Data Security

Dr. Brian Kleiner
October 12, 2021

FORS webinar series

1. **Introduction to data management planning**
2. **Data protection**
3. **Data security**
4. **Data sharing** / November 4th

FORS

Outline:

1. Introduction
2. General framework
3. Technical measures
4. Organisational measures
5. Putting it all together

Introduction

Definition

Data security is the technical and organizational means of ensuring that research data are kept safe from corruption, theft, and loss, and improper access.

Things can go wrong!

Some data security risks

- Hardware failure; software malfunction; degradation of storage media
- Accidental or malicious damage/modification to data
- Theft of data and inappropriate access more generally
- Natural disaster (e.g., fire, flood)

Possible consequences of data security breaches

- Harm to research participants
- Harm to researchers, their reputations, and their work
- Harm to institutions

General framework

General framework: legal obligations

Article 7 of FADP: “Personal data must be protected against unauthorized processing through adequate technical and organizational measures.”

The duty to secure data is one of the most important aspects of data protection.

As data controller, you must take all appropriate measures to prevent unlawful access or accidental loss, in relation to objective risks to participants.

General framework: organisation and controlled access

Data security is also about protecting your work and ensuring key outcomes for your research.

Through proper organization and rules for access, you should provide for the safety and consistency of your data.

Best practices in data management are needed, throughout the research life cycle.

Criteria for taking “appropriate” measures

- The **type** of data
- Level of **sensitivity** of data
- The amount of **personal information** collected
- The **risks in case of security leaks**
- The **technical state of the art**

Some general principles for security

- **Only what is needed** - You should only collect what is needed for your research project
- **Defense in depth** - It is better to have multiple layers of security
- **Weak link** - A system is only as secure as its weakest part. There is no point in putting a lot of effort into one part if you neglect another
- **Simplicity** - If security prevents team members from working, they will bypass security

Technical measures

Technical measures

- Storage
- Password protection
- Encryption
- Backup
- Data disposal

Technical measures

- **Storage**
- Password protection
- Encryption
- Backup
- Data disposal

About storage

Storage is the capture and retention of digital information on a storage-media during a research project.

Almost every research project needs a storage solution.

Storage can be digital or physical. You may need physical storage for paper documents or for digital devices, for example, in a room or cabinet, but projects with no digital content are rare these days.

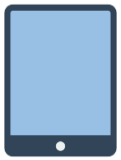
Relevant factors

A storage solution must suit the project, according to:

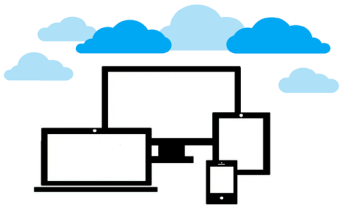
- Data volume
- Data sensitivity
- Team composition and roles
- Institutional setting(s)
- Available infrastructure
- Project duration
- Project resources
- Fieldwork setting(s)
- Non-digital objects
- Need for transmission of files

Storage solutions: 3 main types

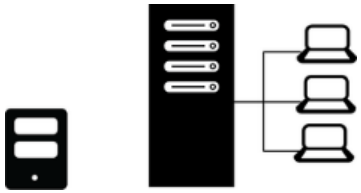
Local devices



Cloud storage



Network drives



Storage solutions

Local storage

Advantages	Disadvantages
• Usually allow easy «portable» transport of data and files without transmitting them over the internet • Full control over files • May be easier to protect against unauthorised access	• Easily lost, damaged, or stolen • Not robust for long-term storage • Possible quality control issues due to versionning • Only the person who has access to the device can access the data and files

Personal devices and physical security

If you need to use local storage:

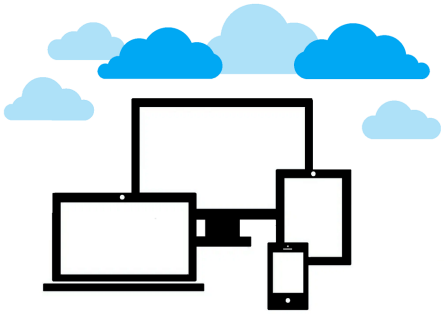
- Use a password-locked screensaver and timeout lock
- Install and maintain antivirus software
- Use a firewall
- Keep operating system and all software up to date
- Don’t install or run programs from untrusted sources

For all local storage devices:

- Keep all computers and devices in secured spaces
- Limit physical movement

Storage solutions

Cloud storage



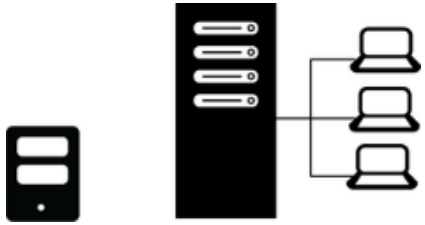
Advantages	Disadvantages
• Shared files from any computer • Automatic backups • Often automatic version control	• Not all cloud services are secure. May not be suitable for sensitive data. • Insufficient control over where the data are stored and how often they are backed up. • Free services by commercial providers may claim rights to use content you manage. • Data can be lost if your account is suspended or accidently deleted or if the provider goes out of business.

Examples of cloud storage solutions

- SWITCHdrive
- Dropbox
- iCloud
- Google Drive
- Microsoft One Drive
- IDrive
- Mega
- Box
- pCloud

Storage solutions

Network drives



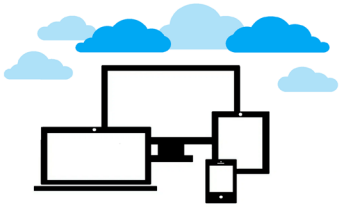
Advantages	Disadvantages
• Data and files are centrally stored • Shared access, remote access for everyone involved in the project possible • Backups can be centrally managed and automated	• Higher security precautions are required to prevent unauthorized access and the accidental deletion or manipulation of data and files. • Access for external project partners can be difficult or impossible

Precautions for personal data

Local storage

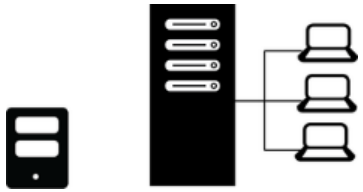
Use in combination with encryption and strong password protection

Cloud storage



Encrypt all personal data before uploading them to the cloud.

Network drives



Use in combination with a suitable security strategy to protect against unauthorised access

Ideal ranking of solutions, all things considered

1) Use an institutional network solution, if available
2) Use a cloud solution if there are no sensitive or personal data
3) Use a local solution, with good backup, versionning, and data transmission strategies

Security during fieldwork

Usually local devices are used for fieldwork, especially where internet access is limited. Special precautions must be taken to protect personal data, especially in high-risk environments:

- Plan in advance how to store and safeguard your data
- Encryption on all devices is recommended
- Backup new or changed files as soon as possible
- Upload files to cloud or network as soon as internet access is available
- Only gather information that is strictly needed for your research questions

Technical measures

- Storage
- **Password protection**
- Encryption
- Backup
- Data disposal

Password protection

Passwords provide the first line of defense in data security.

Passwords are a means by which a user proves that he or she is authorized to use a computing device.

Stronger passwords mean better protection from unauthorized access, hacking, and malicious software.

Choosing passwords

There are several strategies when it comes to choosing a password:

- Passwords should contain at least ten characters and have a combination of characters such as commas, percent signs, and parentheses, as well as upper-case and lower-case letters and numbers.
- A useful way is to make them up of four randomly chosen and altered words, e.g. C.rr3ctHorseBatteryStaple
- You may also use pass phrases, e.g. MyMotherM$kesTheB*stCakes.

Password tools

It may be useful to use a tool for managing your passwords, such as KeePass (windows/Linux) or MacPass (MacOS). Other good solutions include 1Password and Bitwarden.

Technical measures

- Storage
- Password protection
- **Encryption**
- Backup
- Data disposal

Encrypting data

Encryption is the process of encoding or scrambling digital information in such a way that only authorized parties can view it. It is especially useful when transmitting personal data.

Encryption allows you to add a layer of protection where sensitive data are insufficiently safe otherwise.

Encrypting data

For best practice:

- Encrypt sensitive data, especially before transmitting them online, uploading them to the cloud, or transporting them on portable devices.
- Make sure that the key can be accessed by everyone who needs to access it (but only those people).
- Ensure that you do not lose the key to decrypt your files.

Encrypting data

Most encryption software allows you to encrypt at different levels:

- Device-level (whole-disk)
- Folder-level
- File-level

Encryption softwares include BitLocker, FileVault2, Pretty Good Privacy, VeraCrypt, Axcrypt, SafeHouse.

Technical measures

- Storage
- Password protection
- Encryption
- **Backup**
- Data disposal

Back-up

Back-up is crucial since your data are always at risk of loss - a system failure can wipe out all your data, corruption can render your data useless, and an error can lead to permanent deletion.

The only way to tackle data loss incidents is by putting a solid back-up strategy in place.

Your strategy should be established at the beginning of your project and evaluated periodically.

Key questions for a back-up strategy

1. What backup services or tools are available that meet your needs? Does your institution have a backup solution?
2. Will all data or only changed data be backed up? How often will full, differential, or incremental backups be made?
3. How many back-up copies will you have? Where will backups be stored?
4. How long will backups be kept and how will they be destroyed?
5. How much space will be required to maintain the backup schedule?
6. Who will be responsible for doing backups, if back-ups are not automatic?

See CESSDA Data Management Expert Guide's 10-steps for a backup strategy.

Technical measures

- Storage
- Password protection
- Encryption
- Backup
- **Data disposal**

Disposing of data

By law, you should delete any personal information after you no longer need it for your research project.

Also, keeping unnecessary files adds complexity and increases risk of error.

Data disposal (i.e., “erasure”, “sanitizing”, or “wiping”) is an important measure for reducing risk.

Warning about data disposal

Merely hitting the “delete” button is not enough. Deleted files can be recovered.

Even installing a new operating system on top of your old one does not fix the problem. You can still recover many of your previous system's files.

Data erasure

Data erasure is a software-based method of overwriting the data that aims to completely destroy all electronic data on a hard disk drive or other digital media by using zeros and ones to overwrite data onto all sectors of a device.

There are two options for secure disposal of confidential data:

- The physical destruction of the storage medium
- The use of software for secure erasing (e.g. CCleaner, Eraser, WipeFile, File Shredder, TweakNow SecureDelete)

Organisational measures

Organisational measures

- **Controlling access**
- Data transmission
- File organisation

Controlling access

To prevent unauthorized access and possible changes to your data, data security measures need to be implemented. These could include:

- establishing project rules on who can access which files
- limiting access to only what is needed to carry out the work
- de-identify data as soon as possible

Organisational measures

- Controlling access
- **Data transmission**
- File organisation

Data transmission

Data transmission is the process of sending digital or analog data over a communication medium to one or more computing, network, communication or electronic devices.

In a research perspective, it is when you have to “move” files outside of a common storage environment. This could be sending a file within an email, downloading the file to your own computer’s hard drive, or uploading a file to the cloud.

Data transmission and sharing

Every transmission involves a security risk, but sometimes data transmission is necessary – take care to minimize risks when transmitting personal information.

Unsafe methods include:

- email without encryption
- uploading unencrypted data to a cloud service
- mailing or hand-delivering unencrypted media devices

Data transmission

Safe methods include:

- Emailing an encrypted file, and sharing the password separately and securely
- Uploading an encrypted file to the cloud
- Mailing encrypted files loaded onto encrypted devices
- Survey software with encryption features
- Secure Shell File Transfer Protocol (SFTP)

Organisational measures

- Controlling access
- Data transmission
- **File organisation**

File organisation

Proper file organization is a good way to manage data access and enhance security

- In a shared project space, create a common folder with files accessible to all team members
- Data files containing personal information should be stored in password protected folders, with rights-only access
- Keep files with personal data separate from data files (with corresponding unique identifiers)
- For certain storage environments you can distinguish “edit”, “suggest”, and “read-only” rights (e.g., Google Docs) – make use of this where possible, according to need and level of sensitivity
- Final master data files should be stored separately as “read-only”

Putting it all together

Create a security plan for your project

- Establish the storage environment
- Identify additional needed devices (e.g., for fieldwork) and how these will link to the storage environment
- Create rules for encryption, password protection, backup, and data disposal
- Create rules for access to files, data transmission, and file organisation
- Keep communication going: make sure everyone on the team is aware of the risks and is adequately involved in discussons of data security issues and measures in place

Main sources:

CESSDA Expert Tour Guide, chapter four: https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide

DMLaw Tool: https://www.usi.ch/en/feeds/15625

Questions?