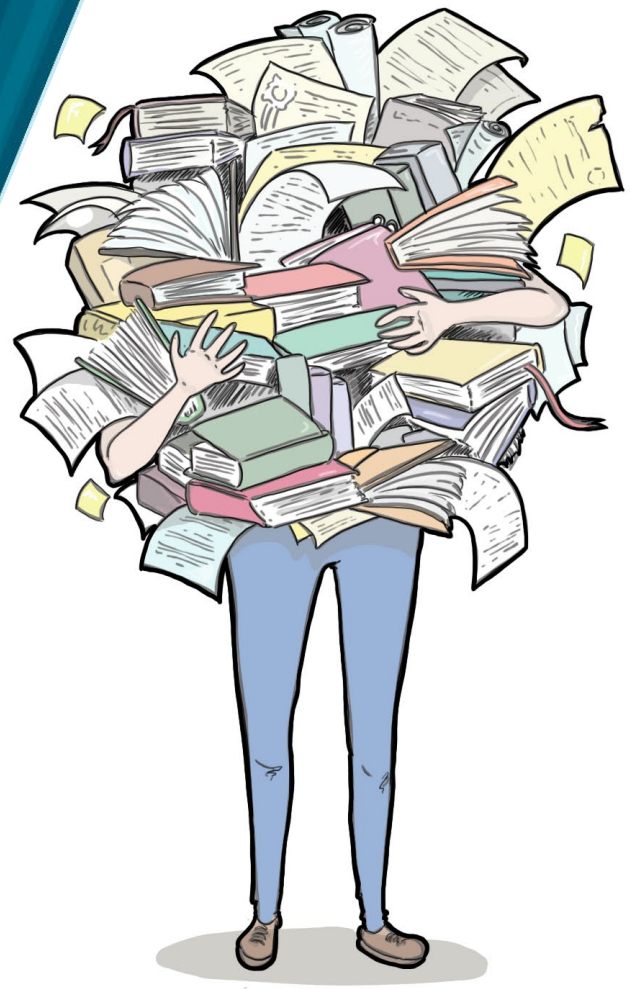
UiT Norges arktiske universitet

How to manage sensitive research data

13.04.2022

Aili Sarre and Huw Grange
Research support at the university library

Margit Ramberg
Legal advisor at the university administration



Agenda

- Definitions
- Responsibility
- The planning phase
- QA
- The active phase
- Archiving/Publishing data
- QA

What is sensitive research data?

Research data:

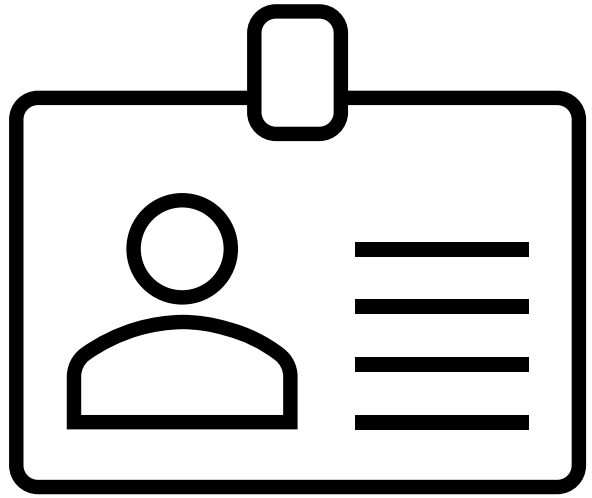
Research data encompass all registrations, records and reports that are generated or occur during research, and that are considered as being of scientific interest and/or having scientific potential. These may include, but are not limited to, numbers, texts, source codes, images, films and sound recordings.

PRINCIPLES AND GUIDELINES FOR MANAGEMENT OF RESEARCH DATA AT UIT §2 (2021)

Sensitive research data is data that needs protection due to legal, security or ethical reasons

- Ecological data – e.g. location data on vulnerable species
- Commercial interest
- Potential misuse of information for unethical purposes
- Research which endangers the researchers
- Personal data – Special categories of personal information, and other types of sensitive information about individuals.
- Data on groups of people

Personal Information



All information that directly or indirectly can identify an individual person.

All processing of personal information is regulated by **The Personal Data Protection Act** and EUs **General Data Protection Regulation (GDPR)**

Anonymized information are not considered personal information and the GDPR does not apply. How to avoid collecting personal data in your project.

Personal data are not necessarily sensitive, but the laws still apply.

Direct identifiers:
Name, ID number, photo, biometrics (fingerprints, voice)

Indirect identifiers: Information which in combination can lead to identification. Address, age, occupation.

Take note: the different pieces can originate from different sources. Read more at Datatilsynet

Special categories data (GDPR Art.9)

- In the previous national Personal Data Act referred to as «sensitive data»

- Race or ethnic origin
- Political views
- Trade union membership
- Religious, spiritual or philosophical beliefs
- Data concerning sexual orientation or sexual life
- Health data
- Biometric or genetic data
- (Data relating to criminal convictions and offences (Art.10))

Resonsibility for data that requires protection

Researchers have a statutory duty to exert caution to ensure that research takes place in accordance with recognized norms of research ethics. (Research Ethics Act §4)

The institutions are legally responsible for ensuring that the research carried out under their purview is conducted in accordance with recognized norms of research ethics. (Research Ethics Act §5)

Responsibility for data that requires protection

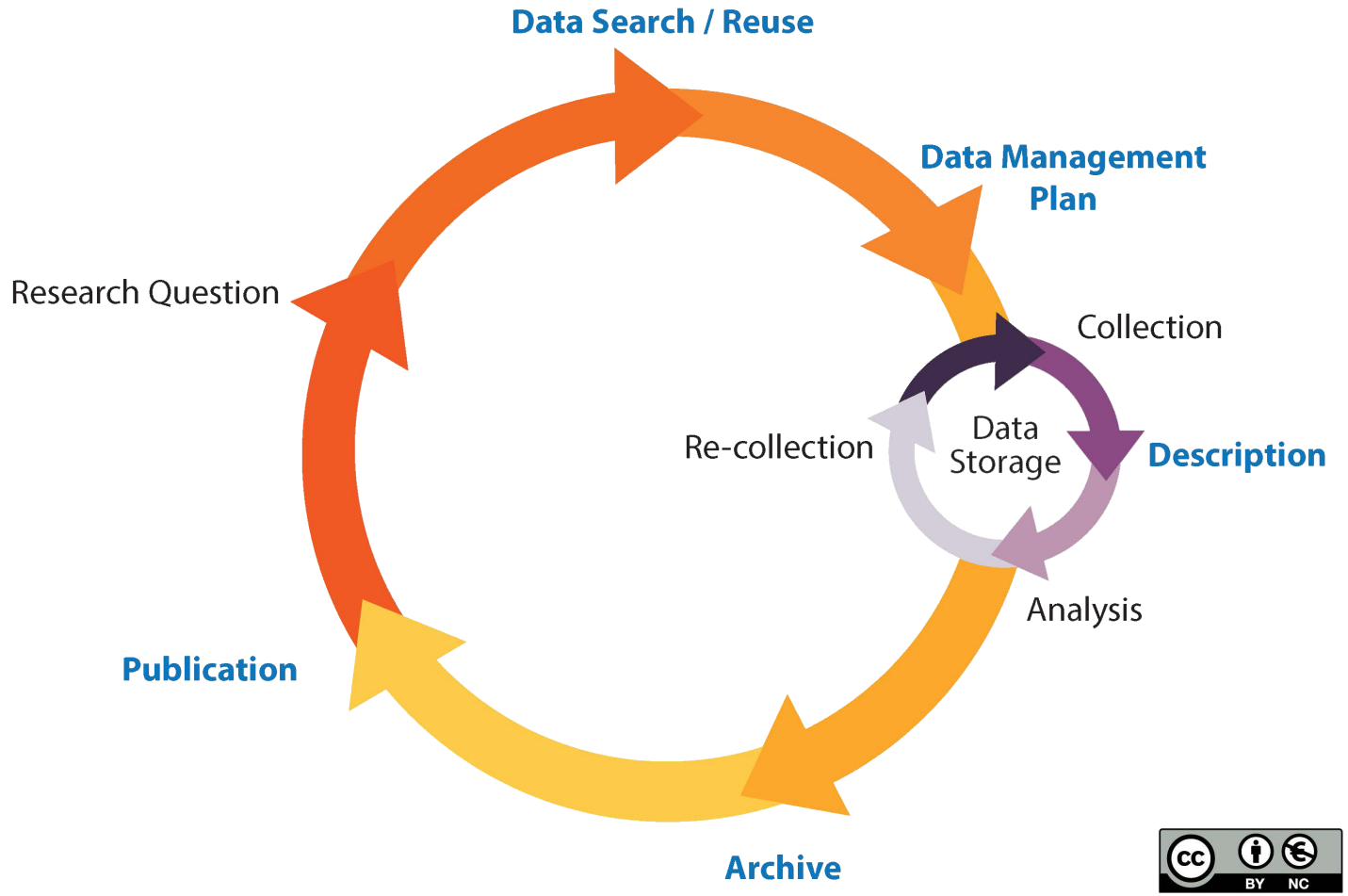
As employees at UiT you are responsible that research data is managed according to best practices, institutional requirements and according to laws and ethical guidelines.

This requires:

- Knowledge of research ethical obligations to:
  - the research community
  - collaborators
  - individuals and/or groups of people
  - Society, nature and the environment
- Identification of potential ethical issues throughout the project period.
- Maintain care and attention to all stakeholders before, during and after the project is finished.

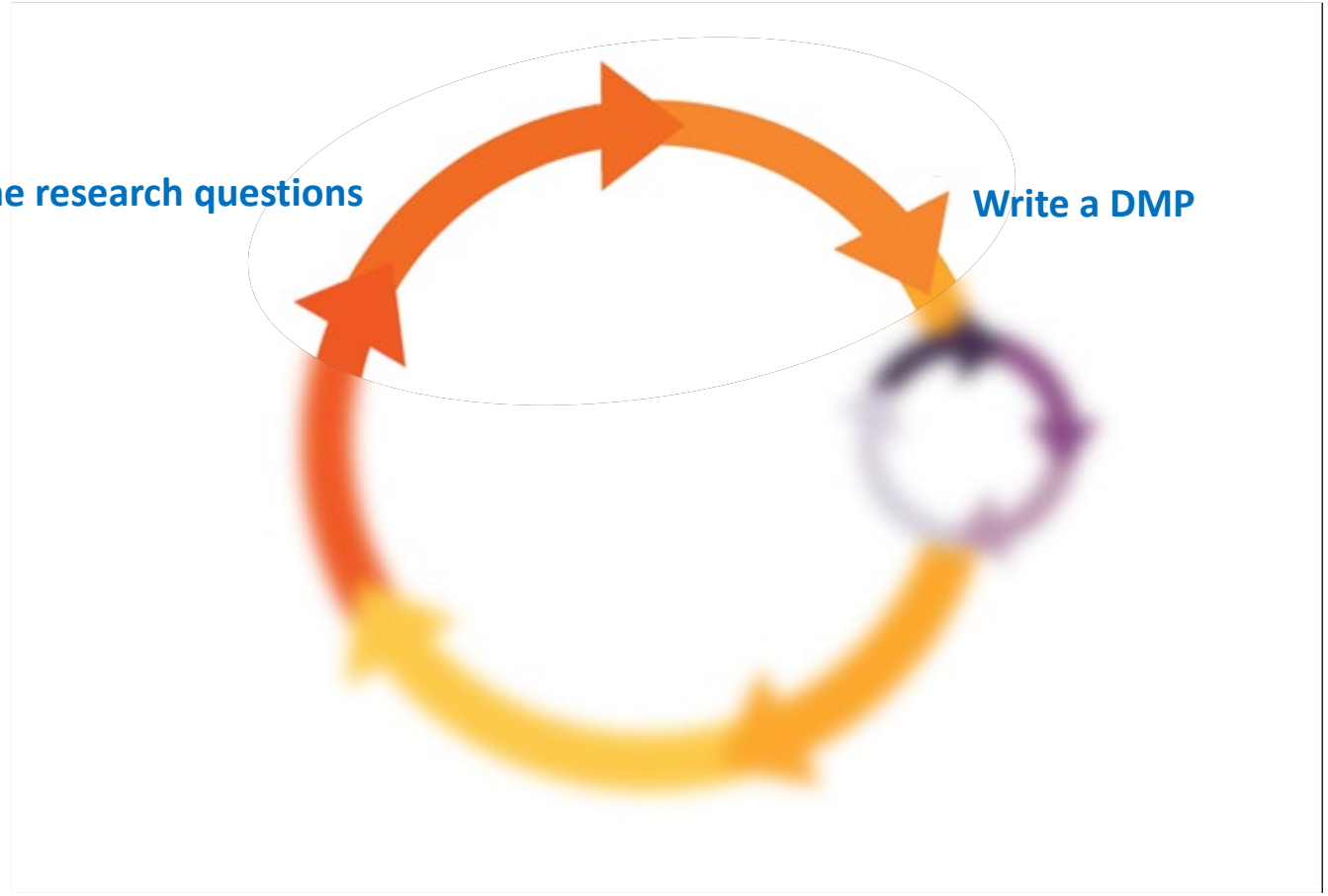
The life cycle of research data

- Planning
- Well-defined roles and responsibilities
- Communication
- Training



Adapted original source:
The University of California, Santa Cruz, Data Management LibGuide, Research Data Management Lifecycle, diagram, viewed May 2, 2016 at <http://guides.library.ucsc.edu/datamanagement>

The planning phase



- Overview of amount and type of data
- Assess sample selection vs. population
- Relevant laws and guidelines
- Ethical reflection or impact assessment (DPIA)
- Notify NSD in case of personal data
- Pre-approvment by REK in case of health research
- Clarify roles, rights and responsibilities.
- Data management plan (DMP)

Relevant laws, guidelines and routines

Laws

- The Research Ethics Act (Lov om organisering av forskningsetisk arbeid)
- The Personal Data Act and GDPR (Personopplysningsloven og personvernforordninga)
- The Health Research Act (Helseforskningsloven)
- Regulations on population-based health research (Forskrift om befolkningsbaserte helseundersøkelser)
- The Copyright Act (Åndsverksloven)
- The Security Act (Lov om nasjonal sikkerhet)

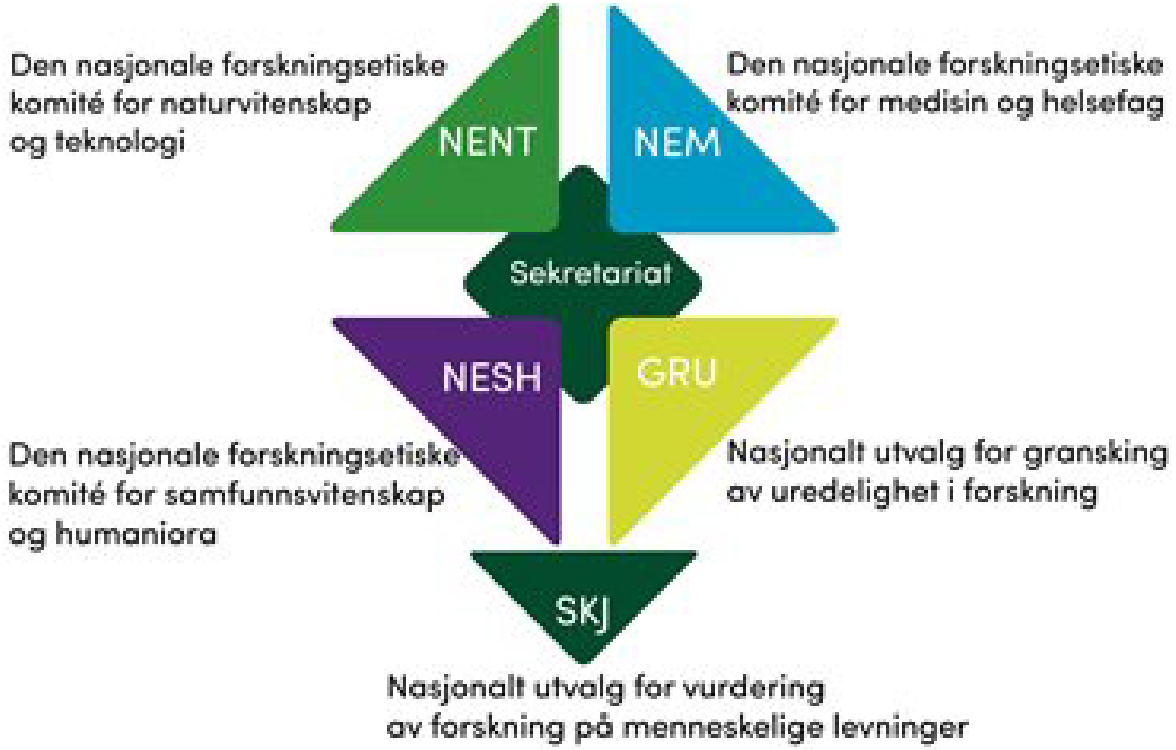
Norwegian resources

- The Research Ethics Library
- Data Protection Services at NSD
- The Norwegian Data Protection Agency

Guidelines at UiT

- Principles and guidelines for management of research data at UiT (2021)
- UiT – Guidelines for handling personal data in research projects
- Information security and data protection at UiT
- Guidelines for research ethics at UiT
- Intellectual Property Rights regulations at UiT
- Ethical guidelines for projects with external collaborators.
- Ethical guidelines for health research on Sami population
- Routines for health research at the Faculty of Health Sciences

The Norwegian National Research Ethics Comittees



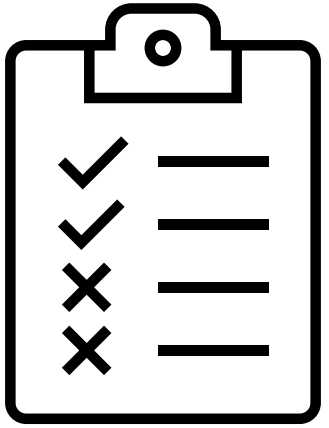
- General research ethical guidelines
- Social science and humanities (NESH)
- Medical and health research(NEM)
- Natural sciences and technology (NENT)
- Guidelines for research on human remains
- Guidelines for internet-research

Relevant groups and committees at UiT

- Privacy and information security office (Faggruppe for informasjonssikkerhet og personvern, IT-department)
- Information Security and Privacy Forum
- The ethics board (Forskningsetisk utvalg)
- The data privacy officer (personvernombud):
  Sølvi Brendeford Anderssen

Approval by ethical boards

- What is the practice at UiT (Norway in general)



- Personal data – notification to Sikt (NSD)
- Health data – Approval from The regional comitees for medical and health research ethics (REK).
- Ethical assessments…

Comitees at UiT

- Forskningsetisk komité ved Institutt for psykologi (IPS)
- *Etisk utvalg ved HSL-fakultetet*

Ethical self-assessment

A systematic ethical process to identify, reflect upon and manage ethical challenges within research projects.

- Define the problem
- Evaluate options for handling them
- Assess the selection of participants
- Assess the data needs
- Evaluate what values are at stake and what considerations are needed for the different alternatives
- Decide which values and considerations to prioritize
- Strategies to minimize risk
- Involve the participants?

The power asymmetry between researcher and participant

- Individuals lacking capacity for consent
- Economically disadvantaged
- Individuals susceptible to undue influence
- Ethnic minorities

→ Stricter ethical rigor:

- Selection, recruitment and initial contact
- Assessment of capacity for consent and voluntariness
- Confidentiality
- Follow-up during and after project ends.

Article on vulnerable groups in The Research Ethics Library

CARE Principles for Indigenous Data Governance

Ethical guidelines for indigenous data.

https://www.gida-global.org

Collective Benefit

Data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.

C1

For inclusive development and innovation

Governments and institutions must actively support the use and reuse of data by Indigenous nations and communities by facilitating the establishment of the foundations for Indigenous innovation, value generation, and the promotion of local self-determined development processes.

C2

For improved governance and citizen engagement

Data enrich the planning, implementation, and evaluation processes that support the service and policy needs of Indigenous communities. Data also enable better engagement between citizens, institutions, and governments to improve decision-making. Ethical use of open data has the capacity to improve transparency and decision-making by providing Indigenous nations and communities with a better understanding of their peoples, territories, and resources. It similarly can provide greater insight into third-party policies and programs affecting Indigenous Peoples.

C3

For equitable outcomes

Indigenous data are grounded in community values, which extend to society at large. Any value created from Indigenous data should benefit Indigenous communities in an equitable manner and contribute to Indigenous aspirations for wellbeing.

Responsibility

Those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples' self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples.

R1

For positive relationships

Indigenous data use is unviable unless linked to relationships built on respect, reciprocity, trust, and mutual understanding, as defined by the Indigenous Peoples to whom those data relate. Those working with Indigenous data are responsible for ensuring that the creation, interpretation, and use of those data uphold, or are respectful of, the dignity of Indigenous nations and communities.

R2

For expanding capability and capacity

Use of Indigenous data invokes a reciprocal responsibility to enhance data literacy within Indigenous communities and to support the development of an Indigenous data workforce and digital infrastructure to enable the creation, collection, management, security, governance, and application of data.

R3

For Indigenous languages and worldviews

Resources must be provided to generate data grounded in the languages, worldviews, and lived experiences (including values and principles) of Indigenous Peoples.

Authority to Control

Indigenous Peoples' rights and interests in Indigenous data must be recognised and their authority to control such data be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, as well as Indigenous lands, territories, resources, knowledges and geographical indicators, are represented and identified within data.

A1

Recognizing rights and interests

Indigenous Peoples have rights and interests in both Indigenous Knowledge and Indigenous data. Indigenous Peoples have collective and individual rights to free, prior, and informed consent in the collection and use of such data, including the development of data policies and protocols for collection.

A2

Data for governance

Indigenous Peoples have the right to data that are relevant to their world views and empower self-determination and effective self-governance. Indigenous data must be made available and accessible to Indigenous nations and communities in order to support Indigenous governance.

A3

Governance of data

Indigenous Peoples have the right to develop cultural governance protocols for Indigenous data and be active leaders in the stewardship of, and access to, Indigenous data especially in the context of Indigenous Knowledge.

Ethics

Indigenous Peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.

E1

For minimizing harm and maximizing benefit

Ethical data are data that do not stigmatize or portray Indigenous Peoples, cultures, or knowledges in terms of deficit. Ethical data are collected and used in ways that align with Indigenous ethical frameworks and with rights affirmed in UNDRIP. Assessing ethical benefits and harms should be done from the perspective of the Indigenous Peoples, nations, or communities to whom the data relate.

E2

For justice

Ethical processes address imbalances in power, resources, and how these affect the expression of Indigenous rights and human rights. Ethical processes must include representation from relevant Indigenous communities.

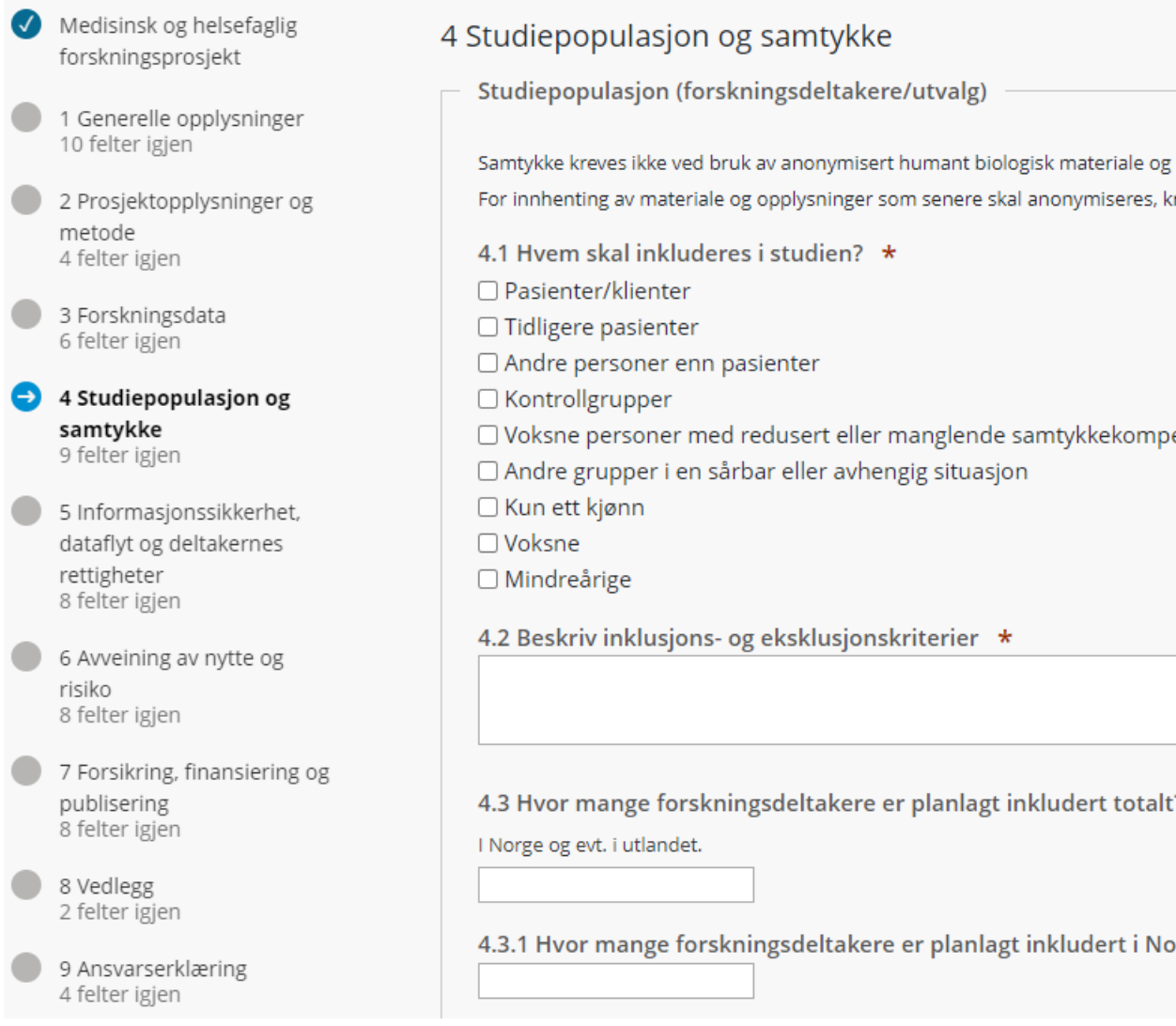
E3

For future use

Data governance should take into account the potential future use and future harm based on ethical frameworks grounded in the values and principles of the relevant Indigenous community. Metadata should acknowledge the provenance and purpose and any limitations or obligations in secondary use inclusive of issues of consent.

Health research data: Preapprovement by REK Nord

- Required for all projects subject to the Health Research Act
- Project leader is responsible
- A project protocol must be provided with the application
- To check if your project needs approval, you may submit a request for pre-assessment (fremleggingsvurdering)



Health data from the Sámi population

Ethical guidelines for sami health research are meant to strengthen the collective rights of the Sámi's as Indigenous Peoples in research projects.

Applies to research

- on Sámi people as groups.
- In municipalities and regions were Sámi make up a majority or a considerable part of the population.
- Where Sámi language, culture, tradition and/or history is relevant.

In Sami health research a collective consent is required.

The authority to give such a consent is delegated by the Sami Parliament to an expert committee.

Application form for Sami collective consent

Personal data: Duty to notify Sikt Personal Data Protection Services (former NSD)

Services provided by Sikt

- Supply legal advise on personal data protection for research purposes.
- Registry of projects handling personal data on behalf of UiT. Sikt perform «pre-approvals» and follow-ups on projects.
- Archiving data about humans and society that cannot be archived in an open archive.

Duty to notify (meldeplikt):

- All projects processing personal information must notify Sikt.
- The form must be submitted at least 30 days before the data collection starts.
- Sikt assess whether the processing of personal information is legal and will give feedback on the notification.
- The feedback from Sikt (to be considered an approval) is your documentation that you are processing the data in a legal way.
- Sikt does no ethical assessments beyond checking the legal aspects related to personal information.
- Read more about the notification form at Sikt

Lawful basis for processing

Definition: **Processing** of personal information encompasses any action relating to personal information, including but not limited to, collecting, registering, storing, analyzing and distributing.

All processing of personal information must have a lawful basis (article 6 nr. 1 in GDPR)

The most relevant lawful basis for research are:

- Concent (6.1.a),
- For tasks carried out in the public interest (6.1.e)

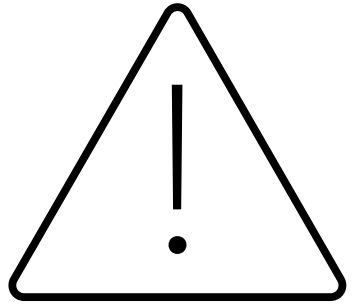
Sikt (NSD) recommends **consent** as the lawful basis of choice

If «task carried out in the public interest» is used as the lawful basis, the data processing must be considered essential in order to perform a task deemed necessary for the interest of the public (a purpose related to scientific or historical research).

As a rule of thumb, it is not legal to process special categories data unless one of the exceptions in article 9 are fulfilled.

Learn more about legal basis at Sikt

Data Processing Impact Assessment (DPIA)



A systematic description of data processing, purpose, risk and measures to minimize risk.

Must assess and convince that the quality and usefulness of the study justifies the privacy disadvantage for the participants, as well as demonstrate measures to minimize the risks as much as possible.

The Personal Data Act (Art.35) requires a DPIA when:

- Special categories data
- Vulnerable groups
- Transfer outside of Europe
- Use of new technology
- Particularly complex or extensive data collection (time, geographic reach, population size)
- In high-risk projects a pre-consultation with The Norwegian Data Protection Agency is required.

The Norwegian Data Protection Agency has guidelines and a checklist for how to perform a DPIA (Datatilsynet)

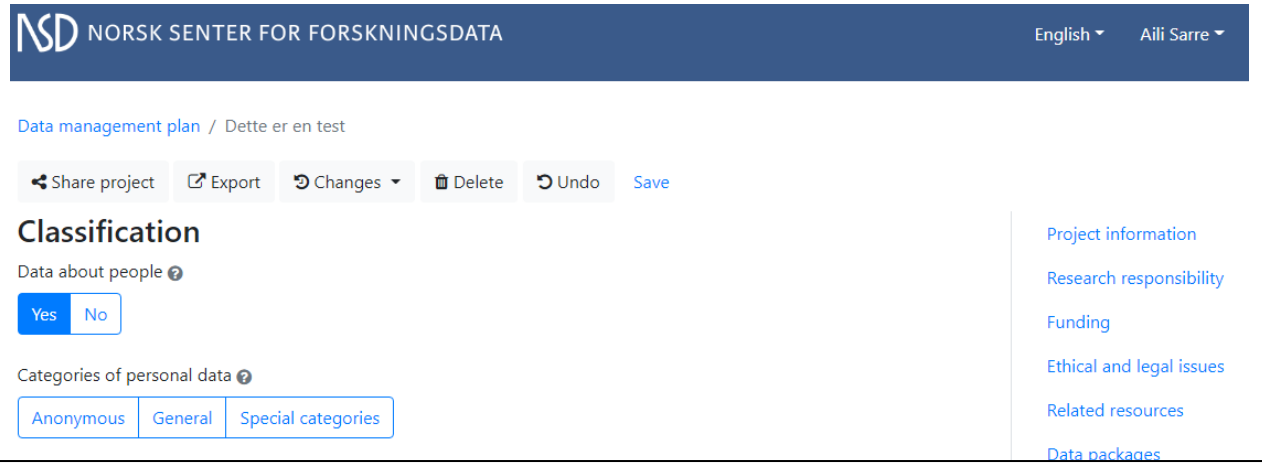
Data Management plans (DMP)

Required according to [Principles and guidelines for management of research data at UiT]

Project subject to notification to NSD,
use their online tool:
https://www.nsd.no/en/create-a-data-management-plan

Project funded by EU,
Use templates on the DMPOnline platform
https://dmponline.dcc.ac.uk/

All other projects, the UiT template:
DMP template in the Research Data Portal



Help and resources:

- Guide with the required contents of a DMP
- The Research Data Portal at UiT
- The webinars on research data managment
- Contact the Library researchdata@hjelp.uit.no for help and feedback.

Collaborations on data handling

- Clarify roles, rights and responsibilities
- Establish a common agreement and routines on how data should be handled within the project.
- GDPR applies only within EU/EØS. For transfers outside of Europe a specific ground for transfer.
- When data is to be processed by external actors, use a Data Processing Agreement – Templates for agreements
- Consider potential costs of data processing and storage

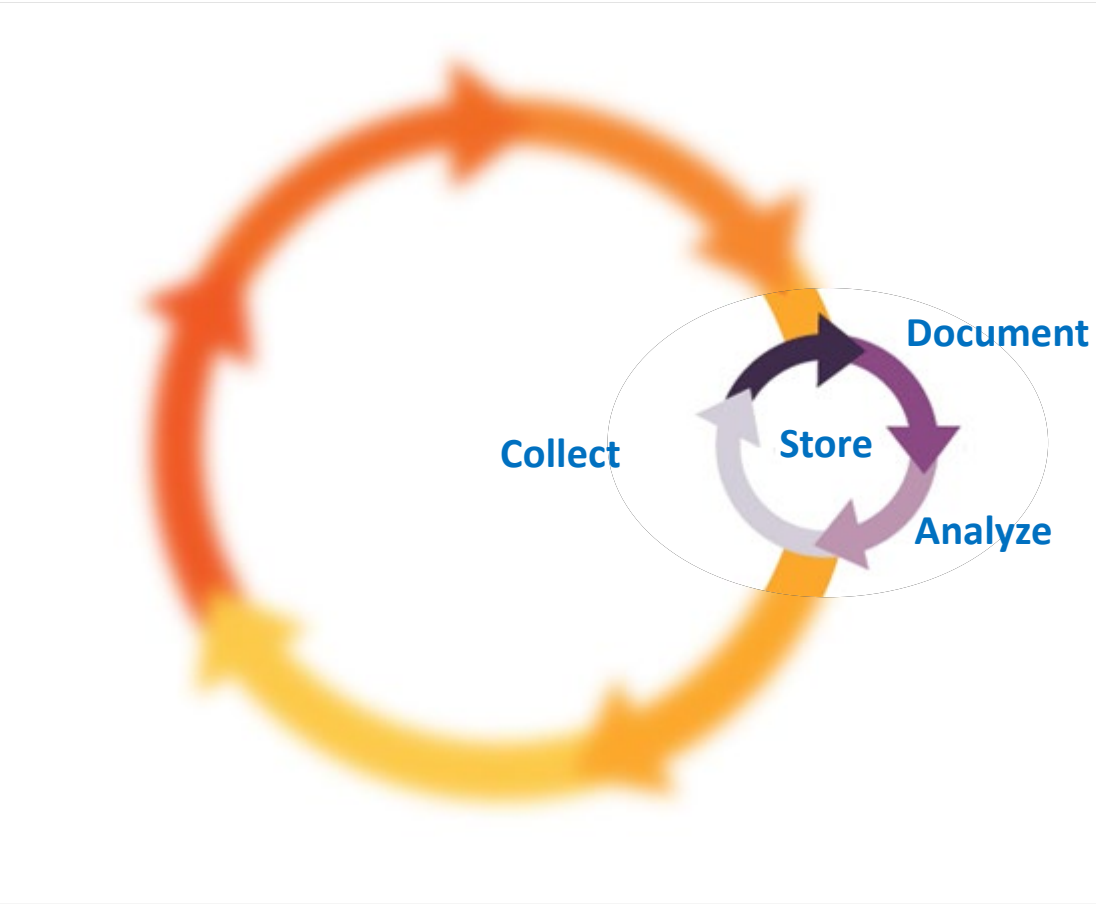
Start with a common overview of the material to be collected

Formats	Storage	Access	Storage	Access
	Analysis & Editing phase **Possibility to withdraw from the study**		**Archiving phase** **No possibility to withdraw from the study** **GDPR rights reserved**	
Video raw (.wav)	Office 365	Researchers within ADLab; Co-researchers & co-creators based on demand	NSD	Accessible by other researchers based on demand & permission Possibility to delete personal data
Conversations raw (.mp3)				
Pictures of artworks, notes, people raw (.jpg)				
Writings, notes, logs (.txt)				
Edited video recordings; non anonymised; consent by editing (.wav)			Published (e.g. Research Film, visual art)	Unrestricted
Edited conversation recordings; non anonymised; consent by editing (.mp3)			Published (e.g. Podcast)	

- Who should have access
- Where to store

[Digital data handling within Artful Dementia Research Lab]

The active phase



- Classifying files
- Secure storage
- Informing participants
- Documentation of consent
- Anonymization

Secure handling and storage

- Guidelines for classifications of files at UiT
- Encrypting of files to be transferred or stored on hardware
- IT-services and systems – what can you use for your data?
- Establish secure routines for transferring data
- Access control
- Store keys and code sheets separate
- Routines when using private equipment
- Nettskjema Survey Dictaphone app
- Tjenester for sensitive data (TSD)

Grønn	Gul	Rød	Svart
Åpen	*Intern*	*Fortrolig*	*Strengt fortrolig*

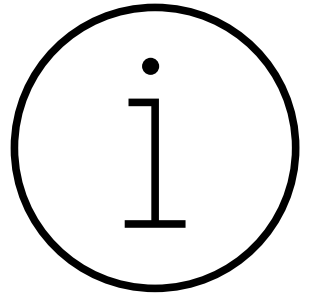
System / tjeneste	Åpen/Grønn	Intern/Gul	Fortrolig/Rød	Strengt fortrolig/Svart
Sharepoint (office 365)	OK	OK	OK [2]	*ikke godkjent*
Stream (office 365)	OK	OK	*ikke godkjent*	*ikke godkjent*
Sway[5] (office 365)	OK	*ikke godkjent*	*ikke godkjent*	*ikke godkjent*
Teams (office 365) - filer	OK	OK	OK [2]	*ikke godkjent*
Teams (office 365) - møter	OK	OK	OK [3]	*ikke godkjent*

More info:

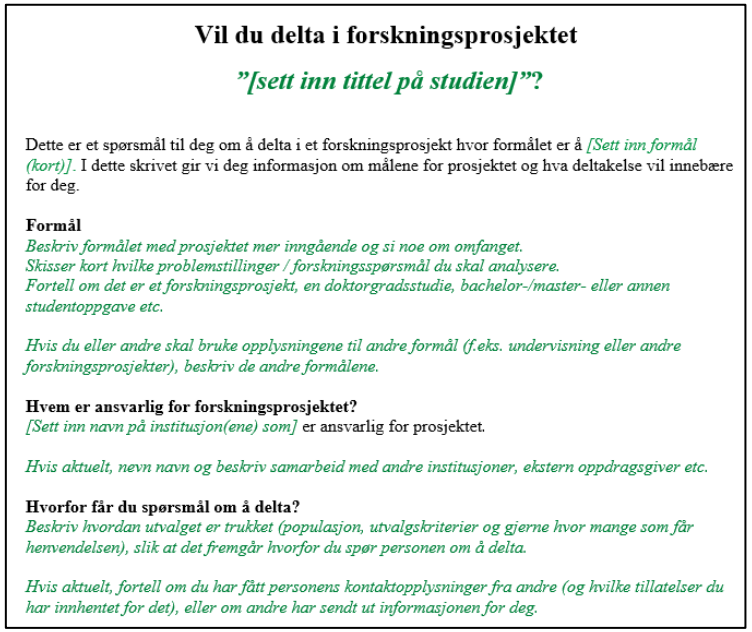
- Information security and data protection at UiT
- Webinar: Storing research data

Duty to inform

The information letter should be short, and easy to comprehend for the target group



- Purpose of the project and possible consequences for the participants.
- How the data will be used, and what will happen to it after the project is finished.
- Inform about the rights of the participants and provide contact information to enquire more information.



Vil du delta i forskningsprosjektet

"[sett inn tittel på studien]"?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å *[Sett inn formål (kort)]*. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål
Beskriv formålet med prosjektet mer inngående og si noe om omfanget.
Skisser kort hvilke problemstillinger / forskningsspørsmål du skal analysere.
Fortell om det er et forskningsprosjekt, en doktorgradsstudie, bachelor-/master- eller annen studentoppgave etc.

Hvis du eller andre skal bruke opplysningene til andre formål (f.eks. undervisning eller andre forskningsprosjekter), beskriv de andre formålene.

Hvem er ansvarlig for forskningsprosjektet?
[Sett inn navn på institusjon(ene) som] er ansvarlig for prosjektet.

Hvis aktuelt, nevn navn og beskriv samarbeid med andre institusjoner, ekstern oppdragsgiver etc.

Hvorfor får du spørsmål om å delta?
Beskriv hvordan utvalget er trukket (populasjon, utvalgskriterier og gjerne hvor mange som får henvendelsen), slik at det fremgår hvorfor du spør personen om å delta.

Hvis aktuelt, fortell om du har fått personens kontaktopplysninger fra andre (og hvilke tillatelser du har innhentet for det), eller om andre har sendt ut informasjonen for deg.

SIKT provides templates for [information letters and consent forms]

Concent

- Voluntary
- Should be as easy to withdraw the consent as giving it – no consequences to withdraw.
- Should be an active opt-in
- Clear and explicit information
- Should be granular and specific
- Should be documented

} Explicit consent when the data is sensitive, or the data will be shared.

SIKT provides templates for [information letters and consent forms]

Issues with informed consent

- In rare cases, informing the participants may invalidate the data.
- Unexpected information – participants can reveal information about 3rd parties or special categories data.
- Lack of capacity for informed consent - collect consent from legal guardian.

SIKT provides templates for information letters and consent forms

Anonymized or pseudonymized

Anonymized data cannot lead to re-identification!

- If re-identification is possible, the data are **de-identified**, not anonymous.

De-identified information (pseudonymized)
Identifiable information is removed or hidden, but can be reconstructed (keys, code sheets).

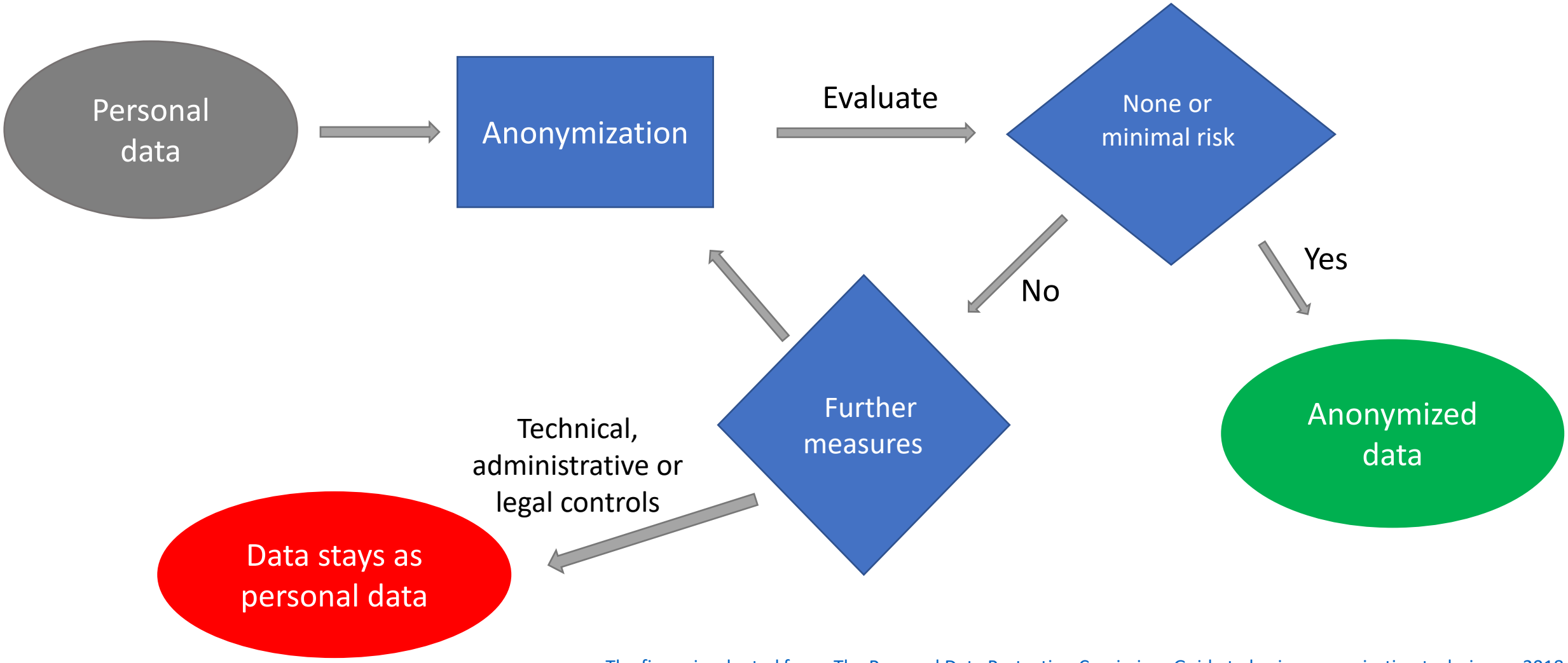
De-identified information is not anonymous in the legal sense. GDPR still applies.

Methods for anonymization

- Removing information (permanent)
- Masking
- Pseudonymization (permanent or reversible)
- Generalization
- Aggregation
- Synthetic data

The Personal Data Protection Comission: Guide to basic anonymisation techniques, 2018

The process of anonymization can be iterative



The figure is adapted from: The Personal Data Protection Comission: Guide to basic anonymisation techniques, 2018

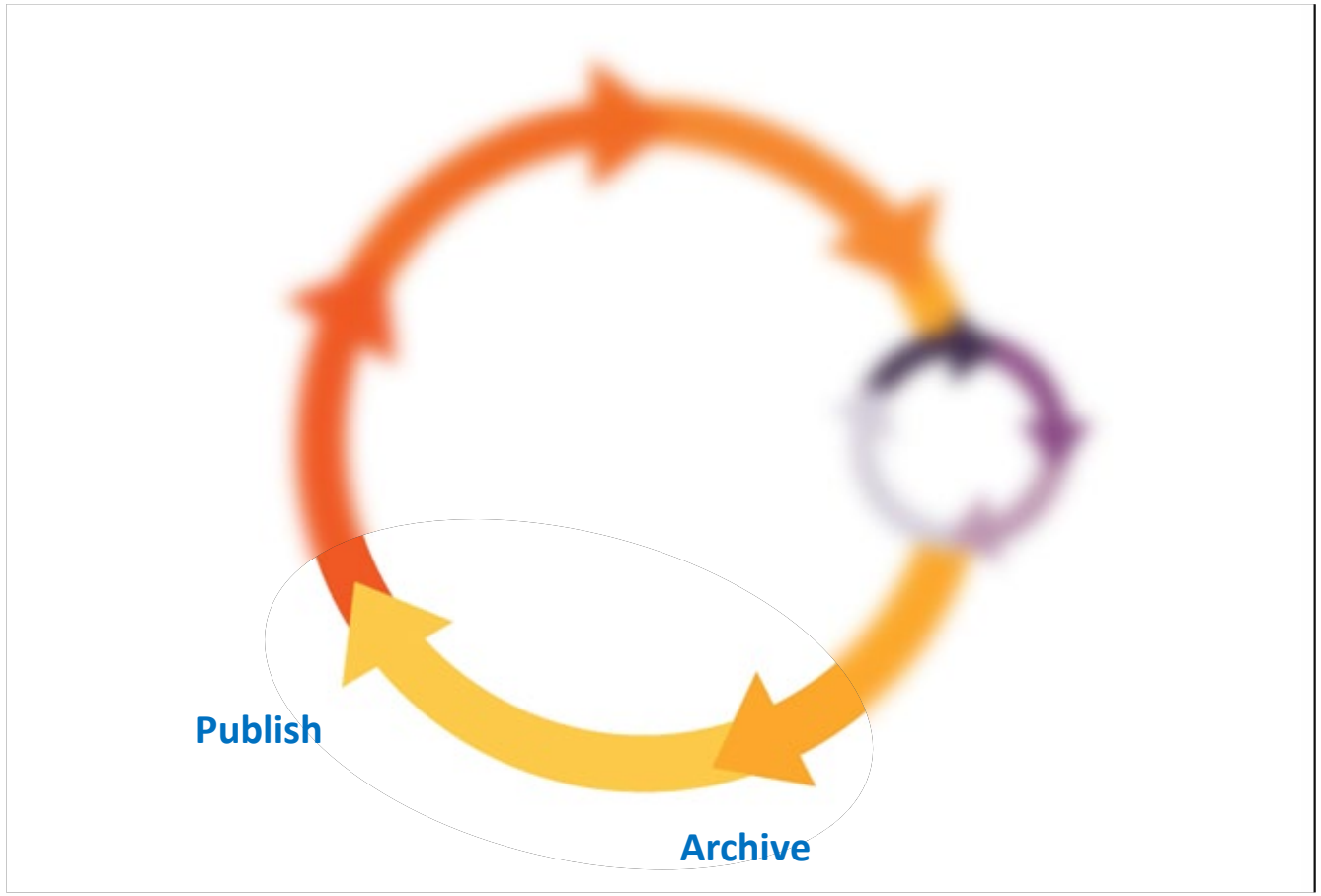
Challenges with anonymization

- Assess carefully the size of your selection and population.
- If possible, apply a quantitative measure of anonymity (k-anonymity).
- Anonymization will in most instances reduce the quality of the data.
- Anonymizing qualitative data is particularly time consuming/challenging.
- The data protection officer at UiT:

> «*Assessing whether data are anonymous requires extensive knowledge of the particular data in question as well as knowledge of the discipline.*»

Important: The researcher is responsible for making sure the data is sufficiently anonymized!

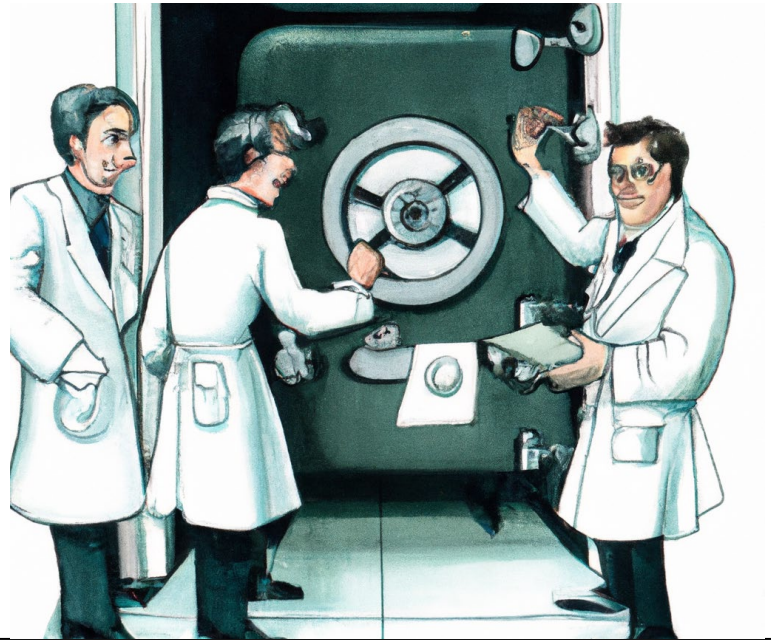
Project end: Archiving and sharing



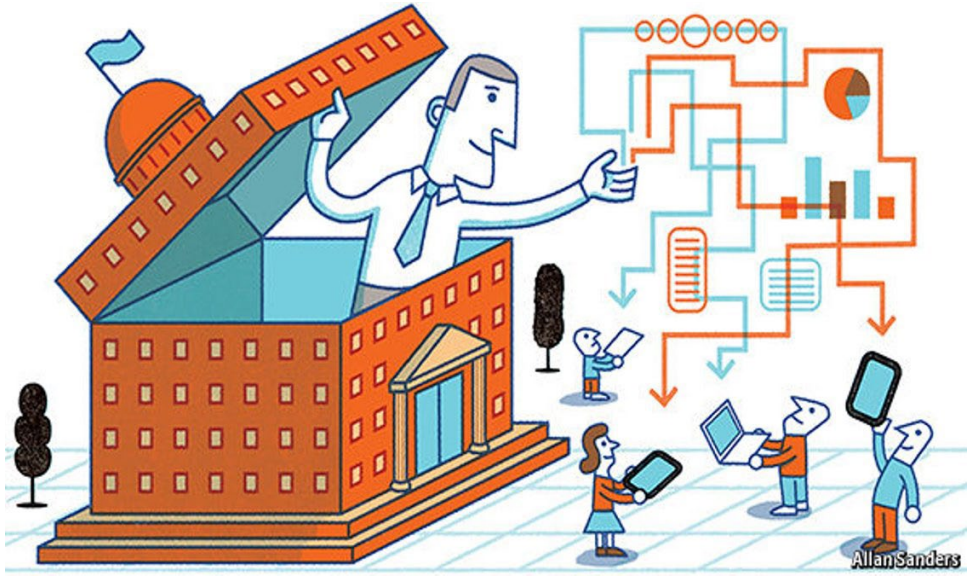
- Deleting vs. preserving
- Long-term storing vs. archiving
- Open archiving vs. regulated access
- **Essential question: What have the participants consented to?**

Long-time storage or archiving

Long-time storage – The researcher and/or group stay in control of the data. The data is stored on a closed data storage unit according UiTs guidelines.



Archiving – The data is deposited with the purpose that others should be able to reuse them. The archive can be open or closed.



Requirements and expectations on data reuse and sharing

Norwegian Ministry of Education and Research

1: Research data must be as open as possible, as closed as necessary.

2: Research data should be managed and curated to take full advantage of their potential.

3: Decisions concerning archiving and management of research data must be taken within the research community.

National strategy on access to and sharing of research data

UiTs Principles and guidelines for RDM

The researcher shall make research data openly available for further use to all relevant users, providing there are no legal, ethical, security or commercial reasons for not doing so.

Principles and guidelines for RDM at UiT

Challenges when archiving data from studies on human participants

Foundational principles of the GDPR:

- Purpose limitation – personal data should only be collected for specified, explicit, and legitimate purposes.
- Data minimization – only collect the smallest amount of data you'll need to complete your purposes.
- Storage limitation – justify the length of time you're keeping each piece of data you store.

- The identifiable individuals have the rights to access, rectification, and erasure of their data.

Read more at Datatilsynet

Deletion or preservation

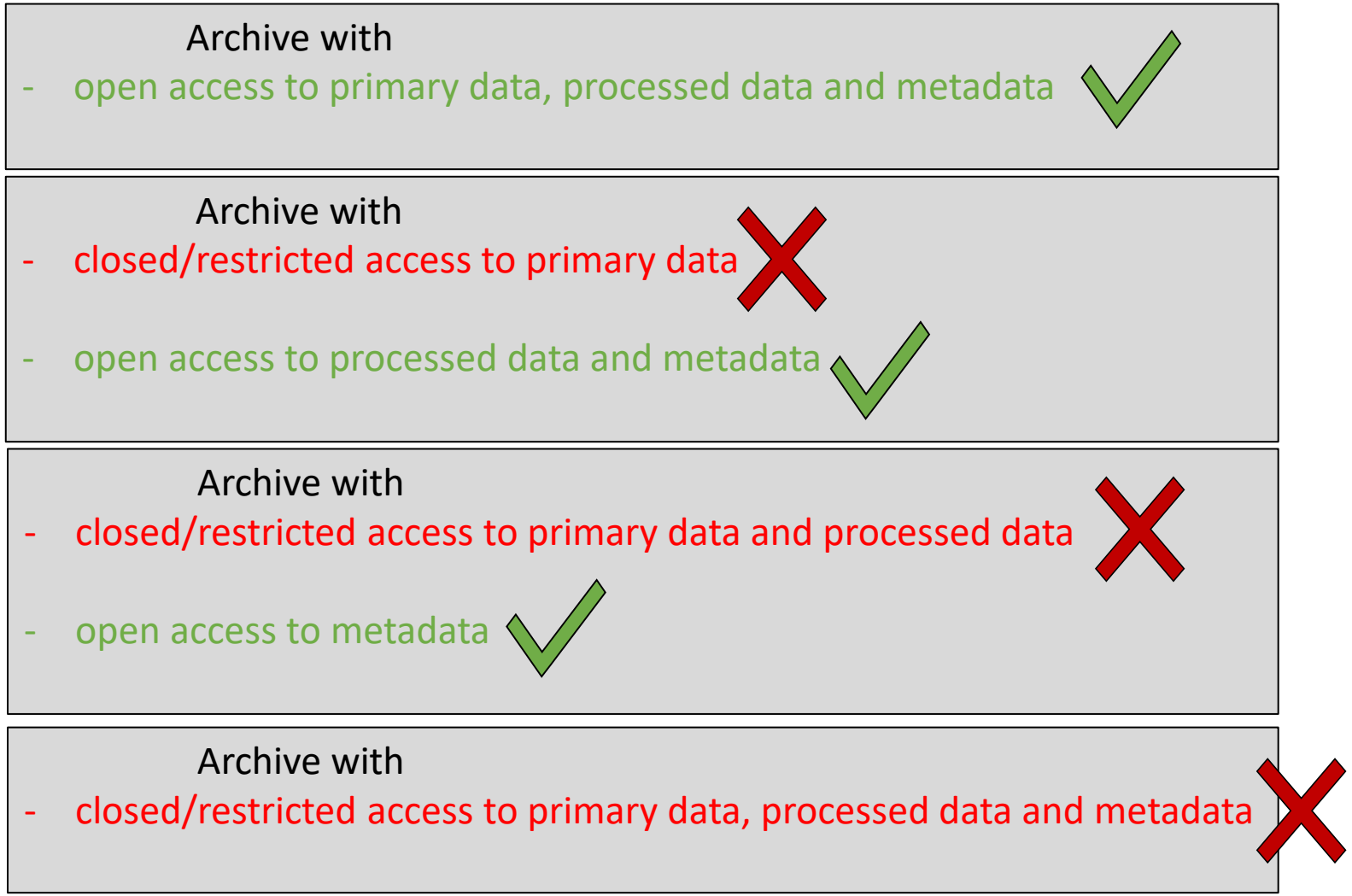
- Personal information and health data should not be stored longer than necessary to complete the project. A date for the deletion or anonymization is given in the decision from REK/NSD.
- Important to establish routines for deleting what needs to be deleted.
- Lang-term storage of research data for follow-up studies or related studies must be approved by NSD/REK.
- In some instances, REK vil require storage for 5 years for the purpose of potential inspection (The Health Research Act § 38).
- If research data is to be stored longer than stated in the original consent, **a new consent** should be collected from the participants.

McCrae, & Murray, J. (2008). When to Delete Recorded Qualitative Research Data. *Research Ethics Review*, *4*(2), 76–77. https://doi.org/10.1177/174701610800400211

How to archive data from studies on human participants

- Plan ahead.
- Use Sikt/NSD as an advisor.
- It must be clear to the participants what will happen to the data after the project is done.
- The consent should be granular and specific.
- Assess whether parts of the material can be anonymized. Even though these data would be legal to publish, it remains ethically dubious if the participants were not informed of this intent with the data.
- Choose the right archive

Which data can be archive openly and which needs access regulation?



Archiving data with controlled access

Data on humans and society

- To be archived at Sikt (NSD - Norsk senter for forskningsdata)
- NSD will accept data to be shared **as open as possible but as closed as necessary**. Data can not have more restrictive terms than necessary.
- Sikt will advice on documentation and how the access can be regulated.
- https://www.nsd.no/en/archiving-research-data

Data from projects with commercial interests:

- Find an archive that offer an embargo period before the dataset is openly published.

Archiving datasets in research data archives

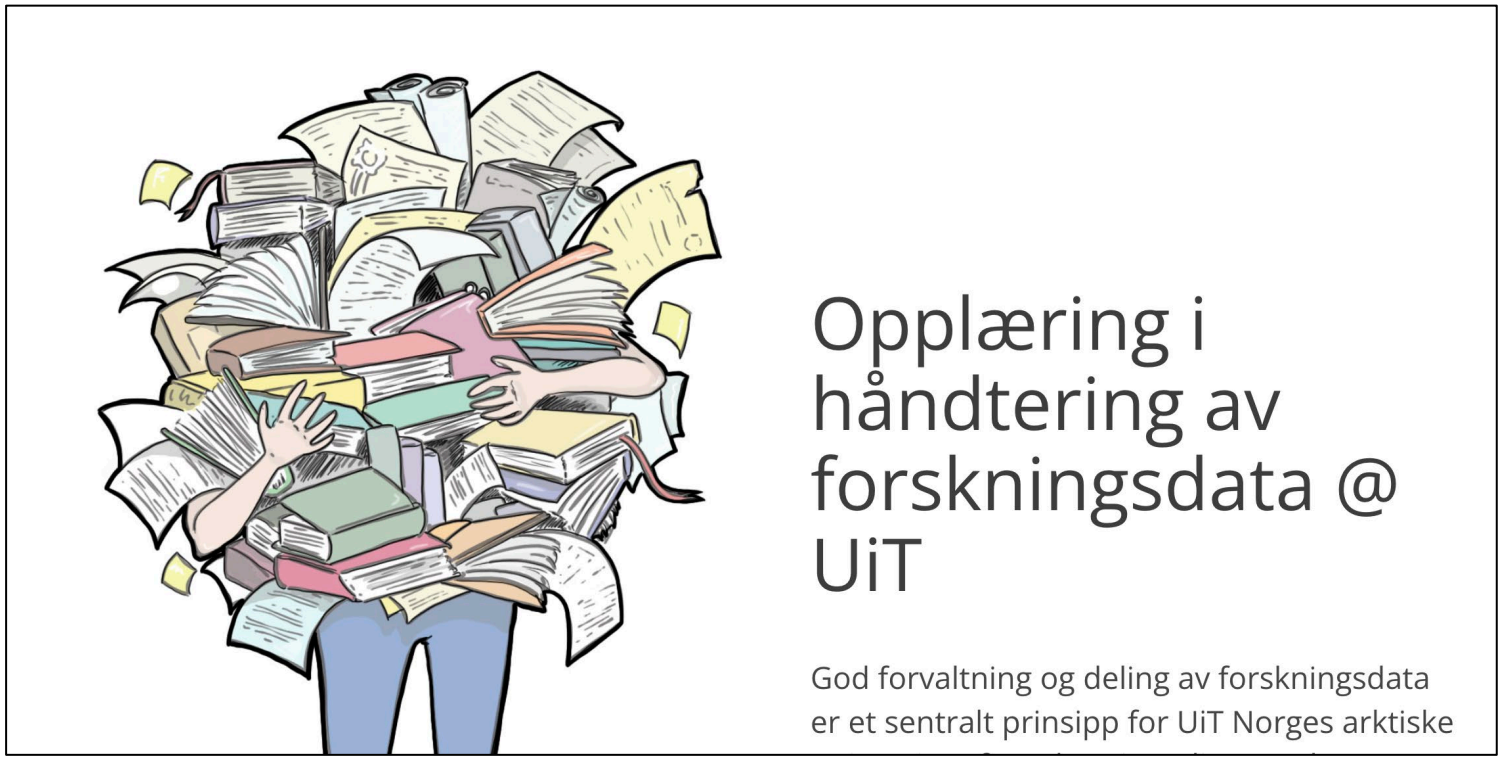
- Data files
- Metadata
- Readme-file

Other supplementary documentation that should accompany the data if relevant:

- Notification form and evaluation from NSD
- Terms for reuse
- Information provided to participants and consent form.
- Data Processing Impact Assessment or ethical evaluation
- Data Processing Agreement

Learn more about archiving data and chosing the right archive on webinar on April 18th

Help and training in research data managment



Support from the publication and research support group at the Library: https://en.uit.no/ub/researchandpublish

Series of open webinars on RDM: Site.uit.no/rdmtraining

Webinars on research data management

April 2023

Sign up here:
https://uit.no/forskning/forskningsdata/art?p_document_id=721540

How to structure and document research data
17. APRIL 2023

How to store research data
17. APRIL 2023

Data cleaning
17. APRIL 2023

How to archive research data
18. APRIL 2023

How to archive data in UiT Open Research Data
18. APRIL 2023

How to search and cite research data
19. APRIL 2023

Research data: Rights and licenses
19. APRIL 2023

How to use an electronic lab notebook
20. APRIL 2023

How to write a Data Management Plan
20. APRIL 2023

Workshop: Data cleaning/Datavask OpenRefine
25. APRIL 2023

Kick-off-seminar datarøkternettverk (data steward network)
3. MAI 2023

Kick-off seminar data steward network
3. MAI 2023

Feedback:

- skjema.uio.no/ubevalno

- Date: 13.04.2022
  Emne: Sensitive data