to the most common option.

25. The individual has the right to ask you to amend incorrect information about them. If anyone asks you to amend information that is incorrect, then you must do so, even in a test dataset. In some cases this will be done by overwriting the data with a new copy of the real data. This is acceptable provided that the overwriting takes place within a reasonable time of receipt of the request to correct the incorrect information, normally within 10 working days.

26. The individual has the right to receive a copy of the information held about them in a test dataset. If anyone asks you for a copy of the information about them held in the test dataset, then you must provide one within 40 calendar days. Clearly mark the copy as test data, manually if necessary.

27. You must ensure that you store personal data so that it cannot be accessed, amended or deleted by someone who has no need to do so. In the case of test
data, this means that the data must be protected with security levels at least as high as the original data. Only staff working on the tests may see the test data, and it must not be used for demonstration purposes. For demonstration purposes, data must either be invented or the individual concerned must have consented to the use of their record in this way.

28. Test data must not be transferred to a country outside the European Economic Area (EEA) without ensuring that appropriate safeguards are in place for the rights and freedoms of the data subjects. In practice this means that they must receive the same rights as they would have under the Data Protection Act in this country.

29. Options for meeting this requirement include:
   - Only transfer data within the European Economic Area, or to countries that have been designated as providing an appropriate level of protection (currently Hungary, Switzerland, Argentina, bodies subject to the Personal Information Protection and Electronic Documents Act in Canada, or organisations that participate in the ‘safe harbour’ initiative in the USA).
   - Only transfer completely anonymised data.
   - Set up a data controller-data processor relationship within the contract. This is relatively straightforward to do, but will make the University legally responsible for any breaches of the Data Protection Act committed by the other party.
   - You amend your contract with your supplier to contain the EC standard contractual clauses for the transfer of personal data to processors established in third countries. Further information about the clauses, their use and the clauses themselves can be found in the European Commission's Decision of 27 December 2001, at: http://europa.eu.int/eur-lex/pri/en/oj/dat/2002/l_006/l_00620020110en00520062.pdf
   - Obtain explicit consent from the people concerned for the transfer of data about them. This means that the people must know to which country their data will be transferred, to which organisation, and what that organisation intends to do with the information.

30. The Records Management Section (recordsmanagement@ed.ac.uk) can provide further advice and guidance on any of these issues.

Susan Graham
June 2005
Appendix One

What to do if you are processing personal data for IT admin purposes

Ensure that the data subjects know that the data is used for these purposes. It is advisable to do this at the time the data is collected as part of the standard data protection statement.

If you are using sensitive personal data, you must obtain the consent of the data subjects. It is advisable to do this at the time the data is collected.

Only take a full copy of the data if you need all the data.

Delete the data as soon as it is no longer needed.

If someone asks you not to use their data for these purposes, ensure that their details are either deleted from the dataset or changed so that they are no longer recognisable. Ensure that measures are in place to do this each time a new version of the data is uploaded.

If someone asks you to correct their data, then do so, either by amending the data in the existing dataset or by overwriting it with data from the real dataset. In any case, the amendments should be made within a reasonable time, normally within 10 working days of receiving the request.

If someone asks for a copy of the data, then you must provide it to them within 40 calendar days. Clearly mark the copy as test data.

Store test data with security measures at least as high as for the original data.

Clearly distinguish between the test data and the real data.

Do not use data about a real person for demonstration purposes unless you have the consent of that person.

Only allow access to test or training data by those staff who have to work on it for testing or training purposes.

Do not export data about identifiable individuals outside the European Economic Area unless:
- You have the express consent of all the data subjects
- You have set up a data controller-data processor relationship with the company to whom you are passing the data
- You have a contract with the overseas company which ensures that the rights of the data subjects are properly protected.