



Requirements

from funding bodies & institutions



Korbinian Bösl
Data manager ELIXIR Norway
4 March 2020



19/12/2017

National strategy on access to and sharing of research data

Research data must be as open as possible, as closed as necessary

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

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

...

The Government expects:

the research institutions to develop procedures for (i) approving data management plans

...



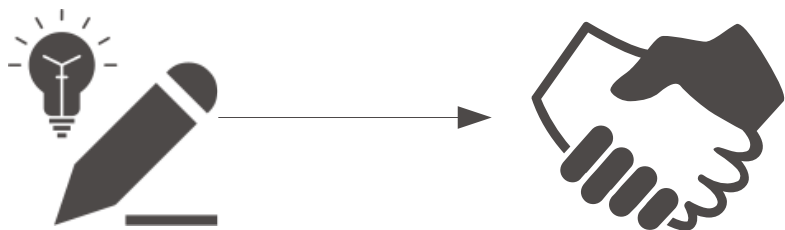


Forskningsrådet





Forskningsrådet



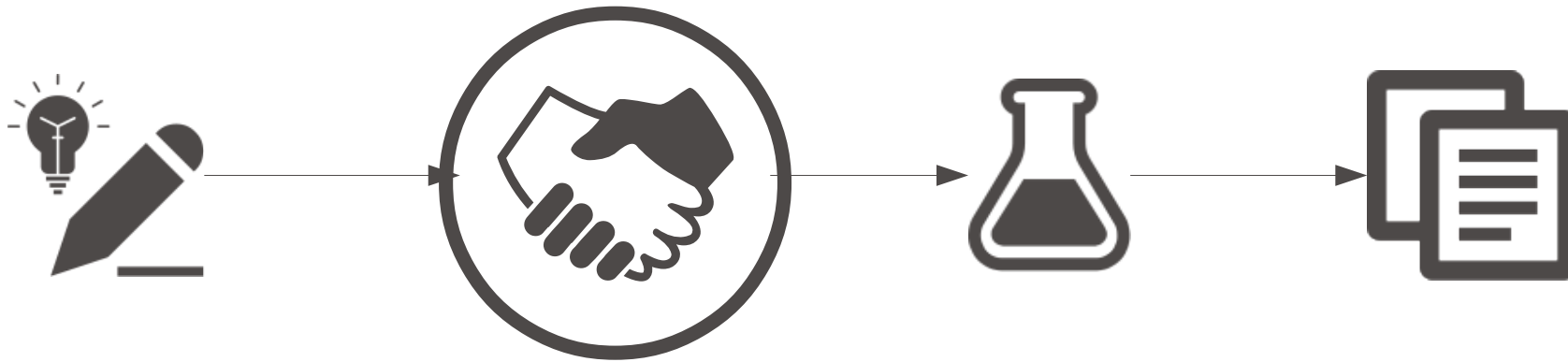


Forskningsrådet





Forskningsrådet



DMP

SCIENCE
EUROPE





Forskningsrådet

- Description & Re-use
- Documentation & Data quality
- Storage & Backups
- Legal & Ethic requirements
- Sharing & Preservation
- Responsibilities & Resources (FAIR principles)

SCIENCE
EUROPE



https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

https://www.scienceeurope.org/media/nsxdyvqn/se_guidance_document_rdmps.pdf

<https://www.forskningsradet.no/en/forskningspolitisk-radgivning/open-science/open-access-to-research-data/>



Description & Re-use

How will new data be collected or produced and/or how will existing data be re-used



Description & Re-use

How will new data be collected or produced and/or how will existing data be re-used

What data (for example the kinds, formats, and volumes) will be collected or produced?



Documentation & Data quality

What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?



Documentation & Data quality

What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?

What data quality control measures will be used?



Storage and backup during the research process

How will data and metadata be stored and backed up during the research process?



Storage and backup during the research process

How will data and metadata be stored and backed up during the research process?



Storage and backup during the research process

How will data and metadata be stored and backed up during the research process?

How will data security and protection of sensitive data be taken care of during the research?



Legal and ethical requirements, codes of conduct

If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?



Legal and ethical requirements, codes of conduct

If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?

How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?



Legal and ethical requirements, codes of conduct

If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?

How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

How will possible ethical issues be taken into account, and codes of conduct followed?



**SCIENCE
EUROPE**



Data sharing and long-term preservation

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?



Data sharing and long-term preservation

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?



Data sharing and long-term preservation

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?

What methods or software tools will be needed to access and use the data?



Data sharing and long-term preservation

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?

What methods or software tools will be needed to access and use the data?

How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

**Open, non-binary
data formats!**



**SCIENCE
EUROPE**



Data management responsibilities and resources

Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?



Data management responsibilities and resources

Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?



NeLS

Norwegian e-Infrastructure for Life Sciences

BioStudies.

ENA
European Nucleotide Archive

IDR

ArrayExpress

MetaboLights

Proteomics Identifications Database

EUROPEAN
GENOME-PHENOME
ARCHIVE

DSW

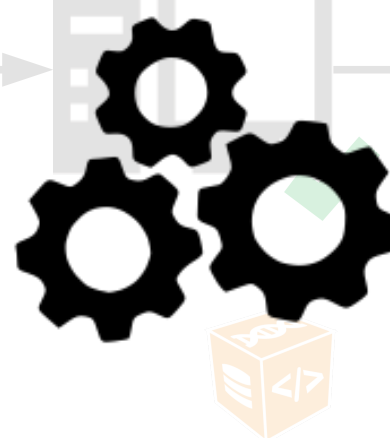
SEEK

Galaxy

F indelible A ccessible I nteroperable R euseable



sensitive
data



OPEN

SUBJECT CATEGORIES

» Research data
» Publication
characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*[#]

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Received: 10 December 2015

Accepted: 12 February 2016

Published: 15 March 2016

FAIR Guidelines

To be Findable:

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

FAIR Guidelines

To be Accessible:

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

FAIR Guidelines

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

FAIR Guidelines

To be Reusable:

R1. meta(data) are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards



Horizon 2020 → Horizon Europe

Open Research Data Pilot (ORD pilot)
Default since 2017 → mandatory 2021



DMP
outline



DMP

- After 6 month
- Update upon significant changes





Horizon 2020 → Horizon Europe

Data summary

FAIR data implementation

Resources

Data security

Ethics



Explicit additional requirements - naming

What [file] naming conventions do you follow?

Will search keywords be provided that optimize possibilities for re-use?



Explicit additional requirements - software

Is documentation about the software needed to access the data included?

Is it possible to include the relevant software (e.g. in open source code)?



Explicit additional requirements - access

If there are restrictions on use, how will access be provided?

Is there a need for a data access committee?

Are there well described conditions for access (i.e. a machine readable license)?

How will the identity of the person accessing the data be ascertained?



Explicit additional requirements - resources

What are the costs for making data FAIR in your project?

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?



Explicit additional requirements - security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Is the data safely stored in certified repositories for long term preservation and curation?

