FAIR, the EOSC vehicle to efficiently use the data infrastructure

EOSC in Progress, September 20th 2018, Amsterdam

Melanie Imming, SURFsara
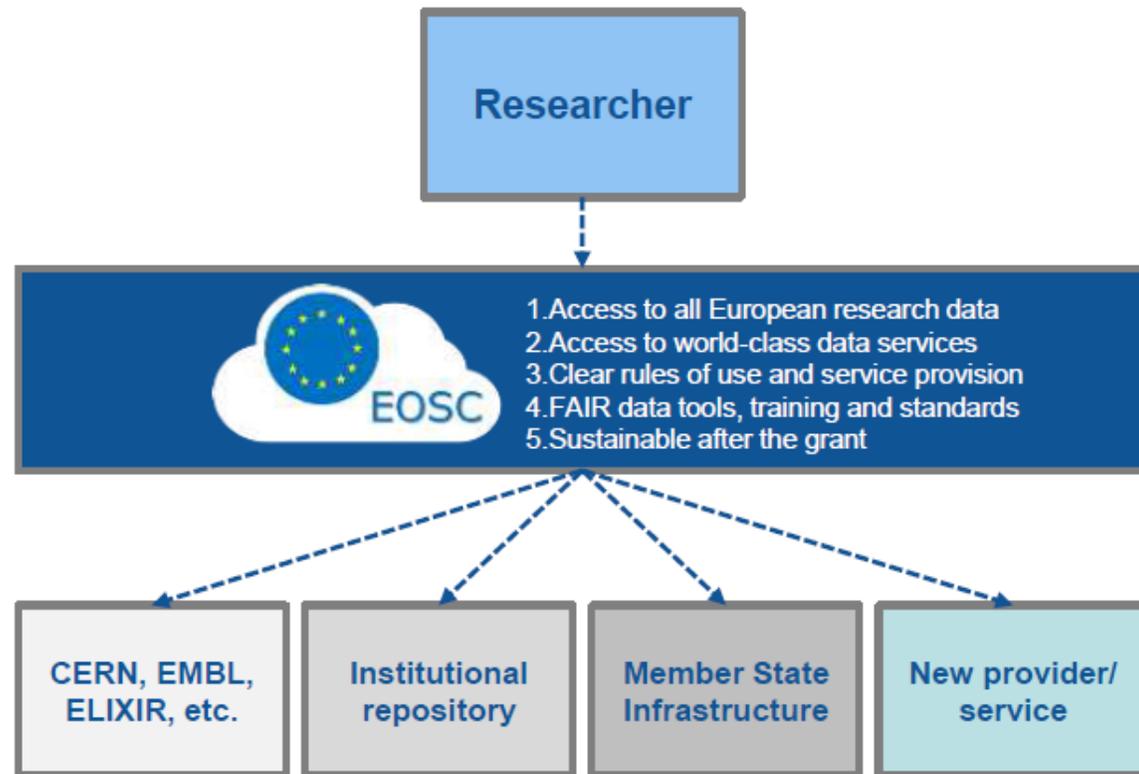
@melimming

FAIR Data Advanced Use Cases: from principles to practice in the Netherlands
zenodo.org/record/1250535

European
Commission

A. The EOSC will allow for universal access to data and a new level playing field for EU researchers

Researcher

1. Access to all European research data
2. Access to world-class data services
3. Clear rules of use and service provision
4. FAIR data tools, training and standards
5. Sustainable after the grant

EOSC

CERN, EMBL, ELIXIR, etc.

Institutional repository

Member State Infrastructure

New provider/ service

o Easy access through a universal access point for ALL European researchers

o Cross-disciplinary access to data unleashes potential of interdisciplinary research

o Services and data are interoperable (FAIR data)

o Data funded with public money is in principle open (as open as possible, as closed as necessary)

o EOSC will help increase recognition of data intensive research and data science

Seamless environment, enabling interdisciplinary research

European Commission

A. The EOSC will allow for universal access to data and a new level playing field for EU researchers

Researcher

1. Access to all European research data
2. Access to world-class data services
3. Clear rules of use and service provision
4. FAIR data tools, training and standards
5. Sustainable after the grant

CERN, EMBL, ELIXIR, etc.

Institutional repository

Member State Infrastructure

New provider/ service

- o Easy access through a universal access point for ALL European researchers
- o Cross-disciplinary access to data unleashes potential of interdisciplinary research
- o Services and data are interoperable (FAIR data)
- o Data funded with public money is in principle open (as open as possible, as closed as necessary)
- o EOSC will help increase recognition of data intensive research and data science
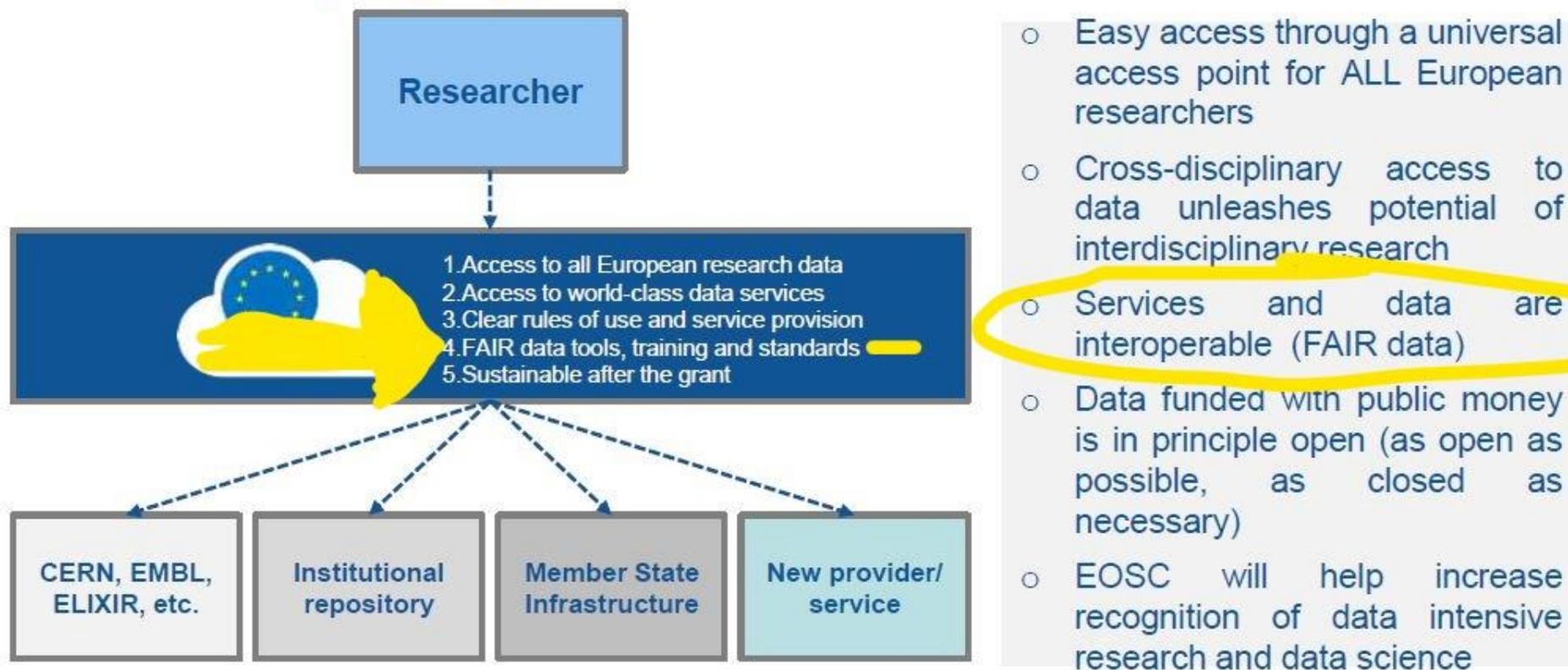
Seamless environment, enabling interdisciplinary research

THE FAIR DATA GUIDING 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

FAIR

UNFAIR

FAIR

UNFAIR

nope

FAIR DATA
ADVANCED USE CASES

FROM PRINCIPLES TO PRACTICE
IN THE NETHERLANDS

(PRELIMINARY) REPORT
MAY 2018
WWW.SURF.NL

SURF

The report

Describes six different approaches within different domains and within a number of projects, institutes and university libraries.

Examples, not instructions!

CC by SURFsara

Conclusions

Implementing FAIR is seen as a series of improvements;
- There are always steps ahead that can improve reuse even further.

FAIR is seen as part of a larger culture change;
- As part of good research data management or data stewardship;
- Towards more openness in research and interdisciplinary cooperation;
- Can't be seen separately of new privacy regulations and policies.

Conclusions

There is a tension between domain specific needs and maximum interoperability;

    No matter the FAIR maturity of the community

    Solutions:

- Trying to get consensus on minimal cross domain standards
- Sharing FAIR examples from different domains to get an understanding of the potential to align standards and workflows

Policies can't be about FAIR compliance alone;

    Any policy in which FAIR is mentioned should be open to discipline specific solutions

The Future

A way forward: integrated approaches with domain specific guidance;

To make it as easy as possible for researchers, the FAIR data principles need to be translated into more practical guidance. There is a tendency to take an integrated approach when doing so, in which domain-specific needs are leading.

FAIR takes effort, but it is worth it!

It takes effort to get to a certain level of FAIRness, but some use cases show that once you have reached that level of FAIRness, a whole world of possibilities opens.