Update on F.A.I.R.

For IWSG Science Gateways

16 October 2017

RDS
www.rds.edu.au

ands
AUSTRALIAN NATIONAL DATA SERVICE

nectar

Working on four transformations

- Building a data advantage
- Innovative approaches and tools
- Increase (inter)national collaboration
- Translating research outcomes

Requires FAIR data

What are the FAIR data principles?

Findable

(Identifiers, discovery metadata, catalogs)

Accessible

(Access mechanisms, services over the data)

Interoperable

(Standard formats and vocabs, link to code)

Reusable

(Licensed, Provenance information, contextual metadata)

Working on FAIR Data Partnerships 2017-18

Through general awareness raising, policy work, through projects, engagement with community, data providers and data and tool producers

- Policies and procedures
(e.g. Research Data management policies and plans)
- Technical infrastructure
(e.g. repositories, virtual labs, data clouds)
- Human infrastructure
(e.g. trained support staff)

FAIR data in Science Gateways

- Make it easy for researchers to make their data FAIR
- By linking into National Data services/initiatives

  Findable: DOI/Handle minting, National Data Discovery Service
  Accessible: Data services IG, advice on ethics
  Interoperable: Vocab service
  Reusable: advice on licensing, Provenance IG
- Raising awareness and skills

Making the software itself FAIR?

The FAIR Guiding Principles	OSS recommendations
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	"R2. Make software easy to discover by providing software metadata via a popular community registry" aligns with the Findability principle, helping to increase visibility and helping software providers to think about how to describe software metadata (versions, identifiers, contributors, citations, etc.)
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	"R1. Make source code publicly accessible from day one" focuses on openness including accessibility. The FAIR accessible principle instead opens the door to data that is restricted access e.g. for privacy reasons. Since such reasons do not apply for software, the OSS recommendations prefer to direct towards openness instead, supporting open science to the maximum extent.
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	This OSS recommendations do not aim to address software interoperability directly but contribute to a more homogenous description of software by encouraging software providers to register software metadata into registries providing specific metadata 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	"R3. Adopt a license and comply with the licence of third-party dependencies" aligns with the Reusability principle, helping to define to what extent the source code can be used and reused by the community, as a standalone software or as part of other software. Open availability of tools and libraries working with data formats can be a great help in making data interoperable: e.g. reuse of the same tools to read and write data can prevent subtle interoperability problems. Reproducibility of experiments and reuse of data is facilitated by the open availability of the associated software which is part of the provenance. All of the OSS recommendations thereby facilitate data Reusability.

From: Wilkinson et al 2016, Four simple recommendations to encourage best practices in research software

Does it make sense to apply FAIR to software?

- Yes, for the metadata of the software and identifiers
- Licensing
- Accessible and Interoperable are different

Things that are not covered

- Governance around the development
- Sustainability and community
- Machine readable?

Software is critical in making the data reusable

FAIR related international initiatives

Research Data Alliance

GoFAIR
(GoChange, GoBuild, GoTrain)

FAIR metrics

More information on FAIR

www.ands.org.au/fair

AUSTRALIAN NATIONAL DATA SERVICE

ands.org.au

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program. Monash University leads the partnership with the Australian National University and CSIRO.

Keith Russell

keith.russell@ands.org.au