UiO : ARENA Centre for European Studies
University of Oslo

Data management and GDPR for researchers

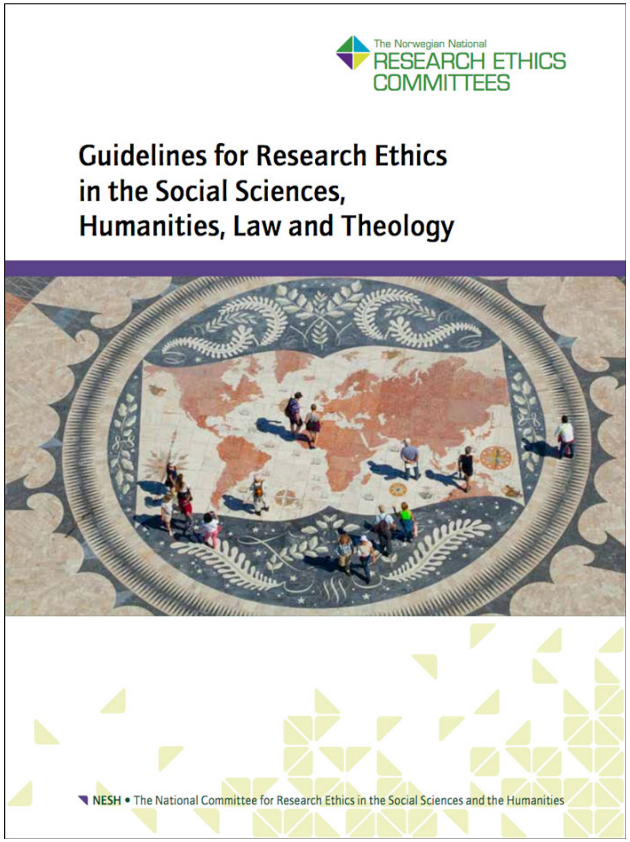
Staff info meeting, 19 March 2019

Marit Eldholm

Research ethics (NESH)

The Norwegian National Research Ethics Committee (NESH): Guidelines
(4th ed, June 2016)

- Respect for individuals
  - Privacy, duty to inform, consent, confidentiality, limited reuse, storage…
- Respect for groups and institutions
  - Public administration: restricted access
- The research community
  - Co-authorship, plagiarism, citation practice, data sharing, impartiality, integrity, student-supervisor relationship…
- Commissioned research
- Dissemination of research



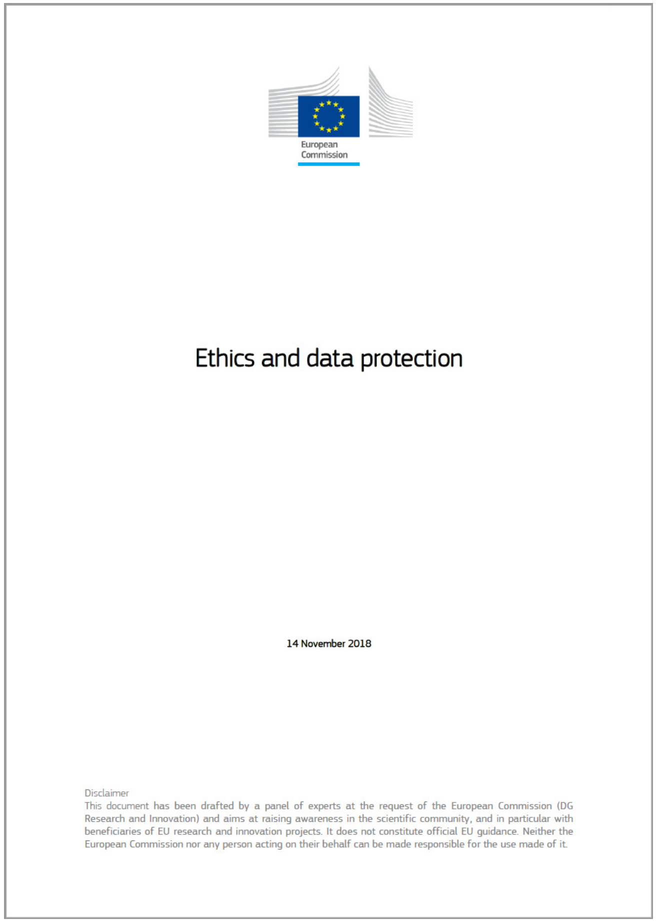
The Norwegian National
RESEARCH ETHICS
COMMITTEES

Guidelines for Research Ethics
in the Social Sciences,
Humanities, Law and Theology

NESH • The National Committee for Research Ethics in the Social Sciences and the Humanities

European Commission

“The fact that your research is legally permissible does not necessarily mean that it will be deemed *ethical*. Crucially, if your research proposal involves the processing of any personal data, whatever method is used, you – and all of your partners, collaborators and service providers – must, if called upon, be able to demonstrate compliance with both **legal** and **ethical** requirements. Such requests could come from **data subjects**, **funding agencies** or **data protection supervisory authorities**. When developing and implementing your proposal, it is your responsibility to identify the appropriate legal provisions and ensure compliance.”

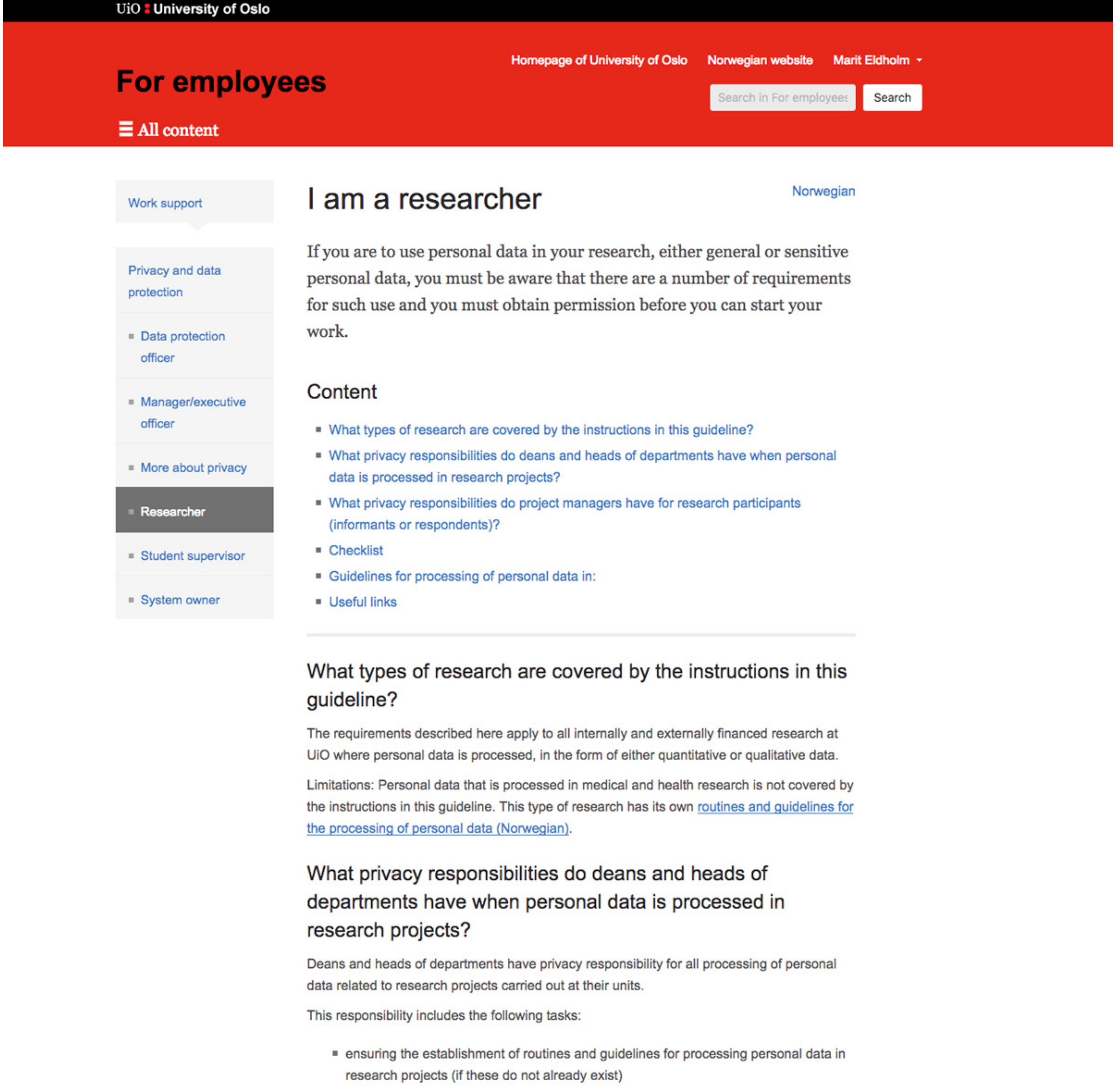


European Commission

Ethics and data protection

14 November 2018

Disclaimer
This document has been drafted by a panel of experts at the request of the European Commission (DG Research and Innovation) and aims at raising awareness in the scientific community, and in particular with beneficiaries of EU research and innovation projects. It does not constitute official EU guidance. Neither the European Commission nor any person acting on their behalf can be made responsible for the use made of it.



I am a researcher

If you are to use personal data in your research, either general or sensitive personal data, you must be aware that there are a number of requirements for such use and you must obtain permission before you can start your work.

Content

- What types of research are covered by the instructions in this guideline?
- What privacy responsibilities do deans and heads of departments have when personal data is processed in research projects?
- What privacy responsibilities do project managers have for research participants (informants or respondents)?
- Checklist
- Guidelines for processing of personal data in:
- Useful links

What types of research are covered by the instructions in this guideline?

The requirements described here apply to all internally and externally financed research at UiO where personal data is processed, in the form of either quantitative or qualitative data.

Limitations: Personal data that is processed in medical and health research is not covered by the instructions in this guideline. This type of research has its own routines and guidelines for the processing of personal data (Norwegian).

What privacy responsibilities do deans and heads of departments have when personal data is processed in research projects?

Deans and heads of departments have privacy responsibility for all processing of personal data related to research projects carried out at their units.

This responsibility includes the following tasks:

- ensuring the establishment of routines and guidelines for processing personal data in research projects (if these do not already exist)

Why ARENA guidelines?

- EU-funded projects, increasingly also national:
  - ➢ Research ethics reflections in proposals
  - ➢ Data Management Plans required
  - ➢ Data sharing expected (justified if not)
- Two sets of guidelines (For employees)
  - ➢ GDPR: Personal data in research
  - ➢ Managing and sharing research data

Who is responsible?

- UiO (‘data controller’)
  - Data owner (!), legally responsible, Data Protection Officer
  - Data subjects can claim compensation for injury
- Head of department (ARENA director)
  - Privacy responsibility for all processing of personal data related to research projects at their unit
  - Routines and guidelines, training resources, controls, internal UiO audit
- Researchers (‘project manager’)
  - Privacy of the research participants at all phases

25 May 2018

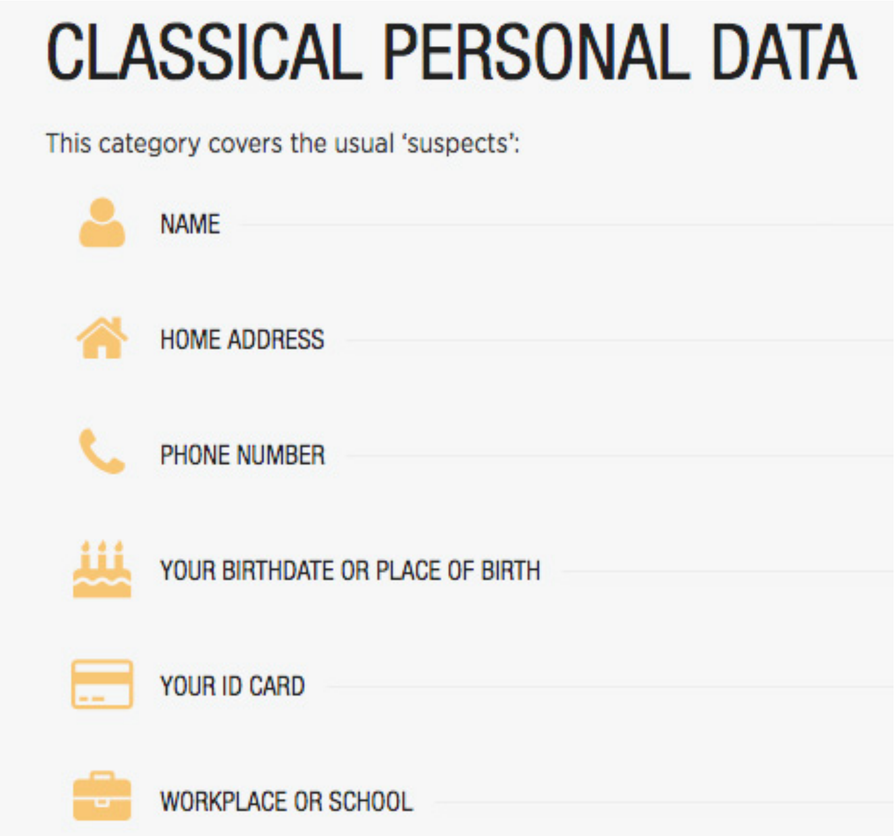
- Europe-wide data protection legislation
- Allow citizens to have greater control over how their data is stored and managed by third parties

GDPR: Possible pitfalls

- Misunderstanding of what counts as personal data
- Misinterpretation of ‘anonymity’
- Not sufficient information to research participants/informants
- Definition of sensitive personal data

Personal data

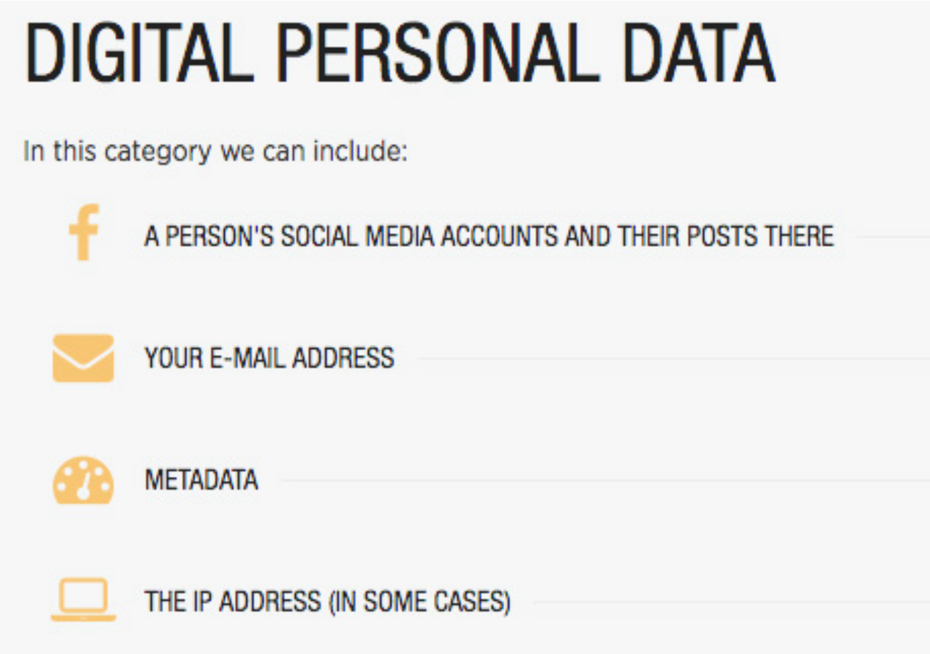
= anything that discloses your identity, that is unique to you (GDPR: ‘any information relating to an identified or identifiable natural person’)



CLASSICAL PERSONAL DATA

This category covers the usual ‘suspects’:

- NAME
- HOME ADDRESS
- PHONE NUMBER
- YOUR BIRTHDATE OR PLACE OF BIRTH
- YOUR ID CARD
- WORKPLACE OR SCHOOL



DIGITAL PERSONAL DATA

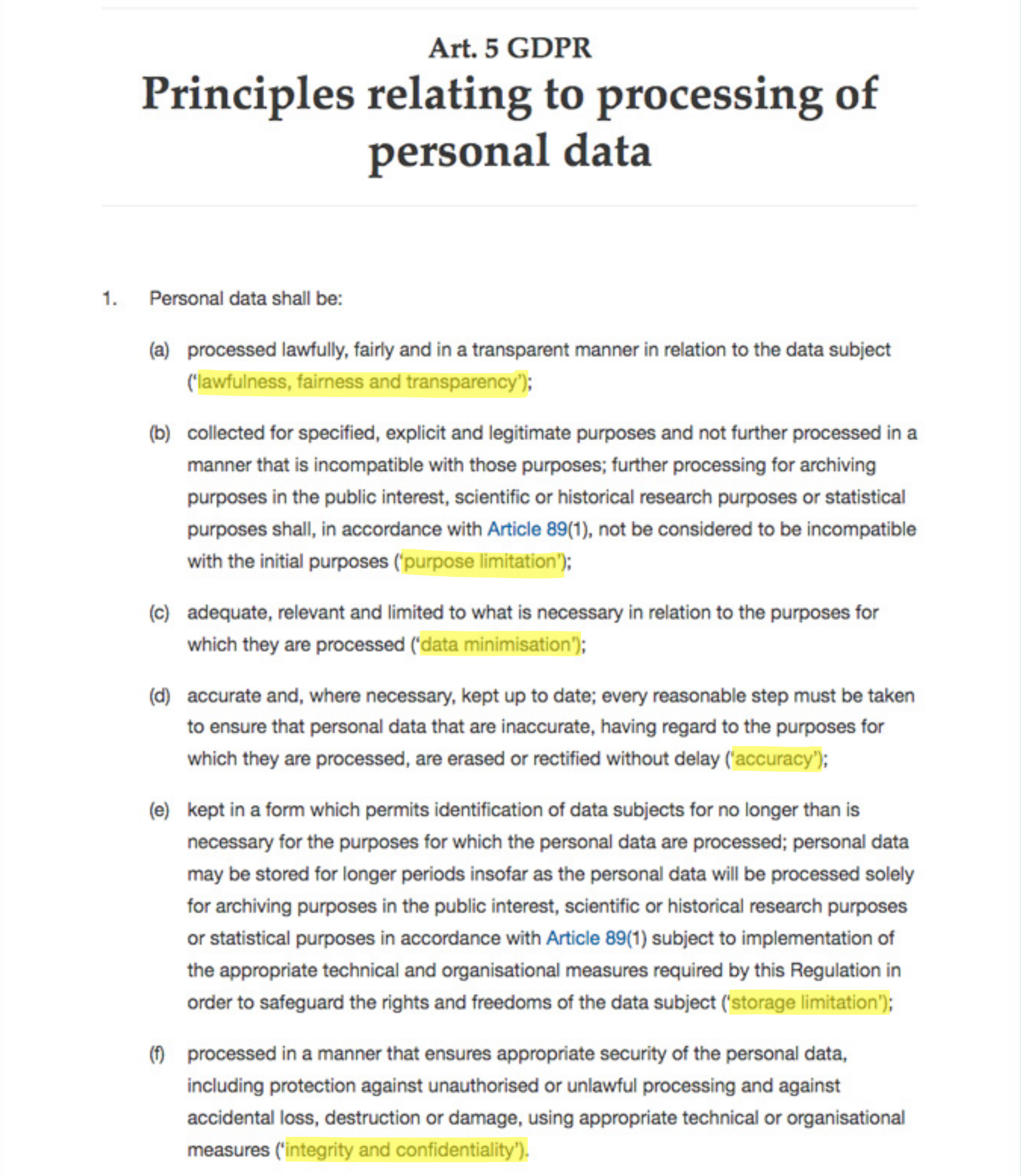
In this category we can include:

- A PERSON'S SOCIAL MEDIA ACCOUNTS AND THEIR POSTS THERE
- YOUR E-MAIL ADDRESS
- METADATA
- THE IP ADDRESS (IN SOME CASES)

Notification to NSD (Norwegian Centre for Research Data)

- Will you process personal data?
  - ✓ Notification Form to NSD (“meldeskjema”)
  - ✓ At least 30 days prior to commencing data collection
  - ✓ You must await their assessment before you start collecting data

- A project must be notified even if you only publish anonymous data
- It is <u>how you will process personal data</u> in the project as a whole, from the beginning of data collection to the publishing of results, that determines whether you have to submit a Notification Form



Art. 5 GDPR

Principles relating to processing of personal data

1. Personal data shall be:

   (a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');

   (b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');

   (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');

   (d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy');

   (e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');

   (f) processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality').

Important distinction

Anonymisation

Removing the possibility of identifying individuals in a dataset, with the tools that can reasonably be expected to be used, also for researcher

Outside the scope of GRPR

- Still have to inform about how the data is handled *before* anonymization

GDPR Recital 26

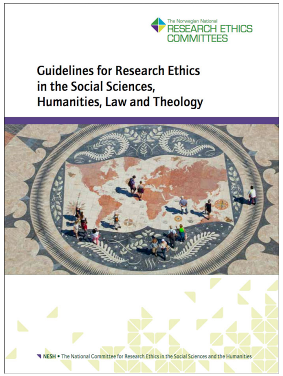
Pseudonymisation

The data can no longer be attributed to a specific data subject without the use of additional information
= de-identification

GDPR only partially applies

- data can be used beyond the original purpose
- data subjects' rights to access, rectification, or erasure no longer apply
- participants must be informed that such rights will be annulled upon pseudonymisation

Re-use of data beyond original purpose



The Norwegian National
RESEARCH ETHICS
COMMITTEES

Guidelines for Research Ethics
in the Social Sciences,
Humanities, Law and Theology

NESH • The National Committee for Research Ethics in the Social Sciences and the Humanities

Generally, re-use of identifiable personal data requires the consent of the participants.

This does not apply to anonymised data, acquired for example for use in statistics, where the researcher cannot link persons and data. When the data have been anonymised, the researcher does not know which person the data and the material come from. However, anonymity must not be confused with de-identified data, where personal data are removed, so that no unauthorised persons are able to establish who the research subjects are, but where the researcher is able to link individuals and data. Re-use of such de-identified data requires consent if researchers supplement registry studies with data obtained through active contact with the participants. When re-using and linking this type of data set, for example in registry studies that are large-scale, of a long duration, or which use geodata, it may also be possible to locate or identify individuals indirectly. In such cases, researchers should make renewed attempts to obtain consent, even though this is difficult in practice. If researchers do not find it possible to obtain consent, they have a particular responsibility to explain why the research is of such great benifit that it justifies deviating from this principle. In such cases, researchers have a general responsibility to inform the persons involved and the general public.

NESH guidelines, p. 18 (art. 10)

GDPR

Legal basis
‘lawfulness of processing’

Informed consent

GDPR Art. 6(1)(a)
The data subject has given consent to the processing of his or her personal data for one or more specific purposes

Public interest

GDPR Art. 6(1)(e)
Processing is necessary for the performance of a task carried out in the public interest

Legitimate interest

GDPR Art. 6(1)(f)
Processing is necessary for the purposes of the legitimate interests pursued by the controller […], except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data […]

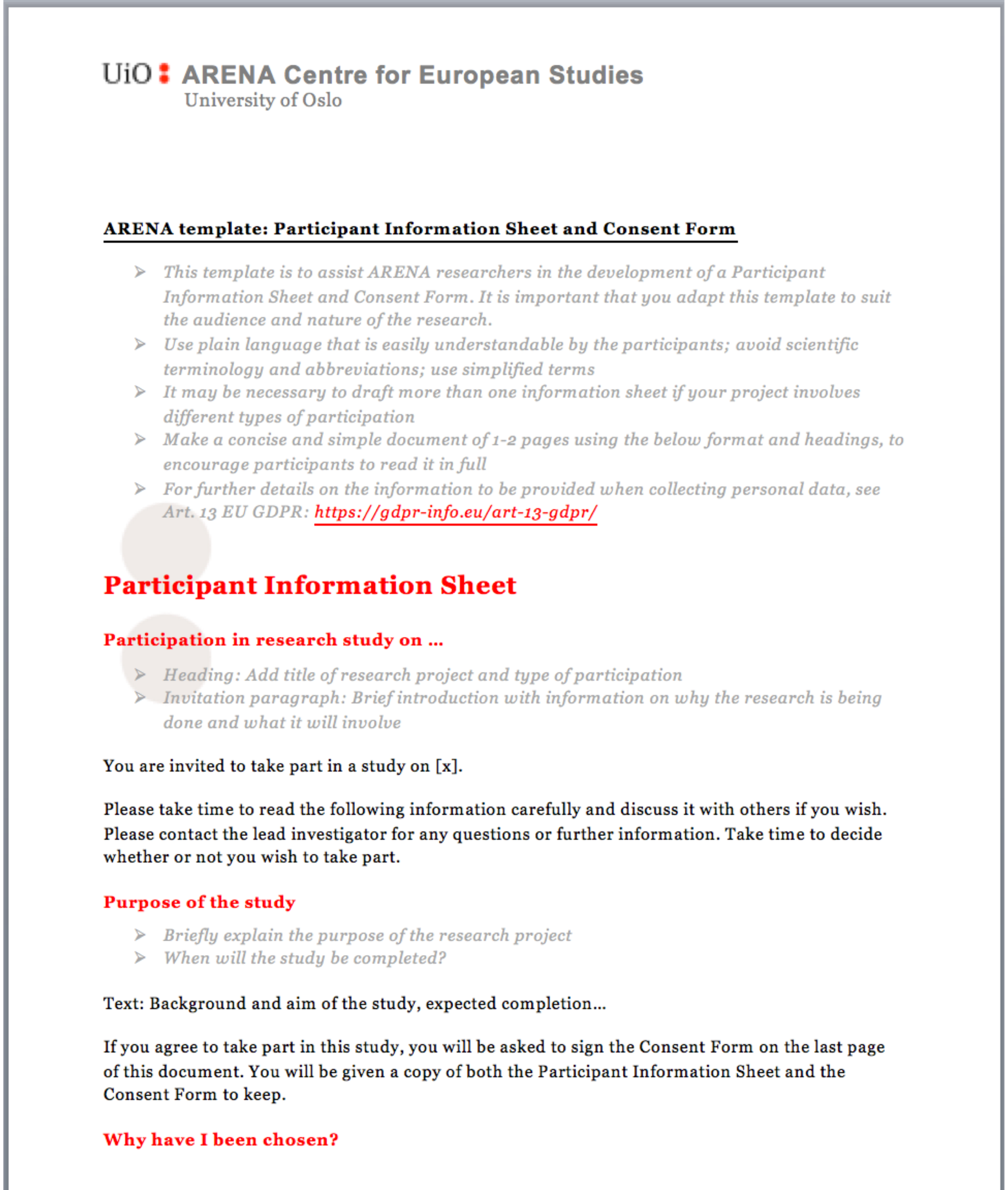
Researcher’s responsibility to consider, NSD to decide if well-grounded

Information to data subjects

- Art. 13 GDPR: Information to be provided where personal data are collected from the data subject
- Addressed in participant information sheet and consent form
    - **GDPR-compliant ARENA template**
    - To be adapted to the audience and nature of the research, plain language, simplified terms

Participant Information Sheet

- Purpose of the study
- Why has the informant been chosen?
  - ➢ Define population, sampling, recruitment procedure
- How and when data collected
- Recording, storage, deletion
- Pseudonymisation, anonymisation
- Confidentiality
- Right to access, rectification, erasure
- Contact details



UiO ARENA Centre for European Studies
University of Oslo

ARENA template: Participant Information Sheet and Consent Form

- *This template is to assist ARENA researchers in the development of a Participant Information Sheet and Consent Form. It is important that you adapt this template to suit the audience and nature of the research.*
- *Use plain language that is easily understandable by the participants; avoid scientific terminology and abbreviations; use simplified terms*
- *It may be necessary to draft more than one information sheet if your project involves different types of participation*
- *Make a concise and simple document of 1-2 pages using the below format and headings, to encourage participants to read it in full*
- *For further details on the information to be provided when collecting personal data, see Art. 13 EU GDPR: https://gdpr-info.eu/art-13-gdpr/*

Participant Information Sheet

Participation in research study on …

- *Heading: Add title of research project and type of participation*
- *Invitation paragraph: Brief introduction with information on why the research is being done and what it will involve*

You are invited to take part in a study on [x].

Please take time to read the following information carefully and discuss it with others if you wish. Please contact the lead investigator for any questions or further information. Take time to decide whether or not you wish to take part.

Purpose of the study

- *Briefly explain the purpose of the research project*
- *When will the study be completed?*

Text: Background and aim of the study, expected completion…

If you agree to take part in this study, you will be asked to sign the Consent Form on the last page of this document. You will be given a copy of both the Participant Information Sheet and the Consent Form to keep.

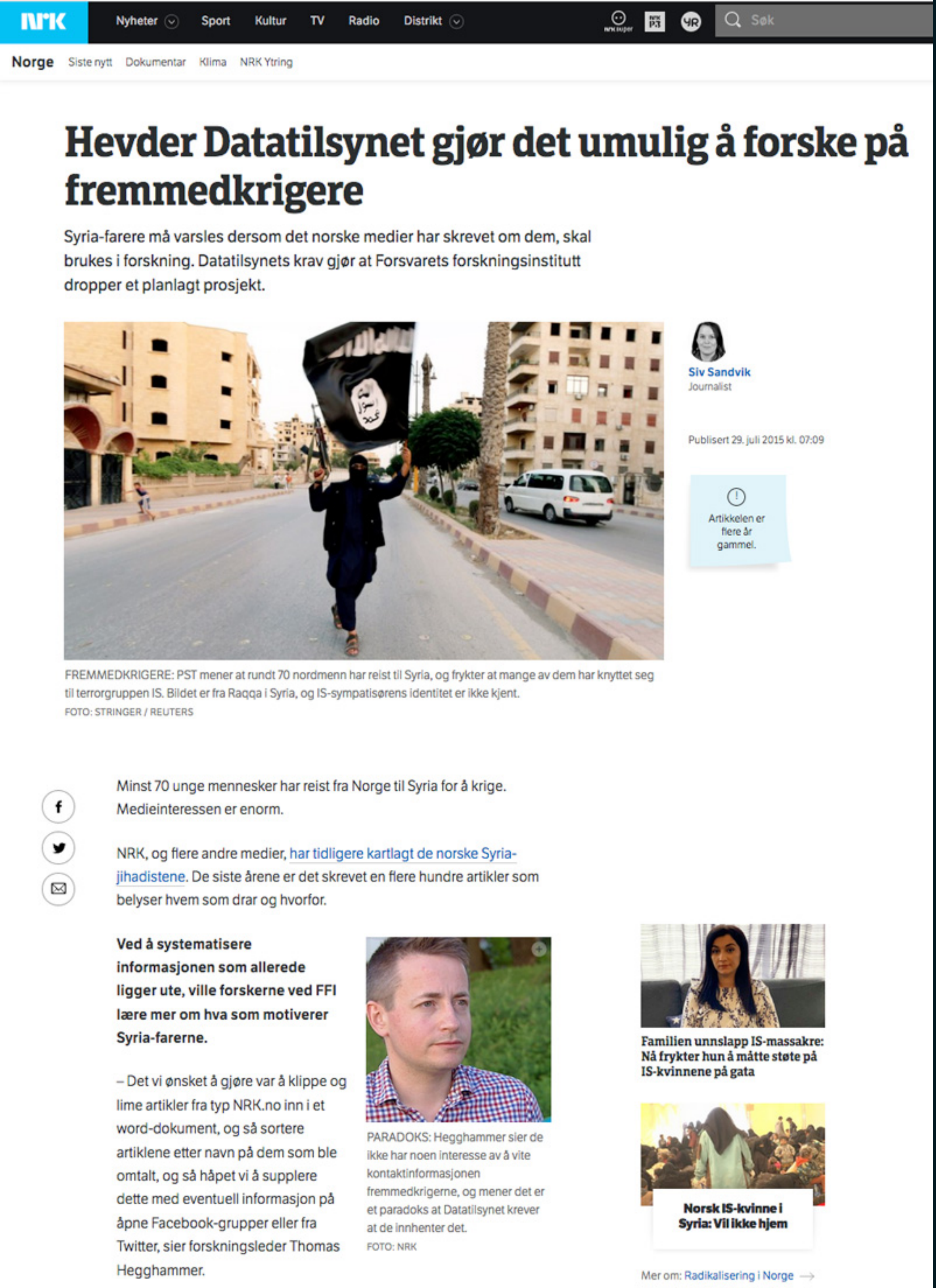
Why have I been chosen?

Consent - Art. 6(1)(a)

- The consent must be documented
  - Signed consent forms kept on file with the researcher
  - Recorded (audio): Can you prove who is talking?
- Store the consent along with and as long as you store the personal data
- Consent forms themselves may be sensitive; follow the UiO storage guide

Consent or public interest?

- Ethical and legal considerations
- Challenges - some examples
  1. Listhaug case: Consent withdrawn
  2. FFI study of foreign fighters: Public interest, but duty to inform made it practically impossible
  3. CREX: Extremists as vulnerable persons (Norwegian Data Protection Authority)
- Ethical Guidelines for Internet Research (English, Dec 2014) – Norwegian updated Sep 2018



Hevder Datatilsynet gjør det umulig å forske på fremmedkrigere

Syria-farere må varsles dersom det norske medier har skrevet om dem, skal brukes i forskning. Datatilsynets krav gjør at Forsvarets forskningsinstitutt dropper et planlagt prosjekt.

FREMMEDKRIGERE: PST mener at rundt 70 nordmenn har reist til Syria, og frykter at mange av dem har knyttet seg til terrorgruppen IS. Bildet er fra Raqqa i Syria, og IS-sympatisørens identitet er ikke kjent.
FOTO: STRINGER / REUTERS

Siv Sandvik
Journalist

Publisert 29. juli 2015 kl. 07:09

Artikkelen er flere år gammel.

Minst 70 unge mennesker har reist fra Norge til Syria for å krige. Medieinteressen er enorm.

NRK, og flere andre medier, har tidligere kartlagt de norske Syria-jihadistene. De siste årene er det skrevet en flere hundre artikler som belyser hvem som drar og hvorfor.

Ved å systematisere informasjonen som allerede ligger ute, ville forskerne ved FFI lære mer om hva som motiverer Syria-farerne.

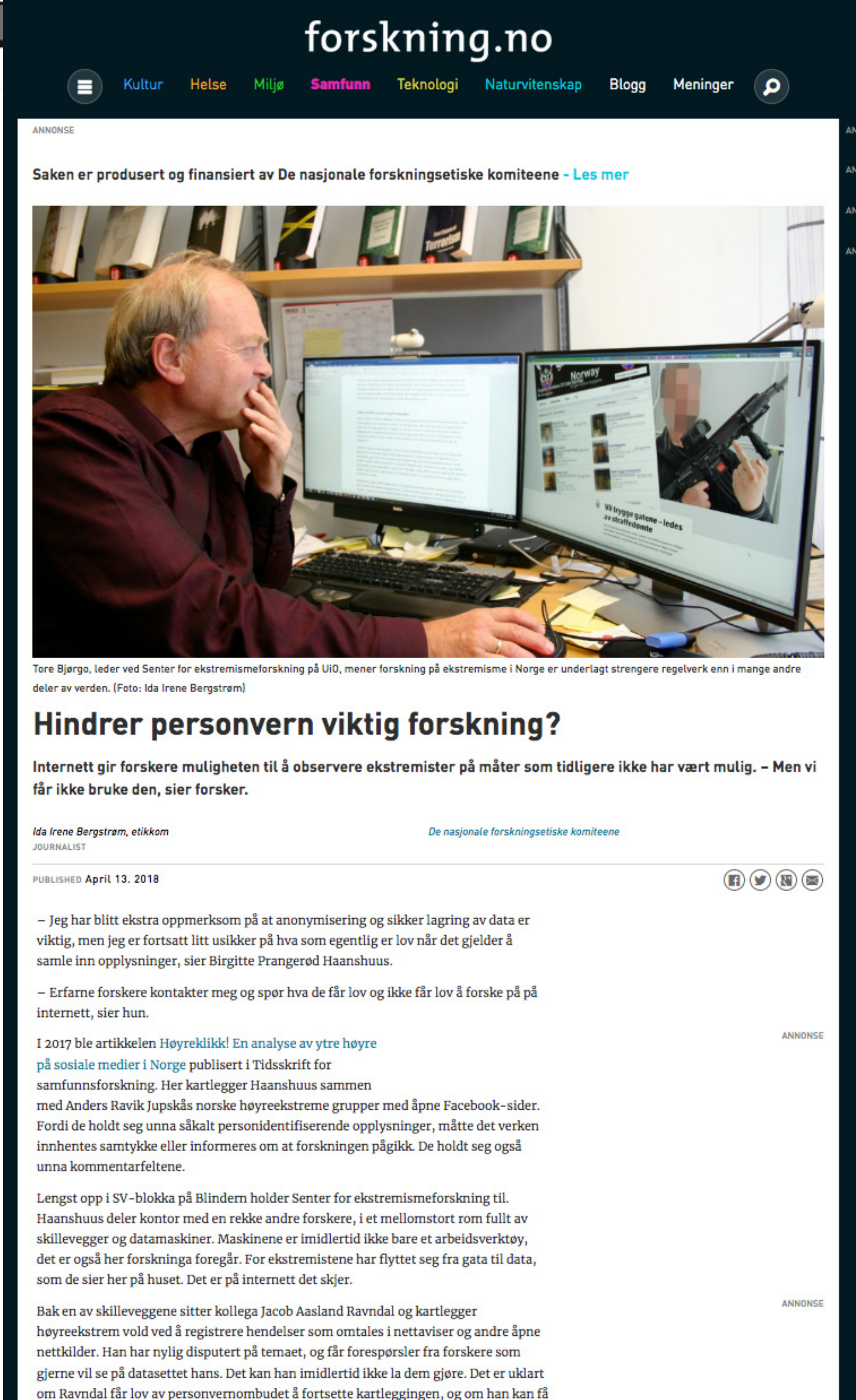
– Det vi ønsket å gjøre var å klippe og lime artikler fra typ NRK.no inn i et word-dokument, og så sortere artiklene etter navn på dem som ble omtalt, og så håpet vi å supplere dette med eventuell informasjon på åpne Facebook-grupper eller fra Twitter, sier forskningsleder Thomas Hegghammer.

PARADOKS: Hegghammer sier de ikke har noen interesse av å vite kontaktinformasjonen fremmedkrigerne, og mener det er et paradoks at Datatilsynet krever at de innhenter det.
FOTO: NRK

Familien unnslapp IS-massakre: Nå frykter hun å måtte støte på IS-kvinnene på gata

Norsk IS-kvinne i Syria: Vil ikke hjem

Mer om: Radikalisering i Norge



ANNONSE

Saken er produsert og finansiert av De nasjonale forskningsetiske komiteene - Les mer

Tore Bjørgo, leder ved Senter for ekstremismeforskning på UiO, mener forskning på ekstremisme i Norge er underlagt strengere regelverk enn i mange andre deler av verden. (Foto: Ida Irene Bergstrøm)

Hindrer personvern viktig forskning?

Internett gir forskere muligheten til å observere ekstremister på måter som tidligere ikke har vært mulig. – Men vi får ikke bruke den, sier forsker.

Ida Irene Bergstrøm, etikkom
JOURNALIST

De nasjonale forskningsetiske komiteene

PUBLISHED April 13. 2018

– Jeg har blitt ekstra oppmerksom på at anonymisering og sikker lagring av data er viktig, men jeg er fortsatt litt usikker på hva som egentlig er lov når det gjelder å samle inn opplysninger, sier Birgitte Prangerød Haanshuus.

– Erfarne forskere kontakter meg og spør hva de får lov og ikke får lov å forske på på internett, sier hun.

ANNONSE

I 2017 ble artikkelen Høyreklikk! En analyse av ytre høyre på sosiale medier i Norge publisert i Tidsskrift for samfunnsforskning. Her kartlegger Haanshuus sammen med Anders Ravik Jupskås norske høyreekstreme grupper med åpne Facebook-sider. Fordi de holdt seg unna såkalt personidentifiserende opplysninger, måtte det verken innhentes samtykke eller informeres om at forskningen pågikk. De holdt seg også unna kommentarfeltene.

Lengst opp i SV-blokka på Blindern holder Senter for ekstremismeforskning til. Haanshuus deler kontor med en rekke andre forskere, i et mellomstort rom fullt av skillevegger og datamaskiner. Maskinene er imidlertid ikke bare et arbeidsverktøy, det er også her forskninga foregår. For ekstremistene har flyttet seg fra gata til data, som de sier her på huset. Det er på internett det skjer.

ANNONSE

Bak en av skilleveggene sitter kollega Jacob Aasland Ravndal og kartlegger høyreekstrem vold ved å registrere hendelser som omtales i nettaviser og andre åpne nettkilder. Han har nylig disputert på temaet, og får forespørsler fra forskere som gjerne vil se på datasettet hans. Det kan han imidlertid ikke la dem gjøre. Det er uklart om Ravndal får lov av personvernombudet å fortsette kartleggingen, og om han kan få

What could possibly go wrong?

- Violations of privacy:
  - Sensitive data about research respondents/informants goes astray
  - Data on informants or respondents in research is used for entirely different and irreconcilable purposes than what they consented to

UNIVERSITAS Nyhet Kultur Magasin Debatt PDF Utgivelsespla

MILLIONBOT: På grunn av slurvete omgang med data, kan UiO måtte belage seg på flere millioner i bo
Daae

UiO risikerer millionbot

Datatilsynet mistenker at UiOs håndtering av sensitiv personda
helseforetakene bøter på til sammen 7,2 millioner kroner.

Nyhet
Datasikkerhet

fre 02. feb. 2018

Tekst:
Mikkel Ihle Tande

Les UiOs tilsvar her.

Etter en brevkontroll har Datatilsynet bedt UiO om å endre sine sikkerh
behandlingen av sensitive personopplysninger som fødselsnummer, in
menneskemengder og informasjon som behandlingsansvarlig har klas
hensyn til konfidensialitet.

– Hvis vi finner områder vi mener har en svakhet gir vi varsel om påleg
endringsforslag til et alvorlig sikkerhetshull, sier Veronica Jarnskjold B

Glemte å strø: *Da fikk UiO bot*

Behandler UiOs svar

Hun holder for tiden på å behandle UiOs svar i saken. Etter at saken er
Datatilsynet godtar UiOs redegjørelse eller om de pålegger universitete
avgjort er detaljene omkring hva pålegget gjelder foreløpig unntatt offe

> « Det kan dreie seg om alt fra et lite endring
> Veronica Jarnskjold Buer

Aftenposten
Fredag 18. januar 2019

Kultur | 5

Gyldendal Akademisk trekker bok

Sehesteds plass i Oslo med forlagshuset Gyldendal t.h. FOTO: TOR G. STENERSEN

Edel Bakkemoen

Boken *Helsefremmende skriving* er trukket tilbake fordi en av bidragsyterne mener hans tekst og personopplysninger ble brukt på en måte han ikke hadde gått god for.

Jeg ble forskrekket over plutselig å finne en del av en tekst jeg skrev på et kurs i skriving som terapi for seks år siden, satt inn i en historie jeg ikke kjenner meg igjen i. Her er det gitt både feilaktige og helsemessige personopplysninger, fortalte Ivar Kamsvåg til Aftenposten like før jul.

Han er lettet over at Gyldendal Akademisk har tatt sakprosaboken *Helsefremmende skriving* ut av salg. Boken er laget av Kjersti Wold og Gunn-Marit Uverud.

Vurderer å utgi en justert versjon

Uverud tok for flere år siden initiativet til to kurs i skriving som terapi, hvor hun var prosjektleder og Kjersti Wold var skrivepedagog. Kurset «Svart gull» ble holdt i årene 2010-12 med finansiering fra Nasjonalt senter for erfaringskompetanse innen psykisk helse. Deltagerne skrev personlige tekster om et vendepunkt i livet, og samtykket samtidig i at Gyldendal Akademisk kunne bruke tekstene. Men i boken som kom seks år senere, er Kamsvågs opprinnelige tekst forkortet og satt inn i en ny sammenheng. Han fikk ikke godkjenne det som gjaldt ham før boken ble utgitt.

Kamsvåg er en av 21 kursdeltagere som fikk sin tekst helt eller delvis gjengitt i boken.

- Boken er trukket fra markedet inntil videre. Vi tar nå kontakt med alle kursdeltagerne for å be om samtykke på nytt. Vi vil se an hva slags tilbakemeldinger og samtykker vi får før vi eventuelt endrer boken deretter. Hvis det kommer en ny utgave, blir det mest sannsynlig en justert versjon. Vi regner ikke med at Kamsvåg vil være representert der, sier redaksjonssjef Berit Gjendem i Gyldendal Akademisk.

Ivar Kamsvåg reagerte på måten hans seks år gamle tekst ble brukt på i boken *Helsefremmende skriving*.

Anonymiserte bidragsytere

Aftenposten er kjent med at flere av de øvrige bidragsyterne er godt fornøyde med prosessen og hvordan boken ble til slutt.

- Jeg føler at jeg har bidratt til en viktig bok og opplevde fremgangsmåten fra forfatterne som god, sier Audhild Hjellup Lønne.

Selv om alt er anonymisert i boken, hevder Kamsvåg at han er blitt gjenkjent. Bokens forfattere tilbød først Kamsvåg å endre det som er skrevet rundt hans tekst, samt å publisere hele kursteksten som et tillegg i neste opplag.

Beklager at det kom skjevt ut

Prosjektet skiftet retning og forlagsredaktører flere ganger i årenes løp. Uverud sier at hun ikke trodde det var nødvendig med ytterligere samtykker, da hun regnet med at den opprinnelige avtalen fra prosjektstart var gyldig.

- Det har aldri vært gitt løfter om bruk av uavkortede tekster i bokprosjektet. Jeg beklager hvis kommunikasjonen rundt dette ikke har vært tydelig nok. Som prosjektleder er jeg lei for at dette kom skjevt ut og håper at vi kan få til et nytt produkt som alle er fornøyd med, sier hun.

Gyldendal innrømmer at forlaget kunne fulgt opp bedre.

- **Har dere fulgt vanlig praksis her?**

- Dette er et unikt prosjekt. Når vi utgir medisinske bøker med anonymiserte pasienthistorier, innhentes tillatelse fra dem det gjelder. Her ble den jobben gjort av dem som holdt skrivekurset. Vi innhenter ikke samtykke så lenge forfatterne har gjort det, og når stoffet er anonymisert. Men det kan hende vi skulle tenkt annerledes, sier redaksjonssjef Gjendem.

What is Research Data Management?

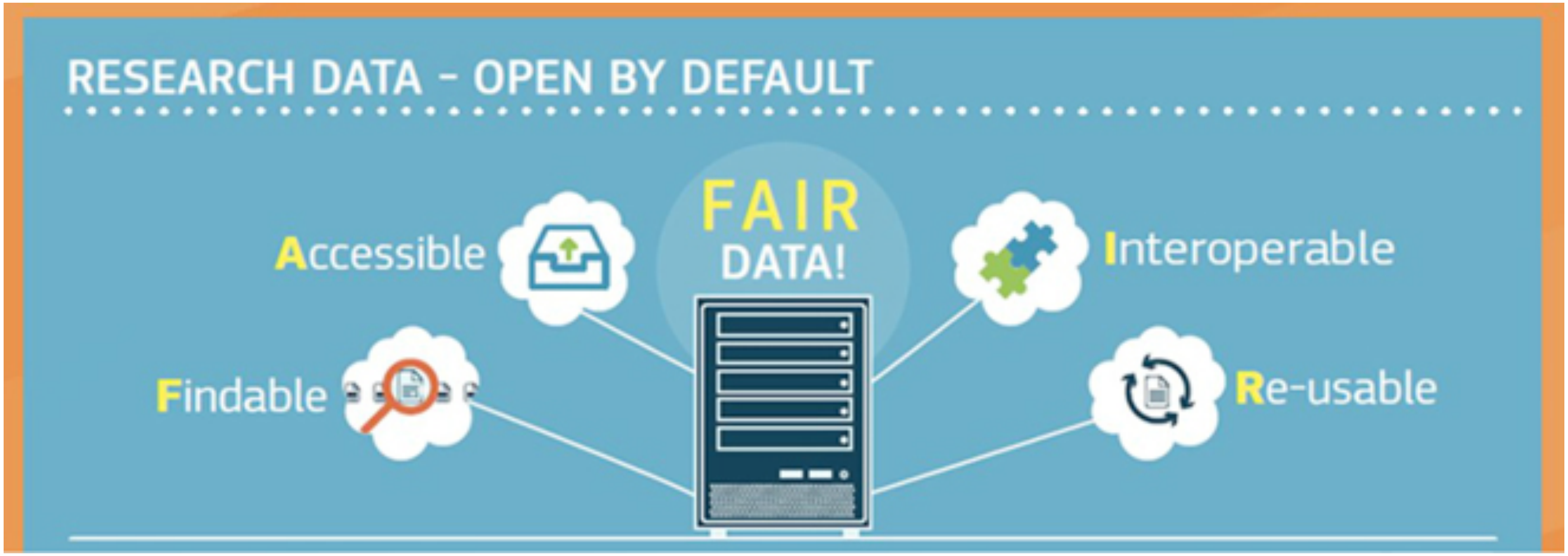
The active management of data throughout the lifecycle

- Data Management Planning
- Creating data
- Documenting data
- Accessing / using data
- Storage and backup
- Selecting what to keep
- Sharing data
- Data licensing and citation
- Preserving data
- ...

FOSTER

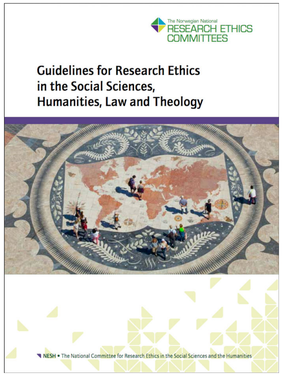
Source: Sarah Jones, Open Science Days 2015:
[FOSTER: Managing and sharing research data]

Principles of FAIR data management



Why share data?
Considerations to be made in individual cases, but need to explain if <u>not</u> share-worthy

Data sharing


The Norwegian National RESEARCH ETHICS COMMITTEES

Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology

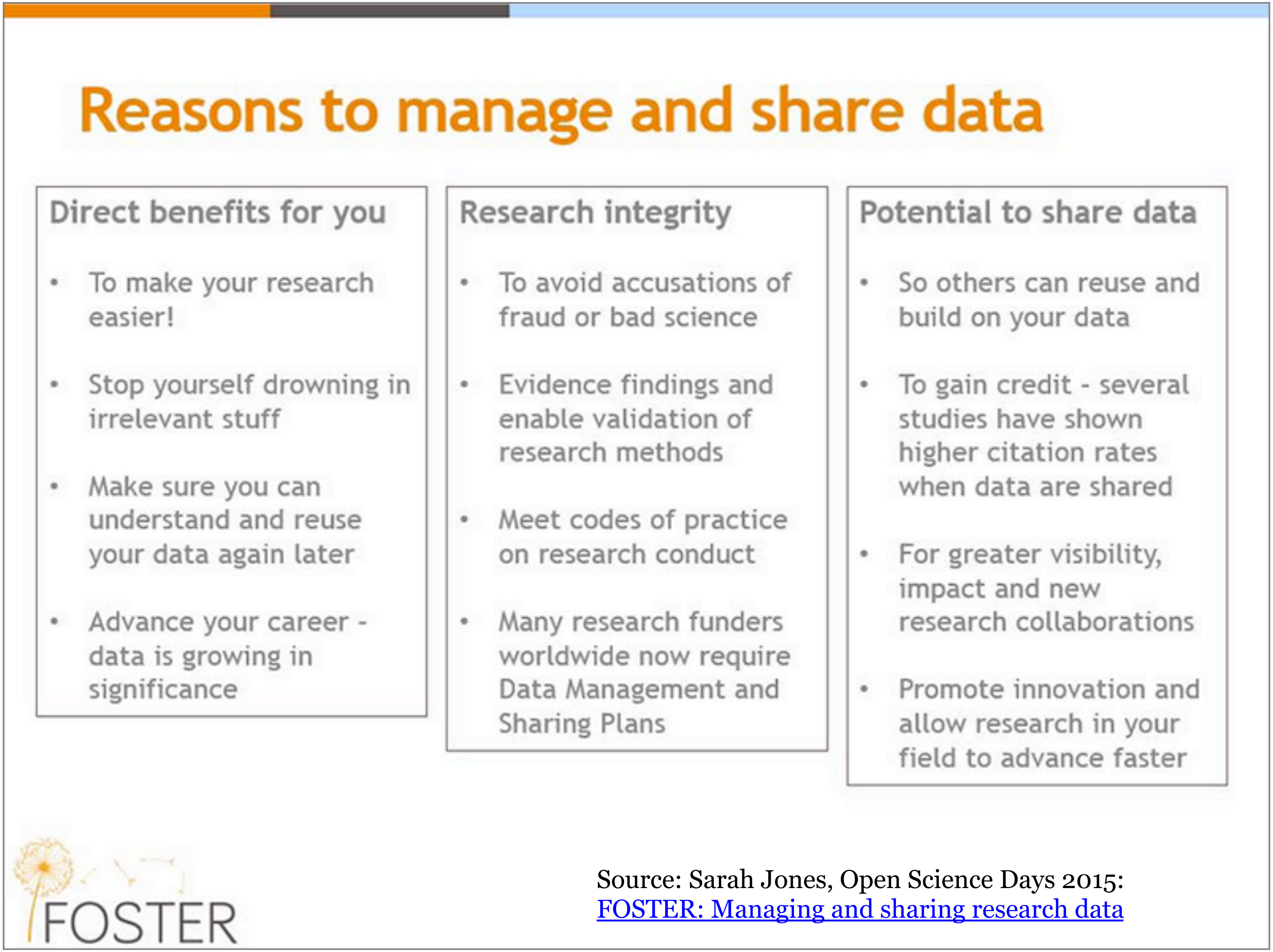
NESH • The National Committee for Research Ethics in the Social Sciences and the Humanities

Research material should be made available to other researchers for secondary analysis and further use. Sharing of research data is often a prerequisite for building up knowledge, comparing results and critically testing the work of others. Improved openness and quality assurance can be achieved by sharing data.

At the same time, data sharing gives rise to **ethical challenges relating to privacy and confidentiality**. Therefore, the norm of transparency and data-sharing, particularly in large-scale registry research, should be balanced against other considerations and requirements of research ethics.

Generally, those responsible for collecting material have the priority right to use it in analyses and in publications. Data acquired with the aid of public funding must be made publicly available after a short period.

NESH guidelines, p. 30 (art. 29)



Reasons to manage and share data

Direct benefits for you

- To make your research easier!
- Stop yourself drowning in irrelevant stuff
- Make sure you can understand and reuse your data again later
- Advance your career - data is growing in significance

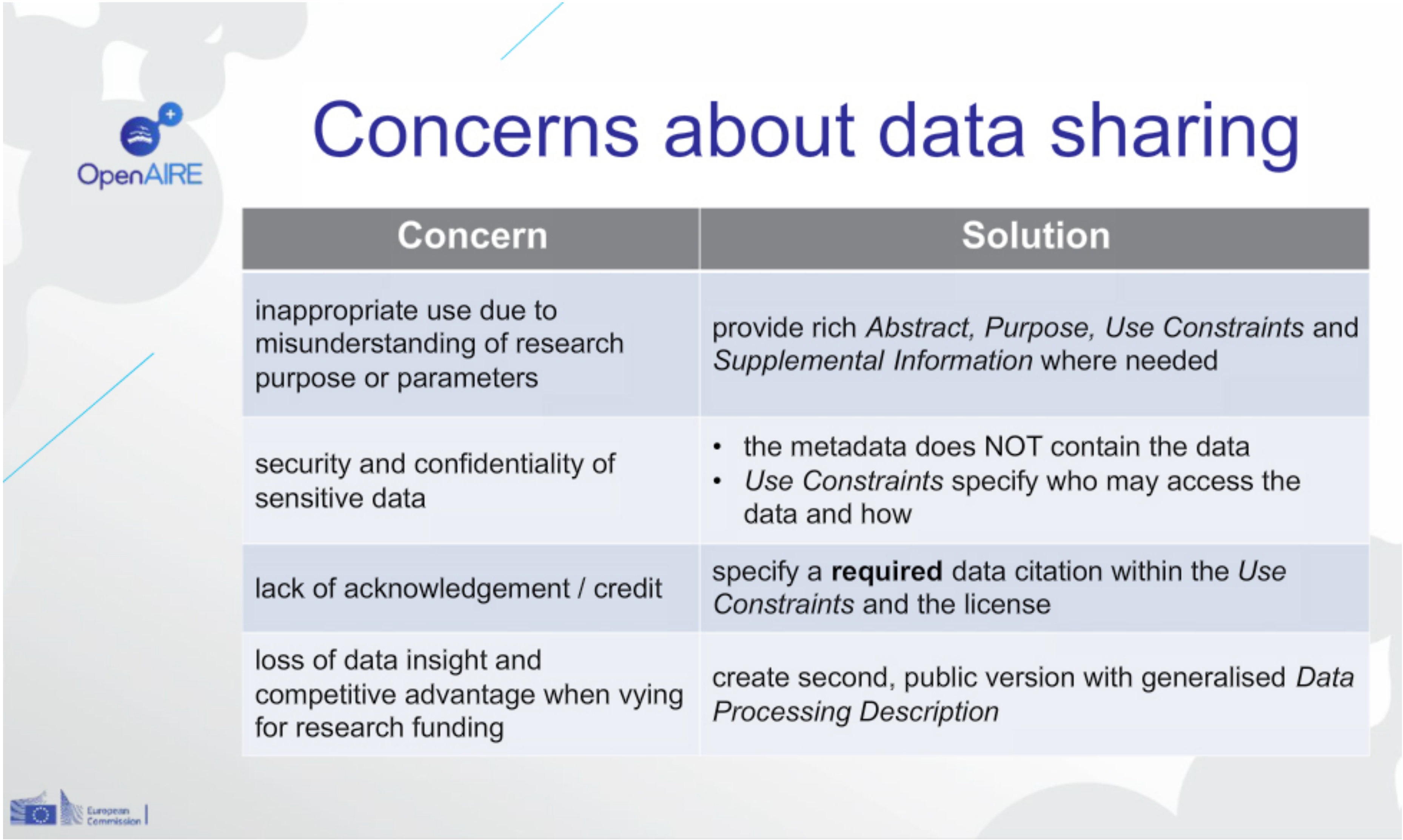
Research integrity

- To avoid accusations of fraud or bad science
- Evidence findings and enable validation of research methods
- Meet codes of practice on research conduct
- Many research funders worldwide now require Data Management and Sharing Plans

Potential to share data

- So others can reuse and build on your data
- To gain credit - several studies have shown higher citation rates when data are shared
- For greater visibility, impact and new research collaborations
- Promote innovation and allow research in your field to advance faster

Source: Sarah Jones, Open Science Days 2015:
[FOSTER: Managing and sharing research data]()



Concerns about data sharing

Concern	Solution
inappropriate use due to misunderstanding of research purpose or parameters	provide rich *Abstract, Purpose, Use Constraints* and *Supplemental Information* where needed
security and confidentiality of sensitive data	• the metadata does NOT contain the data • *Use Constraints* specify who may access the data and how
lack of acknowledgement / credit	specify a **required** data citation within the *Use Constraints* and the license
loss of data insight and competitive advantage when vying for research funding	create second, public version with generalised *Data Processing Description*

Source: OpenAIRE webinar on Open Research Data in Horizon 2020, 30 May 2017 (Marjan Grootveld, DANS)

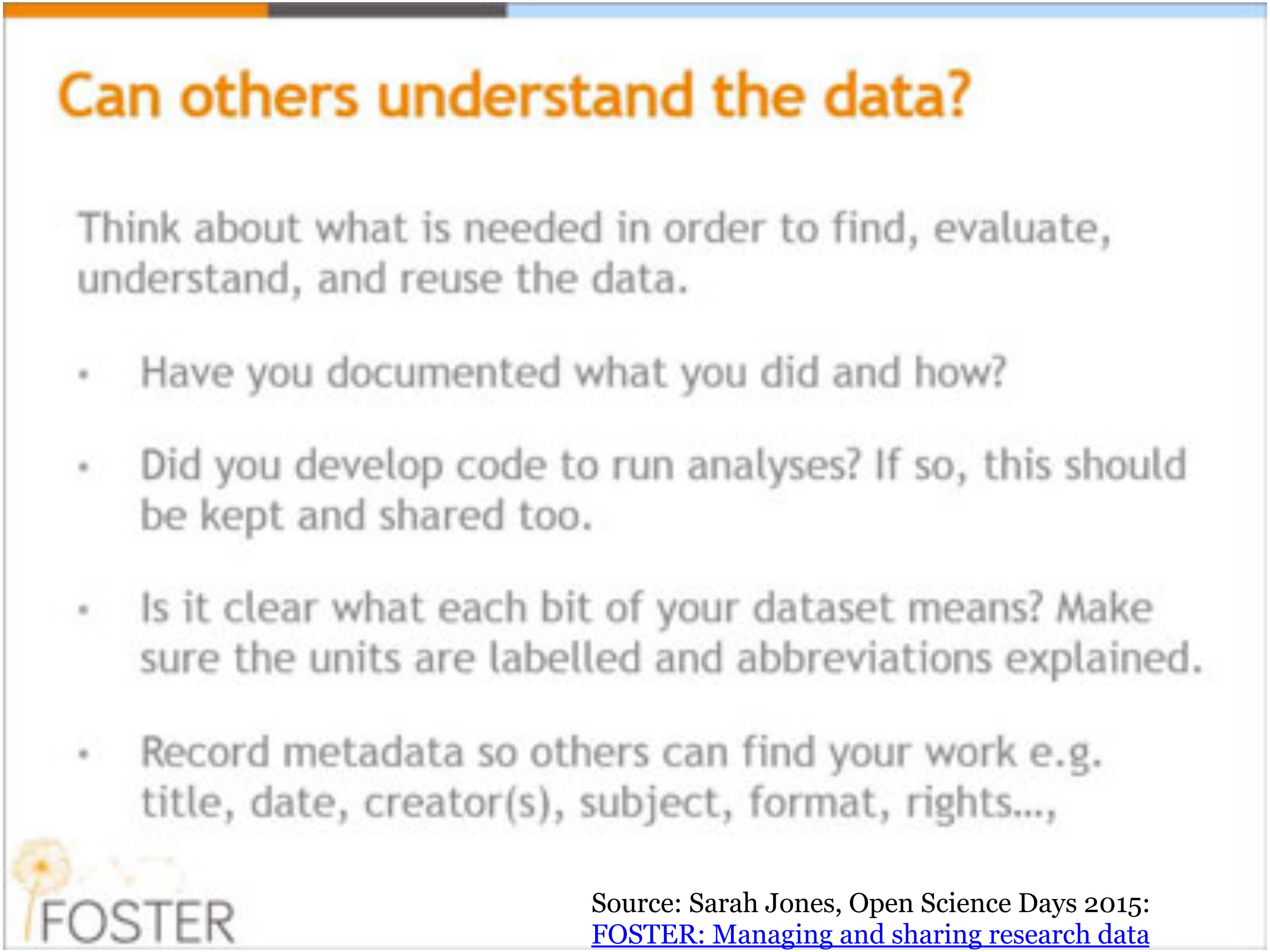
Data Management Plans

- Requirement for H2020 collaborative projects (GLOBUS, PLATO, EU3D)
  - Overall plan with dataset descriptions (made by individual researchers, one description per dataset)
  - Examples and templates available (ARENA intranet)
- NRC projects
  - New requirements (BENCHMARK, LEGOF…)

Data Management Plan contents

- Data types
- Metadata
- Storage and archiving
- Sharing

See Managing and sharing research data web page (ARENA intranet) for links and further information



Can others understand the data?

Think about what is needed in order to find, evaluate, understand, and reuse the data.

- Have you documented what you did and how?
- Did you develop code to run analyses? If so, this should be kept and shared too.
- Is it clear what each bit of your dataset means? Make sure the units are labelled and abbreviations explained.
- Record metadata so others can find your work e.g. title, date, creator(s), subject, format, rights...,

FOSTER

Source: Sarah Jones, Open Science Days 2015:
[FOSTER: Managing and sharing research data]

Coming up…

- Ask an expert lunch meeting – 2 April 2019
    - UiO's data protection officer (Maren Magnus Voll)
    - IT experts (processing, storage etc)
    - Send your questions to Marit!
- Report missing information or own experiences to revise the guidelines
- FAQ page for ARENA researchers
- ARENA support: Marit (personal data protection)