Data Stewards and Data Management Principles

(LTAT.02.014)

Priit Adler
Research Fellow, University of Tartu
Elixir EE, BIIT

2021.04.06
Recap

Sensitive data
• Data protection law
• Informed consent
• Data safeguard
• DPIA
Data storage

storing, securing and backing up your research data.
Learning objectives

After completing this unit you will:

• Be aware of the options available to you to safely store your data.
• Recognise the importance of data backups.
• Be motivated to avoid your own data loss scenario.
• Understand password safety guidelines.
• Know how to encrypt and destroy sensitive data when required.

https://mantra.edina.ac.uk/storageandsecurity/
Introduction

“Your data are the life blood of your research. If you lose your data, recovery could be slow, costly or worse, it could be impossible.”

• ensure that all your research data, regardless of format, are stored securely, backed up and maintained regularly.

• estimate the volume of data required for your project at an early stage (while drawing up your data management plan).
  • consider including costs for data storage in funding proposals.

https://mantra.edina.ac.uk/storageandsecurity/
Where to store your data?

• Networked drives

• Personal computer / laptop

• External storage devices

https://mantra.edina.ac.uk/storageandsecurity/
Networked drives

- It is highly recommended that you store your research data on regularly backed up networked drives
  - [https://sisu.ut.ee/juhendid/failid-pilves](https://sisu.ut.ee/juhendid/failid-pilves)
  - Institute or research group managed network drives or file serves

This way you will ensure that your data will be:
- Stored in a single place and backed up regularly.
- Available to you as and when required.
- Stored securely minimising the risk of loss, theft or unauthorised use.

[https://mantra.edina.ac.uk/storageandsecurity/](https://mantra.edina.ac.uk/storageandsecurity/)
Personal computers and laptops

Personal computers (PCs) and laptops are convenient for storing your data while in use.

However, they should not be used for storing master copies of your data.

Local drives may fail or PCs and laptops may be lost or stolen leading to an inevitable loss of your data.

https://mantra.edina.ac.uk/storageandsecurity/
External storage drives

External storage devices such as hard drives, USB flash drives (also known as memory sticks, USB keyrings or pen drives), Compact Discs (CDs) and Digital Video Discs (DVDs)

• Their longevity is not guaranteed, especially if they are not stored correctly, for example CDs, DVDs and magnetic tapes degrade over the long term. They can be easily damaged, misplaced or lost.
• Errors writing to CDs and DVDs are fairly common.
• They may not be big enough for all the research data, so multiple disks or drives may be needed.
• They pose a security risk due to their portability.

https://mantra.edina.ac.uk/storageandsecurity/
When using external storage device

If you choose to use CDs, DVDs and USB flash drives (for example, for working data or extra backup copies), you should:

- Ensure your master copy is safe and is kept up to date on a networked drive.
- Choose high quality products from reputable manufacturers.
- Follow the instructions provided by the manufacturer for care and handling, including environmental conditions and labelling.
- Regularly check the media to make sure that they are not failing, and periodically 'refresh' the data (that is, copy to a new disk or new USB flash drive).
- Ensure that any private or confidential data is encrypted.

https://mantra.edina.ac.uk/storageandsecurity/
What do you think would be the advantages of storing your data on the central, or school networked drives?

https://mantra.edina.ac.uk/storageandsecurity/
Backups

keep at least 3 copies of your data on at least 2 different media, keeping storage devices in separate locations with at least 1 off-site

https://mantra.edina.ac.uk/storageandsecurity/
Keeping backups is probably the most important data management task.

Real risk of losing data through hard drive failure or accidental deletion.

Keep at least 3 copies of your data on at least 2 different media, keeping storage devices in separate locations with at least 1 off-site.

You should also have a policy for maintaining regular backups.

The 3-2-1-Rule

3. Different copies of data
   - orig
   - .vbk
   - .vbk

2. Different media
   - Drive
   - Disk

1. External medium
   - Cloud

https://www.backupbros.com/2020/03/19/who-takes-care-of-your-tapes-if-you-are-not-allowed-to-go-to-your-data-center/

https://mantra.edina.ac.uk/storageandsecurity/
Backup strategy

• How will you back up your data?
• Will all data, or only amended data, be backed up?
  • (A backup of amended data is known as an "incremental backup", while a backup of all data is known as a "full backup").
• How often will full and incremental backups be made?
• How long will backups be stored?
  • It is important to ascertain the back-up schedule and retention policies of any centralised backup services (e.g. under the Grandfather-Father-Son rotation scheme, files may be available for two to three months before the space is over-written).
• How much hard drive space or how many CDs/DVDs will be required to maintain this backup schedule?
• How will you keep track of different versions of data, especially when backing-up to multiple devices?
  • If using versioning software, which software will you use (e.g., Tortoise, Subversion, Git)?
• If the data are sensitive, how will they be stored securely and appropriately, and how will you manage the destruction of identifying data if required (e.g. at the end of your research)?
• What backup services are available that meet these needs and, if none, what alternatives are available?

https://mantra.edina.ac.uk/storageandsecurity/
Online backups

Remote, online or managed backup services provide users with an online system for storing and backing up computer files.

Typically online backup services:
- Allow users to store and synchronise data files online and between computers.
- Employ cloud computing storage facilities (e.g. Amazon S3).
- Provide the first few gigabytes free and users pay for more facilities, including space.

Online backups

- OneDrive
- tresorit
- box
- iCloud
- FilesAnywhere
- memopal
Online backup: advantages

• No user intervention is required (changing tapes, labeling CDs, performing manual tasks)
• Remote backup maintains data off-site
• Most provide versioning and encryption
• They are multi-platform
Online backup: disadvantages

• Physically located outside the European Economic Area (EEA)?
• Restoration of data may be slow (dependent upon network bandwidth)
• Stored data may not be entirely private (if unencrypted)
• Service provider may go out of business
• Protracted intellectual property rights/copyright/data protection licences
• Vendor lock-in, i.e. vendor’s proprietary formats may make migration to another vendor complex and expensive

https://mantra.edina.ac.uk/storageandsecurity/
Data security

Data security means ensuring that research data are kept safe from corruption and that access is suitably controlled.

https://mantra.edina.ac.uk/storageandsecurity/
Data security

It is important to consider the security of your data to prevent:

- Accidental or malicious damage/modification to data.
- Theft of valuable data.
- Breach of confidentiality agreements and privacy laws.
- Premature release of data, which can void intellectual property claims.
- Release before data have been checked for accuracy and authenticity.

https://mantra.edina.ac.uk/storageandsecurity/
Data security: DMP

You need to consider the following questions for securing your research data:

• How will you manage access arrangements and data security?
• How will you enforce permissions, restrictions and embargoes?
• Other security issues such as
  • sensitive data,
  • off-network storage,
  • storage on mobile devices (laptops, smartphones, flash drives, etc),
  • policy on making copies of data,
  ... etc. where relevant.

https://mantra.edina.ac.uk/storageandsecurity/
Data security

You should always have up to date anti-virus software installed on your office and home computers.

If you have sensitive data that are covered by privacy laws or confidentiality agreements,

   it is best to store them on a computer that is not connected to any network.
   You can also consider encrypting your data or your PC.

A computer that is not connected to a network is still vulnerable to theft and malicious damage/modification to data.

https://mantra.edina.ac.uk/storageandsecurity/
Data security

Highly sensitive data, regardless of physical storage medium (laptop, USB flash drives, CDs/DVDs), should be stored in a locked room or safe when not in use.

Such data should not be stored or transmitted without encryption.

https://mantra.edina.ac.uk/storageandsecurity/
Usernames and passwords

Two issues which are often overlooked but are worth highlighting are 'usernames' and 'passwords'. They are so common users often forget they are still a key part of security on most systems.

If possible:
- Never use your username as your e-mail address e.g. fbloggs27@staffmail.ed.ac.uk.
- Instead use an alias e.g.: Fred.Bloggs@ed.ac.uk.
- Do not write passwords on Post-Its or say them aloud as you type.
- Do not log in to secure spaces on untrusted computers or networks (e.g. internet cafés).
- Do not use the obvious (car reg., phone no., pet’s name)
- Do not use any dictionary words (including foreign) on their own.
- Never use the same password for a cloud service, that you use for a University service.

https://mantra.edina.ac.uk/storageandsecurity/
Encryption

Encryption is the process of converting data into an unreadable code. You must have access to a password or a secret encryption key to be able to read an encrypted file.

Encrypting your data will help ensure your data remain safe from disclosure in the event that a laptop or other portable device such as a USB flash drive/memory stick ends up lost or stolen.

BitLocker (Windows) and FileVault (Mac)

https://mantra.edina.ac.uk/storageandsecurity/
Encryption

Please note the following considerations before encrypting your data:

- Data encryption is not a substitute for other information protection controls.
- Data encryption is reliant on the creation of a strong password.
- Encrypted data cannot be recovered in the event of a failure.
- If the encryption key is lost, the disk image gets corrupted, or the hard disk fails, any encrypted data will be lost.
- Establish a reliable and secure backup procedure for the data and any related passwords.

https://mantra.edina.ac.uk/storageandsecurity/
Data erasure

Once research on data which includes sensitive data no longer requires that the identifiable portions of the data be retained, they should be destroyed and future research be done with de-identified or anonymised data. This applies to paper records, as well as electronic records.

https://mantra.edina.ac.uk/storageandsecurity/
Data erasure

File deletion is not enough to ensure that sensitive electronic data are completely removed from a computer system. File deletion removes only the pointers to the disk sectors in which the data reside. Deleted files can be recovered using commonly available software tools.

To ensure the complete destruction of sensitive data, the main options are:

• data erasure (aka data clearing or data wiping) i.e. removing all data while leaving the hard drive or other storage medium still operable
  • current techniques may not, however, be completely successful on solid-state drives and USB flash drives

• degaussing, which disturbs the magnetic alignment of magnetic storage media
  • in many cases makes newer media (such as hard drives and tapes) unusable

• physical destruction, through disintegration, shredding, pulverizing or incineration.

https://mantra.edina.ac.uk/storageandsecurity/
Discussion and questions
Homework

In Estonian

Read through https://sisu.ut.ee/minuandmed
Answer all “Mõtisklusülesanne” in writing!
indicate chapter and question
Self-test is fun and you should do that, but not compulsory for the homework