Management of sensitive research data

Vilmantas Giedraitis, PhD

2022-11-29

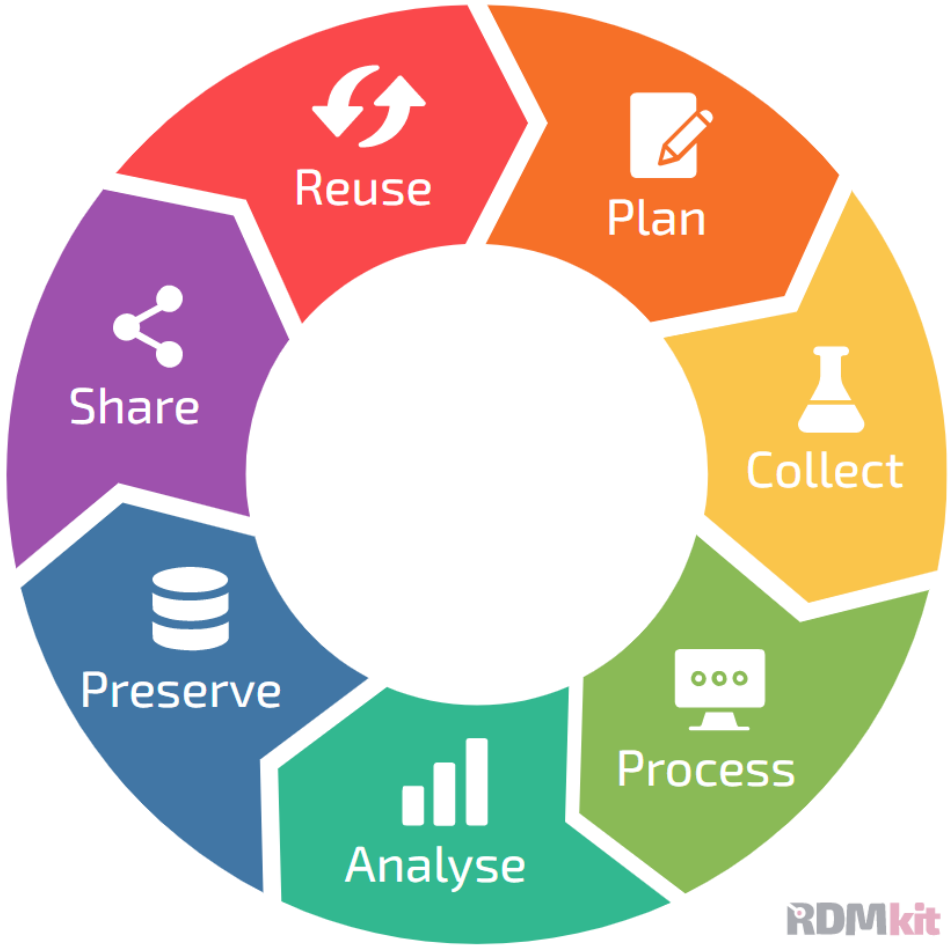
Department of Public Health and Caring Sciences

- Caring Sciences
- Clinical Nutrition and Metabolism
- Disability and Habilitation
- Family Medicine
- Geriatrics
- Health Equity and Working Life / HEAL
- Health Services Research
- Lifestyle and Rehabilitation in long term illness
- Research Ethics and Bioethics
- Social medicine / CHAP
- Speech and language pathology

My own projects with sensitive personal data

- Uppsala-Dalarna Dementia and Gait study (UDDGait)
  - Clinical data and video recordings from patients with dementia and controls
- Uppsala Longitudinal Study of Adult Men (ULSAM)
  - Population cohort study started 1970 and still ongoing
  - More then 450 published research articles and at least 40 doctoral theses
  - More then 20 researchers in several countries are using ULSAM data
- Gene mutation analysis in patients with neurodegenerative diseases
  - Vulnerable patient group and especially sensitive results
- Gene association and biomarker analysis in Alzheimer's disease
  - Project involving several EU countries

Research data life cycle



Personal data and sensitive data

Personal data: name, social security number, location information, IP or e-mail address, photos, video, audio recordings, telephone number.

Sensitive personal data: health, ethnicity, political opinion, religious and philosophical beliefs, sexual orientation, individual sexual life, trade union membership, biometrics, genetics.

Even pseudonymized or encrypted data are sensitive if code and encryption key is available and it is possible to identify individuals.

https://www.imy.se/verksamhet/dataskydd/det-har-galler-enligt-gdpr/introduktion-till-gdpr/personuppgifter/

Information classification

CIA triad (KRT värde): **C**onfidentiality (**K**onfidentialitet)
Integrity (**R**iktighet)
Availability (**T**illgänglighet)

Value	Possible consequences of shortcomings
0	No consequences
1	Could cause discomfort or limited harm / financial loss to individuals, the University or a third party.
2	Could cause major discomfort or extensive injury / financial loss to individuals, the University or a third party.
3	Could cause harm to life or health for individuals or cause major discomfort/injury or financial loss to a large number of people, or very serious harm to the University or a third party.

https://www.du.se/sv/medarbetarwebb/stod-och-service/informationsforvaltning-juridik-ochdataskydd/informationssakerhet/
https://www.msb.se/RibData/Filer/pdf/25602.pdf

Example for information classification

	Konfidentialitet Confidentiality	Riktighet Integrity	Tillgänglighet Availability
Pseudonymized data from surveys, clinical measurements, genetic data from biobank and cohort studies	2	2	1
Video and audio recordings, image data	3	2	1
Pseudonymization key	3	3	1

Application for the processing of personal data to DU Data Protection Officer (Data Protection Officer)

- All structured processing of personal data have to be reported to data protection officer
- *Application is necessary even if personal data are not sensitive*
- *Old projects also needs to be reported*
- *Smaller projects can be grouped in one report*

Data Protection Officer
(Dataskyddsombud)
Yvonne Arvidsson
E-mail: dataskydd@du.se

Storage of sensitive data

Information type / Storage place	Temporal or of limited importance	Critical for the institution	Secret	Sensitive personal data	Personal data
Institutional systems (e.g. Learn, Black board, LADOK, Primula, Agresso Diary, DIVA)					
Central storage (H: or L:, etc.)					
Own computer (C: or equivalent)					
Portable media (Smart phones, tablets, external hard drive, USB, etc.)					
University cloud service (Microsoft Office 365, Teams and One Drive)					
External cloud service (Dropbox, iCloud, Google, Drive, Evernote, etc.)					

- Green boxes show the appropriate storage location
- Yellow can be used for temporary storage if appropriate safety precautions are taken
- Red should not be used

https://www.du.se/sv/medarbetarwebb/stod-och-service/informationsforvaltning-juridik-ochdataskydd/informationssakerhet/

Storage of pseudonymized sensitive data

- On paper in a secure cabinet – several, but not all need to know where the key is.
- Own computer – must be encrypted.
- External HD or USB memory – should be encrypted and / or locked.
- DU central storage – ok to store pseudonymized data from ongoing projects
- The DU archive - stores large amounts of research data in paper format and even digital copies.

Handling of code lists

- On the paper in a secure cabinet (separately from data!) – several, but not all need to know where the key is.
- External HD or USB memory – must be encrypted and locked.
- Own computer – should be avoided!
- DU central storage – code key must be encrypted *and stored on a separate storage area*.

Data encryption

- BitLocker – Installed on Windows 10. Must be turned on if sensitive data is handled on the computer. If computer is stolen no one can access data. Easy to use for the encryption of external HD and USB memory.

- FileVault – Available on Mac OS. Must be turned on if sensitive data is handled on computer. If computer is stolen no one can access data.

- VeraCrypt – free program available for Windows, Mac OS and Linux. Creates encrypted file containers that are easy to copy or backup.

- 7-zip (or any other file archiver) – free program available for Windows, Mac OS and Linux. Compresses and encrypts files.

https://mp.uu.se/sv/web/info/stod/sakerhet/it-sakerhet/kryptering
https://mp.uu.se/sv/web/info/stod/sakerhet/informationssakerhet#IT-security

Data accessibility (for all data types)

- Several but not all researchers in the group should know where data, code lists and encryption keys can be found.
- If many researchers use the same cloud-based storage, there is a risk of intentional or unintentional data deletion. Larger research groups should use separate storage areas for projects or subgroups.
- Data should be saved on at least two different kind of media (server, portable hard drive, cloud)

  … Hard drives and especially USB sticks fails often

  … Cloud-based solutions have limited backup time
- Robust backups need to be automated.

Sharing of sensitive data

- Access to sensitive data (including pseudonymized) must be limited. Share only data needed for the project.
- Do not use email to share sensitive data.
- USB memory can be used, but should be encrypted.
- Cloud services can be used to share pseudonymized data, but data must be encrypted and sharing time limited, so that the link becomes invalid after a certain time period.
- The password must be sent separately.

Data Transfer Agreement (DTA) and Data Processing Agreement

- DTA is always needed if raw data is shared with other universities or companies.
- DTA needed even if data comes from other universities.
- DU Legal Affairs Division can advice in legal matters.
- DTA can be avoided if only data summaries are shared.

- *Is DTA needed if data is analyzed at another department at DU?*
- *Is DTA always needed for joint projects with other Swedish universities?*
- *Is a Data Processing Agreement needed for students?*

De-identification and destruction of the data

- As a general rule primary material must be preserved for unlimited time. If necessary, destruction of the data can be undertaken 10 years after the project completion.
- Data can be de-identified after 10 years or based on informed consent, but then consent must be destroyed.
- The DU archive – stores large amounts of research data in paper format and also digital copies.

https://www.du.se/en/medarbetarwebb/forska-och-utbilda/research-support/research-data-and-personal-data/research-data---support-and-guidance/

Some basic rules for information security

- Keep mobile devices safe and lock the computer when you leave your desk.
- Activate the PIN code on mobile phone.
- Keep software up to date.
- Regularly backup data.
- Do not open SMS/MMS or click on links sent from unknown sources.
- Limit your browsing to pages that use encryption (the address starts with HTTPS).
- Pay attention to warnings that you connection is insecure.
- Do not save the passwords in the browser.

When traveling abroad…

- Backup data before travel.
- Use 4G/5G instead of Wi-Fi.
- Be careful when connecting to open public networks. Use the university's VPN.
- Deactivate services you don't need (location services, Bluetooth, etc.).
- Use your own charger. In some hotels, chargers may be available in the room, but hey might be rigged to plant malware.

Web recourses for data security

- Information security (Swedish) – https://www.du.se/sv/medarbetarwebb/stod-och-service/informationsforvaltning-juridik-ochdataskydd/informationssakerhet/
- Processing of personal data (Swedish) – https://www.du.se/sv/hjalp/personuppgifter/
- Support and guidance for research data – https://www.du.se/en/medarbetarwebb/forska-och-utbilda/research-support/research-data-and-personal-data/research-data---support-and-guidance/
- Data archiving (Swedish) – https://www.du.se/sv/medarbetarwebb/stod-och-service/informationsforvaltning-juridik-ochdataskydd/arkiv-och-e-arkiv/
- General Data Protection Regulation (GDPR) – https://www.imy.se/en/organisations/data-protection/
- Swedish National Data Service (SND) – https://snd.gu.se/en/manage-data/plan/research-material-with-personal-data