Information Security Guidance

1.0 Introduction

The purpose of this document is to set out the information security requirements for managing research data. By data we mean: data within the continuum from the raw, original state, through processing, analysis, synthesis, aggregation to publication of findings; data in all formats and forms (qualitative, quantitative, electronic, audio-visual, paper, artefacts etc.); associated information/documents (e.g. project proposal, protocols and templates, correspondence etc.) Unless otherwise specified, this is the meaning of our use of the word data in the rest of this guidance.

Appropriate information security ensures that data is protected against:

- the consequences of breaches of confidentiality, which can result in reputational damage, claims because of loss of intellectual property, damage to research subjects;
- failures of integrity (i.e. the accuracy and consistency of data), which can undermine research credibility;
- interruptions to the availability of data, which can impact on the research delivery;
- loss through accidental or malicious damage, modification or theft.

Documented data security risk assessments (see section 2.0) provide researchers with evidence of good practice as well as legal compliance and as such can be used to protect against claims of misconduct or intellectual property theft.

It should be noted that this is a guidance document and reference should also be made to Northumbria University's policies:

- *IT Systems Security Policy* (2010). http://www.northumbria.ac.uk/sd/central/its/it_strategy/regs/it_security/

- *Data protection policy*. http://www.northumbria.ac.uk/vc/leservteam/ndp/

2.0 Risk Assessment and Management

At the start of any research project risks assessments should highlight particular data security concerns. Research data should be risk assessed in the context of the damage (operational, financial or reputational) that would be caused to the research team and the University if the data was lost or compromised. Risk assessments will underpin project management decisions and the resources allocated to a project. As a result of risk assessments it may be necessary to purchase specialist services or security equipment such as fire proof safes or encryption software.

Risk assessment comprises comparing the impact of a risk against the likelihood of it happening. Risks can be mitigated or managed. Risks may be:

- tolerated (knowingly and objectively accepted);
- treated (managed by applying appropriate controls), e.g. access controls through encryption, locked filing cabinets, etc;
- transferred to other parties, e.g. sometimes if specialist services are not available in-house then it might be necessary to outsource a part of a project. However, it is not easy to transfer RDM risk;

- terminated (avoided - a decision could be taken to cease an activity associated with a particular risk).

A simple 3 x 3 grid illustrates this, although more sophisticated structures might be developed for complex projects.

An example of a risk assessment grid

Risk Assessment and Management				
Likelihood of risk		Low	Medium	High
	High	Treat	Treat Transfer	Treat Transfer Terminate
	Medium	Tolerate Treat	Treat	Treat Transfer
	Low	Tolerate Treat	Tolerate Treat	Treat
Impact of risk				

Note: the suggested risk management tactics within each box are indicative, not mandatory. The tactics you would apply within each box depends on your risk assessment for a specific research project

High Risk Data

The following types of data are examples of high risk that if compromised might have significant repercussions, not only on the research project itself but also the wider University or participants:

a. Details relating to identifiable individuals the contents of which, if compromised have the potential to cause damage or distress. In particular any personal data which has the potential to enable fraud or identity theft (including, but not limited to, personal contact details, date of birth, parents names and dates of birth) should be identified as part of the risk assessment. These details could reside in the data itself, in 'databases' of participant contact details and characteristics, or in signed consent forms.
b. Any set of data relating to an identifiable individual's sensitive personal details, i.e. health, disability, ethnicity, sex life, trade union membership, political or religious affiliations, or the commission or alleged commission of an offence. For example, raw, unanonymised data; 'databases' of participants contact details and characteristics.
c. Data concerning any vulnerable individual (e.g. children, adults with a learning disability or intellectual impairment). For example, the original audio records of interviews/focus groups with such individuals or their carers/relatives; unanonymised transcripts; their contact details; signed consent and assent forms.

d. Large data sets relating to 1,000 or more identifiable individuals. The loss of large amounts of such data, even if not of a sensitive or distressing nature, can still cause the research project/University reputational damage. For example, the raw, unanonymised data; 'databases' of participants contact details and characteristics; signed consent forms.
e. Research recommendations, before the decision was officially announced, that could result in substantial reorganisation or restructuring proposals that would have a significant impact on more than 50 individuals.
f. Data that, if compromised, would affect contracts with commercial or other partners, or confidentiality and non-disclosure agreements.
g. Information that would compromise patent applications.
h. Any data that is the result of an un-repeatable study. For example, a one-off event; an organisation that no longer exists; people who have died since the study.

It should be noted that the Information Commissioner (the ombudsman for data protection legislation) has the power to impose substantial fines if personal data is mismanaged. The Information Commissioner has established particular penalties for mishandling sensitive information (e.g. personal medical or political data) as well as data sets which relate to large numbers of individuals.

Medium Risk Items

The following types of data are examples of medium risk which if compromised would impact upon the individual research project or the wider University:

a. Any set of unpublished research data as this might undermine future publication.
b. Analysed data that would take significant time and effort to reconstruct.
c. Data relating to 10-50 identifiable people's personal and/or family lives, where the data is not distressing or sensitive and where the individuals are not in a vulnerable group.
d. Information relating to identifiable research participants (not in vulnerable groups), other than information within the public domain. For example 'database' of participants' characteristics.
e. Research recommendations, before the decision was officially announced, that could result in substantial reorganisation or restructuring proposals that would have a significant impact on 10-49 individuals.
f. Project documents and records that contain staffing or financial information (e.g. the project proposal, financial transactions, HR details) or contractual information (e.g. funder's offer letter, agreements with partners).

Low Risk Items

The following types of data are examples which carry a lower level of risk although this does not mitigate the requirement to ensure that it is appropriately managed:

a. Project documents and records, e.g. redacted project proposal, minutes of meetings, reports to funders, outputs and publications.
b. Internal project notes.
c. Research protocols and templates.
d. Analysed data that can be reconstructed without significant time or cost.

Any data with identified medium or high level risks would potentially require controls to minimise the level of risk exposure.

3.0 Storing Research Data Onsite

Wherever possible the University's secure drives should be used to store and access research data. By storing data centrally it will be automatically covered by Northumbria University's backup processes. For a single researcher project the secure drive would be their U:drive. For multi-researcher projects a shared drive should be set up (contact IT services to arrange this). This will ensure that all those in the research team can access and update information. It will limit the production of duplicate copes and thus minimise any confusion regarding version control.

At the start of a research project, IT access controls for the shared drive should be developed in conjunction with IT Services. IT Services will set up security controls around folders. IT Services will also allocate certain administration rights to named team members who can then manage some parts of this process within the team.

Where some documents within a folder are particularly sensitive then it may be necessary to place these within a sub-folder with further additional access controls. Individual documents stored on the central servers should not be password protected or encrypted without the authorisation of IT Services as this may cause problems with the University's regular backup processes. In addition, it is subsequently difficult to manage this data or change the security status.

It is very likely that a research project will also require the retention and management of some physical assets. Consideration should be given to securing these assets, e.g. file paper in locked filing cabinets/cupboards or safes. In addition, it may be advisable to scan hardcopy data (e.g. researcher notes, signed consent forms) into electronic format. These scanned copies can then be placed on the University drives and therefore be covered by back up procedures. However, for official documents (such as signed consent forms or agreements) the hardcopy versions must also be kept as the originals.

4.0 Storing And Accessing Research Data Offsite

If there is a requirement to access data offsite then wherever possible access should be via the University's remote access facilities (such as Desktop Anywhere). This avoids the risks of transporting the data on portable/mobile devices or media or using a third party host/the cloud.

If personal data is being used offsite then consideration should be given to partially or fully anonymising the information to obscure the identities of the individuals concerned.

If there is no option but to use portable/mobile devices or portable media for high and medium risk data, then it is important to use id/password access to devices and to consider purchasing encrypted memory sticks, encrypted laptops or using encryption software.

Personal equipment (such as home PCs or non-encrypted personal USB sticks) or third party hosting services (such as Google Docs) should not be used for high or medium risk material unless there is a strong justification which has been appropriately risk assessed.

When collecting data in the field, then by definition personal details and unique data will be collected via portable/mobile devices (e.g. audio recorders, mobile phones, lap tops, paper notebooks) and may need to be transferred to University facilities by portable media (USB sticks), the cloud (e.g. Dropbox) or even email, depending on the IT facilities available in the

field and their level of security (e.g. unsecured wifi, Internet cafes). Wherever possible encryption should be used. If unsecured data is the only available option, then careful guardianship of laptops/data sticks/paper notebooks etc. will be the only security method available to the researcher.

5.0 Data Transfer

When high or medium risk data is transferred to a third party then it would be necessary to consider whether a contract or agreement should be put in place to determine the ownership, management and constraint of use for the data. This is particularly important in the case of data containing personal information which is subject to data protection legislation.

High or medium risk data should not normally be transmitted by email. If this is strictly necessary it is recommended that consideration is given to encrypting the data.

Data can also be transferred/shared through third party hosts in the cloud. Dropbox is a cloud-based file sharing service commonly used by researchers to share data. Dropbox encrypts data for storage and transfer. In addition, users can encrypt their files themselves before sharing on Dropbox. With all such services, complete security cannot be guaranteed.

6.0 Encryption

Encryption should only be used for data that is being gathered, transferred, or shared externally. Files stored on the central University file servers should not be encrypted.

Where encryption is employed it is critical to put in place procedures to protect and maintain passwords and encryption keys. If these are lost then the encrypted data will be lost forever. This would have major ramifications for data being gathered offsite which had not yet been copied onto the central University servers.

It is strongly recommended that any encryption is agreed at the start of a research project in conjunction with IT Services. There are a number of software packages that enable data to be encrypted and these are evolving continually in response to attacks on encryption systems. Current examples include http://www.truecrypt.org/ and http://www.boxcryptor.com/.

7.0 Research Data Retention And Destruction

The RCUK Code of practice for research stipulates that data should normally be preserved and accessible for a period of ten years (RCUK 2011, p.8). Researchers should be aware that specific professional bodies and research councils may require a longer period of data retention. During this time it is important to actively manage data to ensure that it remains available and complete. UKDA has produced guidance on preferred file formats (refer to http://www.data-archive.ac.uk/create-manage/format/formats-table).

Physical data must be destroyed securely. This may include asking IT Services to wipe laptop memories or crush removable media. In addition any confidential or sensitive data in paper form should be securely shredded. (Note: recycling is not a secure form of destruction).

8.0 Northumbria University's Services

IT Services
Further information is available from IT Services http://www.northumbria.ac.uk/sd/central/its/.
Tel: *0191 227 4242*

Records Management Service
Further information is available from Duncan James, Records and Information Manager, http://www.northumbria.ac.uk/vc/leservteam/

9.0 Further References

Harvard University (2003) *Research data security*. http://www.security.harvard.edu/research-data-security-policy. Note: Harvard University is not subject to UK Data Protection requirements.

JISC. (2005) Information safety briefing paper. http://www.jisc.ac.uk/publications/briefingpapers/2005/pub_infosafetybp.aspx

Northumbria University. (Undated) *Good practice guide to using email policy*. Note: via intranet only

Northumbria University. (2003) *Interception and monitoring policy*. http://www.northumbria.ac.uk/sd/central/its/it_strategy/regs/it_security/

RCUK (2013) *Policy and guidelines on governance of good research conduct*. http://www.rcuk.ac.uk/Publications/researchers/grc/

UKRIO (UK Research Integrity Office) (2009) *Code of practice for research: Promoting good practice and preventing misconduct.* http://www.ukrio.org/what-we-do/code-of-practice-for-research/