Open Science

Karel Luyben

Ministry Rome
12 March 2019
Open Science represents a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools.

Commissioner Carlos Moedas

Open Innovation, Open Science, Open to the World (2015)
European Open Science Agenda:

1. Rewards and incentives
2. Next-generation-metrics
3. FAIR (open) data
4. European Open Science Cloud
5. Future scholarly communication (Open Access)
6. Citizen Science
7. Research integrity
8. Open science skills
9. Educating data stewards / scientists
10. New models for publishing
SCIENCE 1.0

Towards FAIR Data
Towards Open Access
Citizen Science

skills rewards metrics

Stimulation and support

OPEN SCIENCE

SCIENCE 2.0
Development of Open Science

Completion of open science over time:

- 2017: Initial stage
- 2020: Early development
- 2030: Significant growth
- 2040: Continuous increase
- 2050:接近100%完成
Amount

Early stage

Usefulness

Mature stage

Curiosity

Research

Practice

(Grand) challenges / experiences

Information (publications; data; designs)

‘Driving force’

‘Life cycle’
“After 1 January 2020 scientific publications on the results from research funded by public grants provided by national and European research councils and funding bodies, must be published in compliant Open Access Journals or on compliant Open Access Platforms.”
FAIR = Findable; Accessible; Interoperable; Reusable
FAIR DATA PRINCIPLES

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

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

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

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

https://doi.org/10.1038/sdata.2016.18
What FAIR is not …

Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud DOI: 10.3233/ISU-170824

FAIR is not a standard, it’s a guiding principle

FAIR is not only ‘Semantic Web’

FAIR does not mean ‘Open’ or ‘Free’

Data are often Open but not FAIR

Data could be non-Open yet perfectly FAIR

Many data can never be Open

FAIR principles do not directly prescribe data quality, trustworthiness, ethics or responsibilities.
DIGITAL OBJECT
Data, code and other research outputs
At its most basic level, data or code is a bitstream or binary sequence. For this to have meaning and to be FAIR, it needs to be represented in standard formats and be accompanied by Persistent Identifiers (PIs), metadata and documentation. These layers of meaning enrich the object and enable reuse.

IDENTIFIERS
Persistent and unique (PIs)
Digital Objects should be assigned a unique and persistent identifier such as a DOI or URN. This enables stable links to the object and supports citation and reuse to be tracked. Identifiers should also be applied to other related concepts such as the data authors (ORCIDs), projects (RAIDs), funders and associated research resources (RRIDs).

STANDARDS & CODE
Open, documented formats
Digital Objects should be represented in common and ideally open file formats. This enables others to reuse them as the format is in widespread use and software is available to read the files. Open and well-documented formats are easier to preserve. Data also need to be accompanied by the code use to process and analyse the data.

METADATA
Contextual documentation
In order for Digital Objects to be assessable and reusable, they should be accompanied by sufficient metadata and documentation. Basic metadata will enable data discovery, but much richer information and provenance is required to understand how, why, when and by whom the objects were created. To enable the broadest reuse, they should be accompanied by a plurality of relevant attributes and a clear and accessible usage license.
Dutch aims for the phase till 2020

- Open Access for basically all publications (including books).
- An agreed vision on the ‘data-services-infrastructure’.
- Clearity if a support structure is wanted/needed for ‘Citizen Science’.
- Open Science will be explicit in the appraisal of scientific staff (part of HR-policy) in Dutch research performing organisations; in the evaluation of research proposals by funders; and in the evaluation criteria used by research evaluation committees in higher education.
- Researchers will be stimulated and supported to work on/with Open Science.
Coordination team ~ 40 ‘data stewards’

Countrywide expert network
~ 400 institutional ‘data stewards’

e-Infrastructure and e-science experts at
SURF, NLeSC, DANS plus institutional
ICT/data centres
e-infrastructure and e-science experts at SURF, NLeSC, DANS plus institutional ICT/data centres

NL Data Services Infrastructuur MS
- Coordinating national Health-RI team
  ~ 10 ‘data-stewards’ (BBMRI/ELIXIR)
- Network of ‘data stewards’ in institutes & infrastructures
- Web of (FAIR) health/biomedische data collecties

MS domein (Health-RI)
e-infrastructure and e-science experts at SURF, NLeSC, DANS plus institutional ICT/data centres

**NL Data Services Infrastructuur SSH**
- Coordinating national team
  ~ 10 ‘data-stewards’ (CLARIAH/ODISSEI, … )
- Network of ‘data stewards’ in institutes & infrastructures
- Web of (FAIR) data collections in SSH domain
The EOSC is projected to become a reality by 2020 and will be Europe’s virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes.
Dutch relevant issues

- Open access support
- Data services infrastructure
- Data-stewards (curriculum; programmes)
- Step-up in effort (‘shift gear’)
- Financing the transition
Confidence in the future of open science in the Netherlands

“Open science and open access become the norm in scientific research.”

Coalition Agreement
2017 - 2021
BACK-UP