Open Science and FAIR data

NI4OS-Europe Dissemination Event

Prof. Vladimir Trajkovik

November 6, 2020

Who am I?

Vladimir Trajkovikj (trvlado@finki.ukim.mk)
Faculty of Computer Science and Engineering, R. N. Macedonia

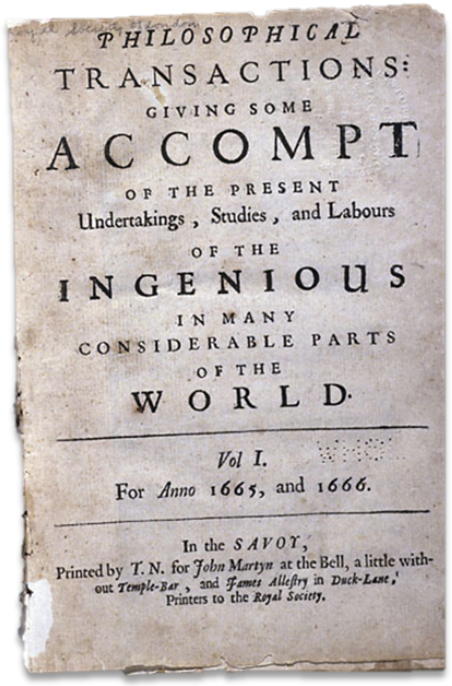
Special thanks for content sources to:
René van Horik & Cees Hof, Data Archiving and Networked Services (DANS), The Netherlands
Prof. Nikola Stikov, Polytechnique Montréal, Canada
Prof. Roberto Barbera – University of Catania – Italy

Special thanks for infrastructure and NI4OS data to:
Prof. Atanas Mishev, FCSE, R.N. Macedonia
Prof Sonja Filiposka, FCSE, R.N. Macedonia

AGENDA for Today

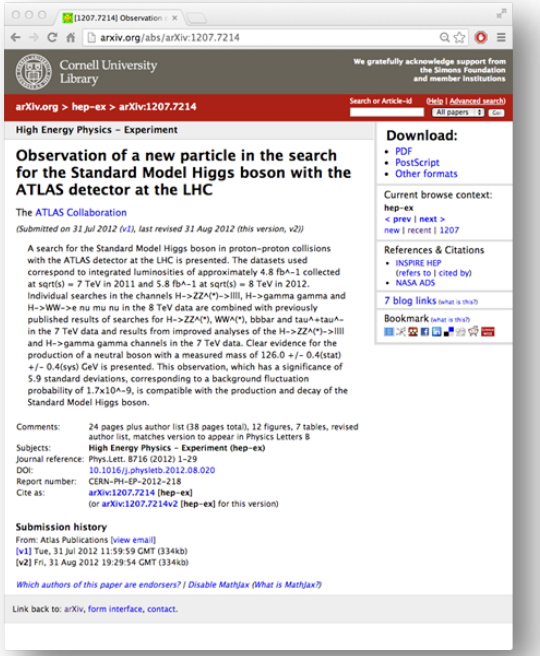
- Scientific challenges today
- Open Science
  - FAIR data and principles
  - Next steps
- Conclusion & Discussion

Scientific output: Challenge 1


PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane, Printers to the Royal Society.



It is the same since almost 4 centuries!

Scientific Foundation

Repeatability

- The closeness of agreement between independent results obtained with the same method on identical test material, under **the same conditions** (same operator, same apparatus, same laboratory and after short intervals of time)
- Affected by random errors

Reproducibility

- The closeness of agreement between independent results obtained with the same method on identical test material but **under different conditions** (different operators, different apparatus, different laboratories and/or after different intervals of time)
- Affected by systematic errors

Scientific Foundation

Repeatability

- The closeness of agreement bet… results obtained with the same method on identic… ame conditions (same operator, same … ter short intervals of time)
- Affe…

Is science really reproducible?

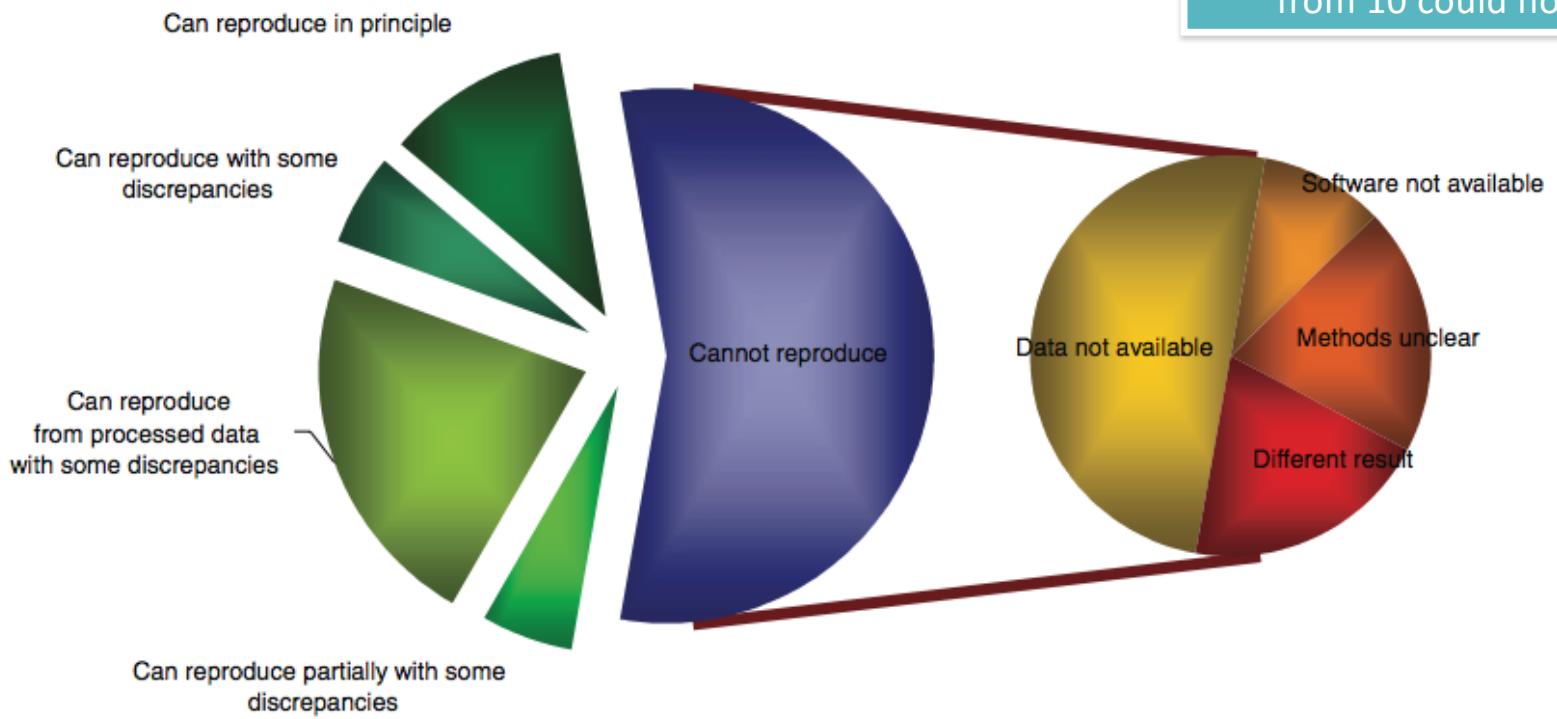
… closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or after different intervals of time)

- Affected by systematic errors

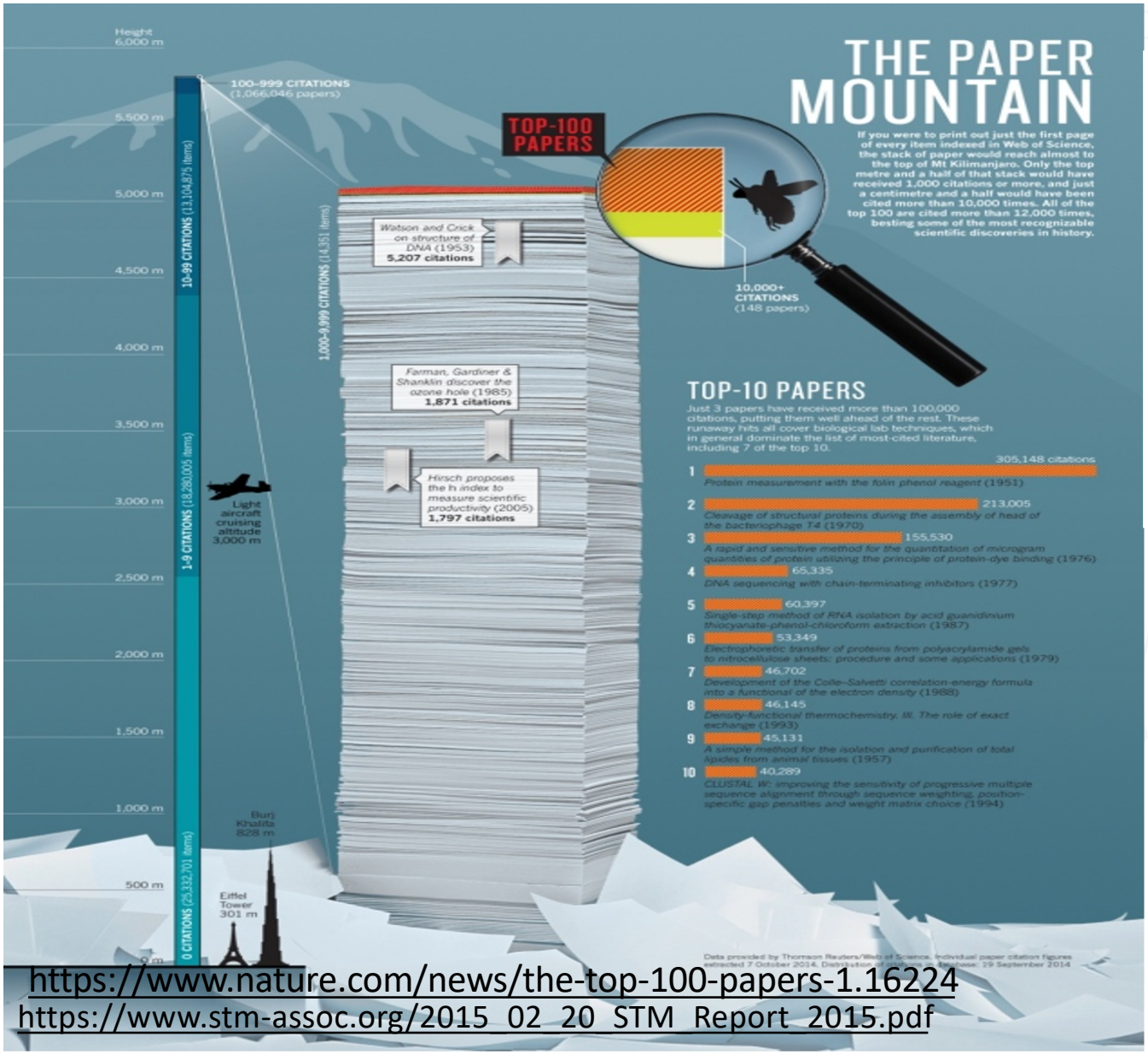
Challenge 2: Scientific Reproducibility

Challenge 2: Scientific Reproducibility

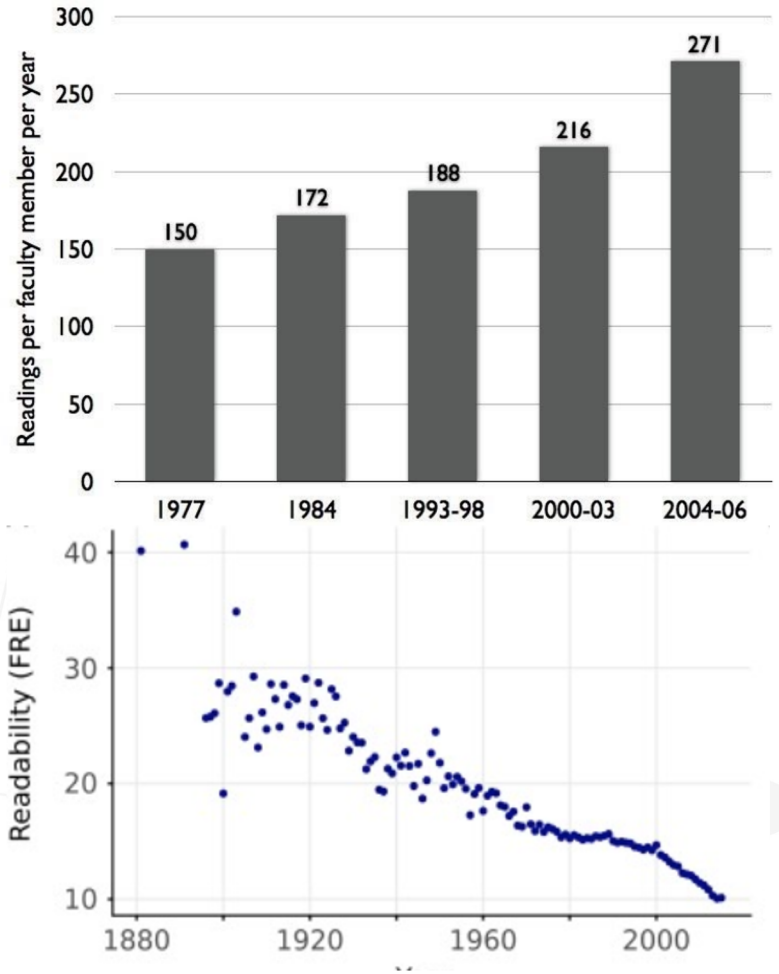
Out of 18 microarray papers, results from 10 could not be reproduced



1. Ioannidis et al., 2009. Repeatability of published microarray gene expression analyses. *Nature Genetics* 41: 14
2. Science publishing: The trouble with retractions http://www.nature.com/news/2011/111005/full/478026a.html
3. Bjorn Brembs: Open Access and the looming crisis in science https://theconversation.com/open-access-and-the-looming-crisis-in-science-14950

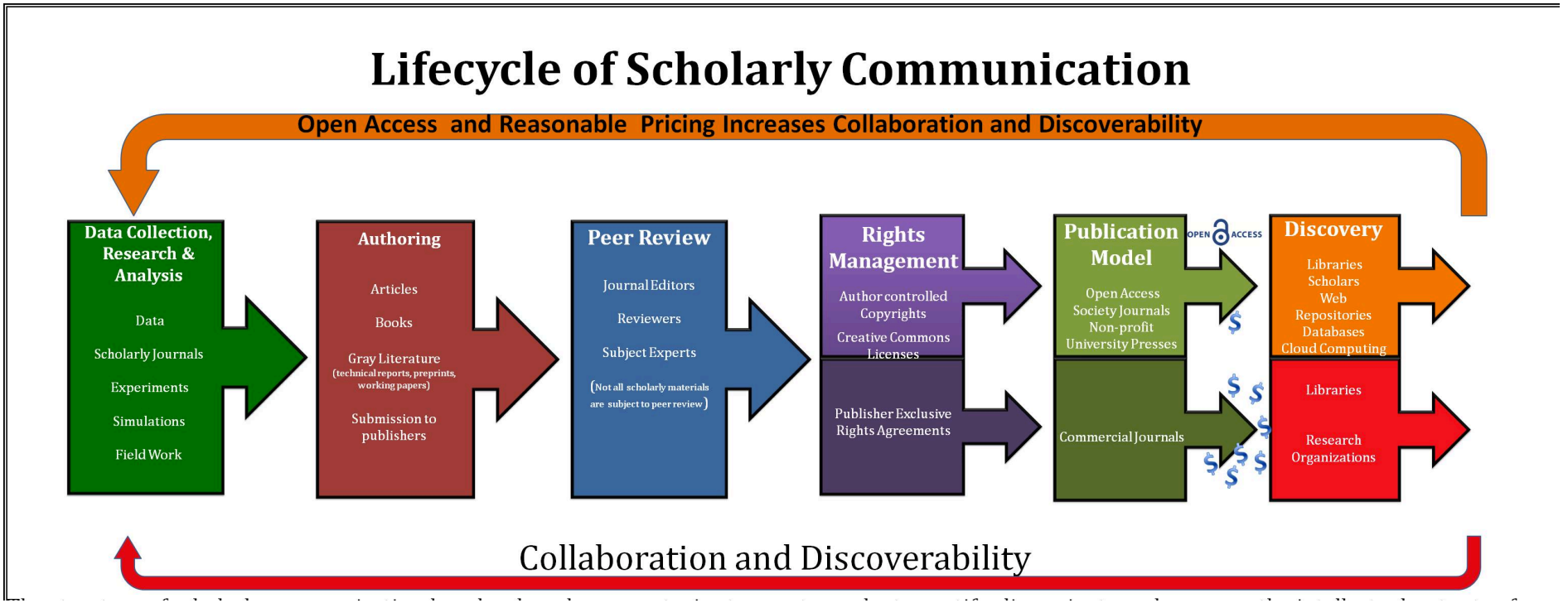


https://www.nature.com/news/the-top-100-papers-1.16224
https://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

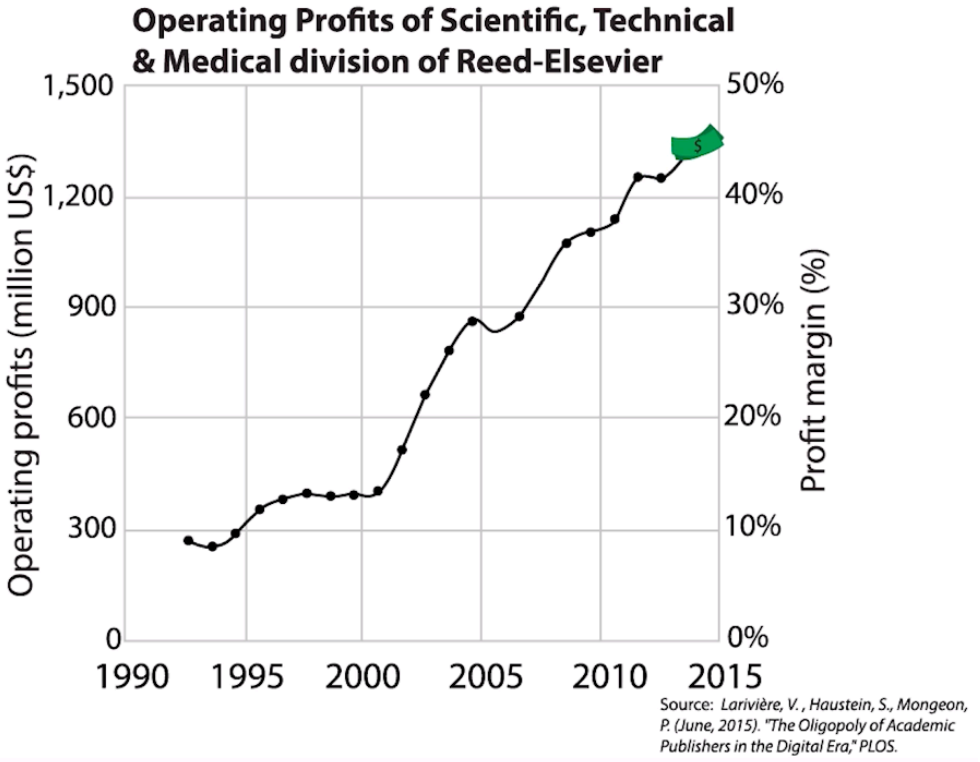


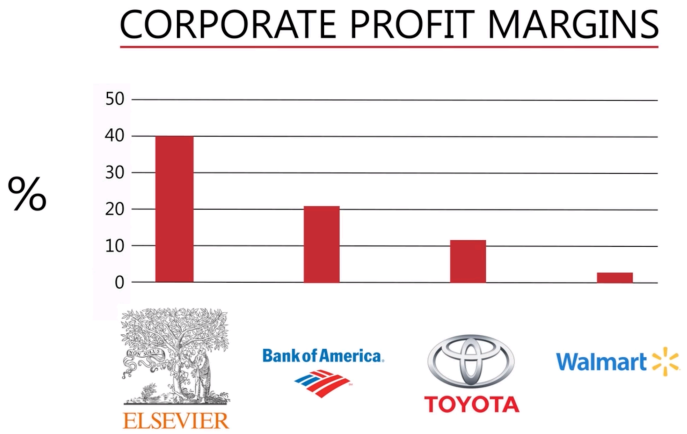
https://elifesciences.org/articles/27725

Challenge 3: Scientific Communication



Challenge 3: Scientific Communication



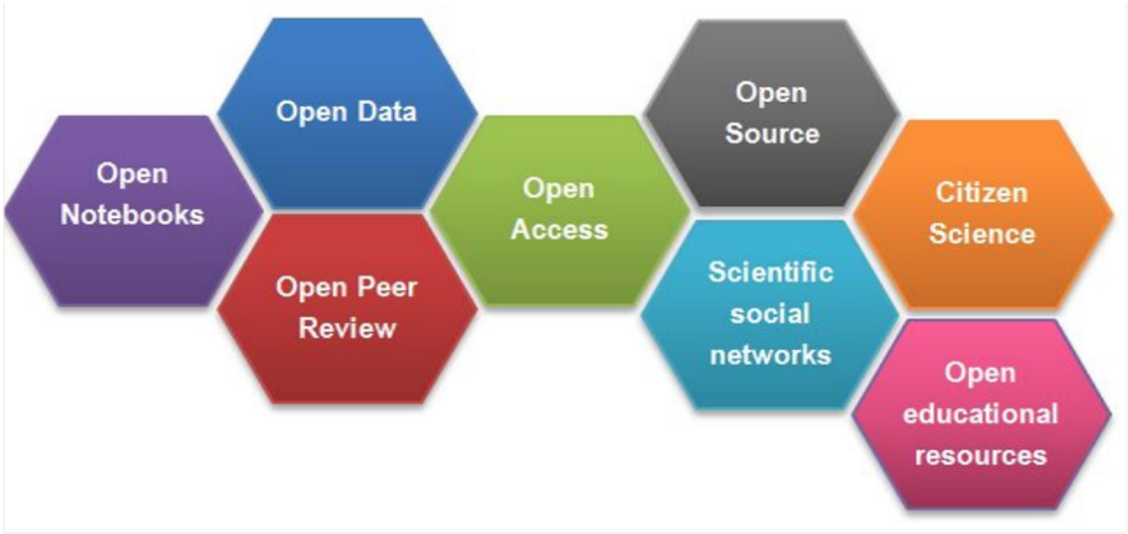


What does this mean for small universities with no significant funding?

OPEN SCIENCE INTRO

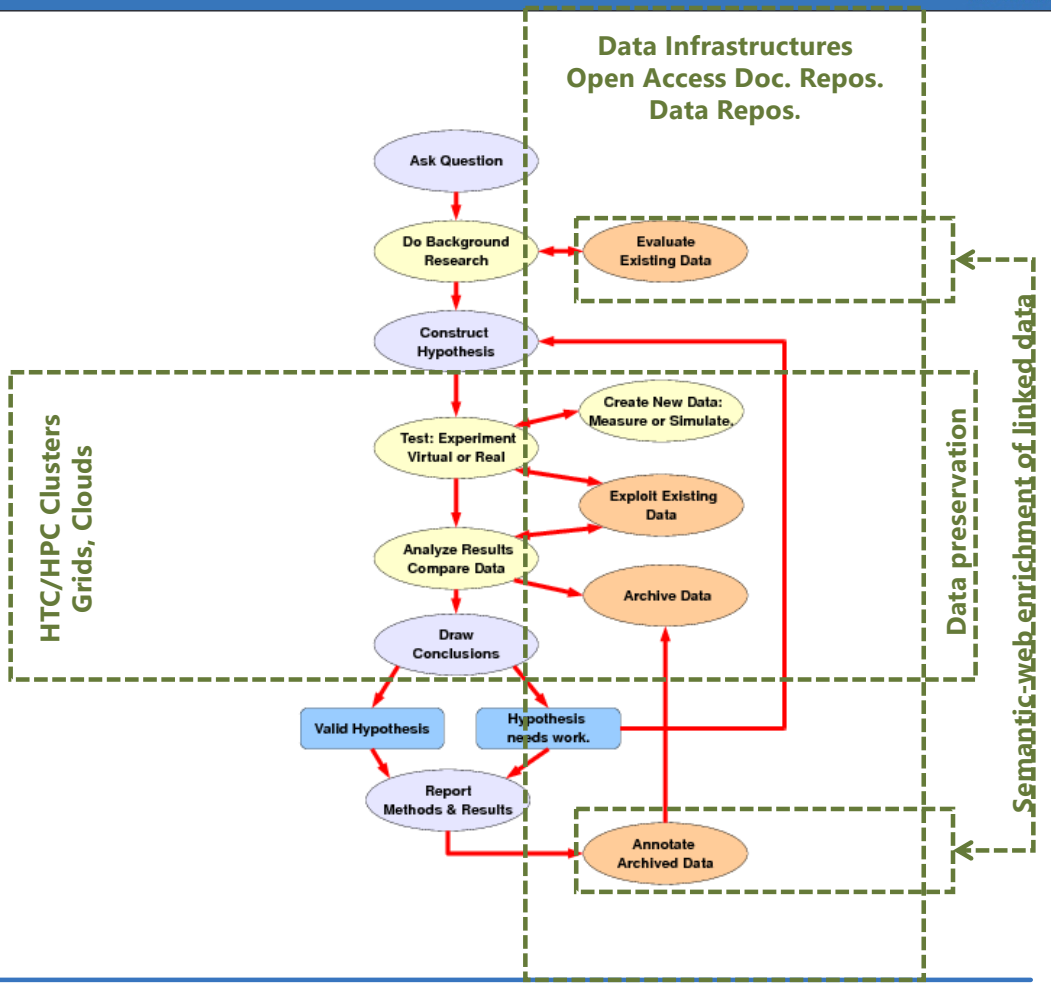
Open Science

Open Science is the movement to make scientific research (publications, software, data) and its dissemination accessible to all levels of society. Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks.



https://www.fosteropenscience.eu/node/1420

How to fit open science in research data lifecycle



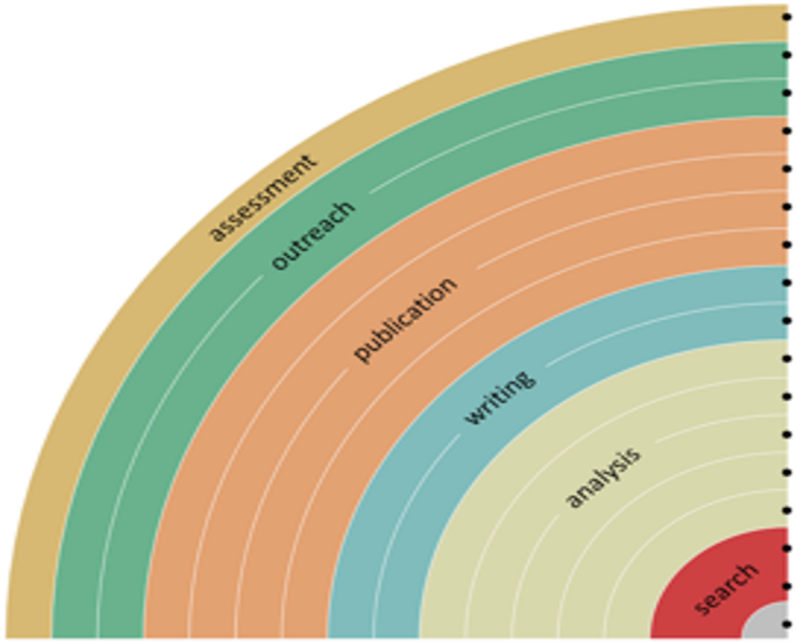
https://www.ukdataservice.ac.uk/manage-data/lifecycle

Open science approach example

https://slides.com/agahkarakuzu/qmrlabjn1

More possibilities?

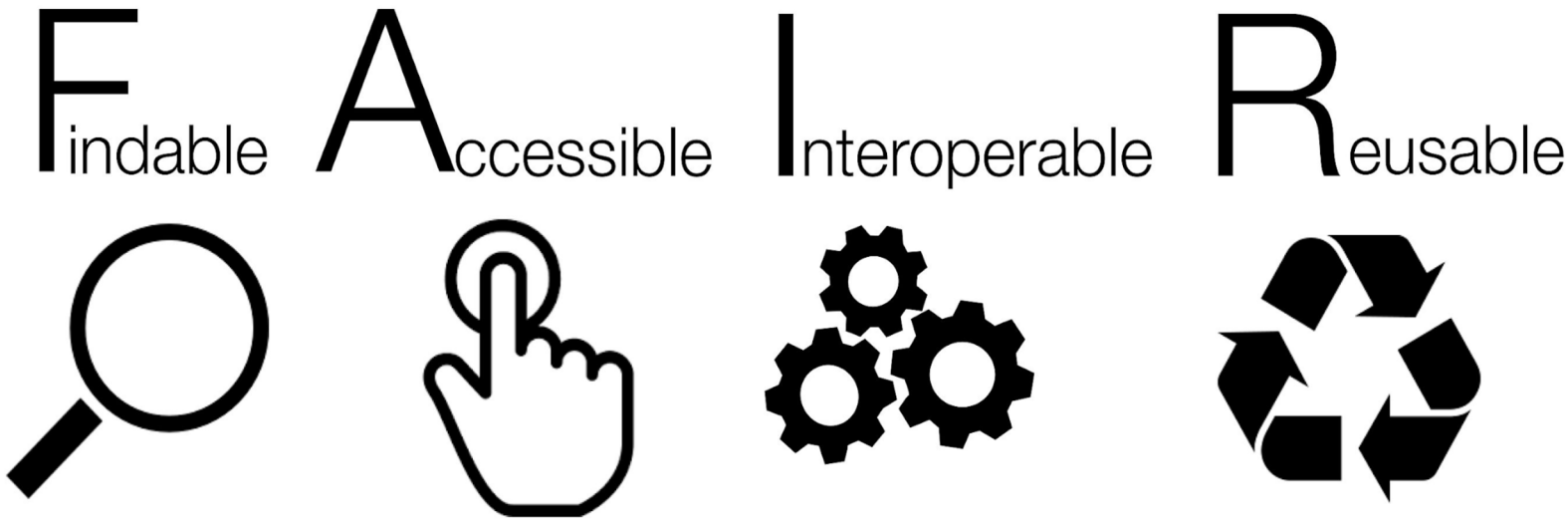
You can make your workflow more open by ...



- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO

DOI: 10.5281/zenodo.1147025

FAIR data principles



https://en.wikipedia.org/wiki/FAIR_data#/media/File:FAIR_data_principles.jpg

FAIR data principles

- **F**indable – Easy to find by both humans and computer systems and based on mandatory description of the metadata that allow the discovery of interesting datasets;
- **A**ccessible – Stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions (Open Access when possible), whether at the level of metadata, or at the level of the actual data content;
- **I**nteroperable – Ready to be combined with other datasets by humans as well as computer systems;
- **R**e-usable – Ready to be used for future research and to be processed further using computational methods

http://www.nature.com/articles/sdata201618
www.force11.org/group/fairgroup/fairprinciples

FAIR data principles: What need to be done?

To be Findable:

- assigned a globally unique and eternally persistent identifier
- described with rich metadata
- registered or indexed in a searchable resource
- metadata specify the data identifier

To be Accessible:

- retrievable by an identifier in a standardized communications protocol
- protocol is open, free, and universally implementable
- protocol allows for an authentication and authorization procedure
- metadata are accessible, even when the data are no longer available

FAIR data principles: What need to be done?

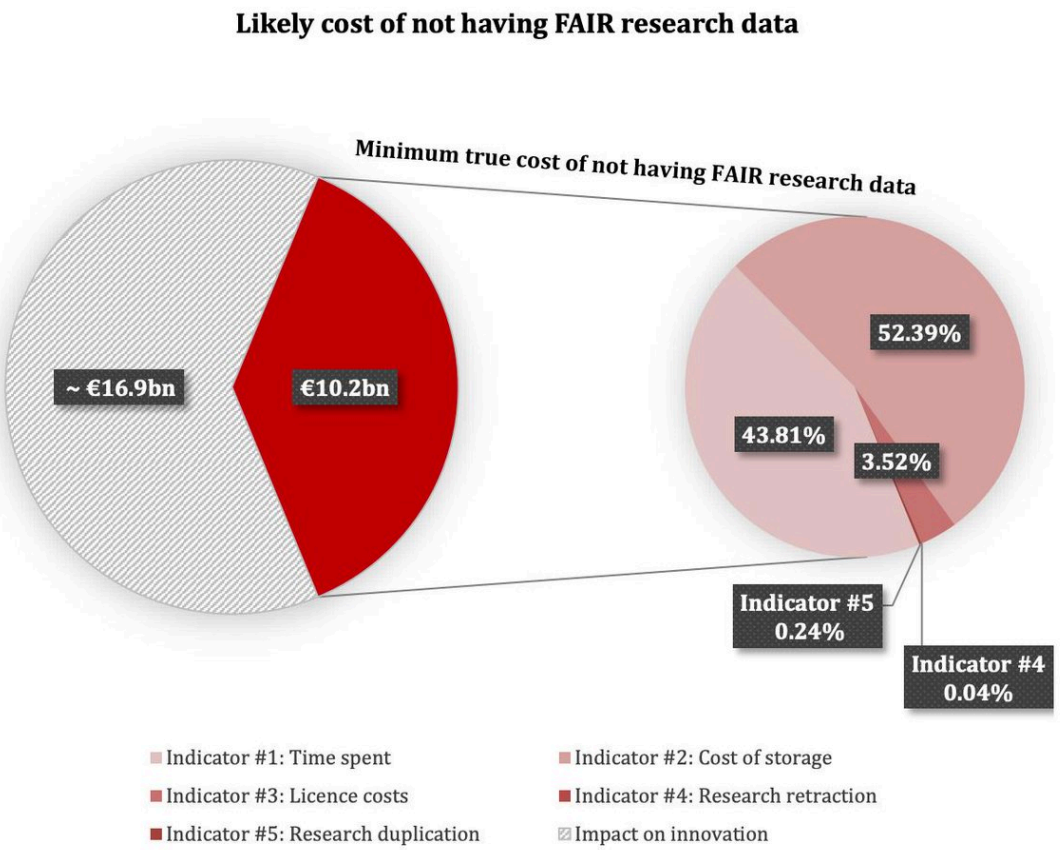
To be Interoperable:

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

To be Re-usable:

- plurality of accurate and relevant attributes
- released with a clear and accessible data usage license
- associated with their provenance
- meet domain-relevant community standards

FAIR data costs



https://publications.europa.eu/en/publication-detail/-/publication/d375368c-1a0a-11e9-8d04-01aa75ed71a1/language-en

Why it is needed?

- Virtual Research Communities
- e-Infrastructure (NI4OS)
- Applications
- Data
- Instruments/sensors

NEXT STEPS

Initiatives concerning FAIR

FAIRsFAIR (Fostering FAIR data practices) - www.fairsfair.eu

(aims to supply practical solutions for the use of the FAIR data principles throughout the research data life cycle)

GO FAIR initiative - www.go-fair.org

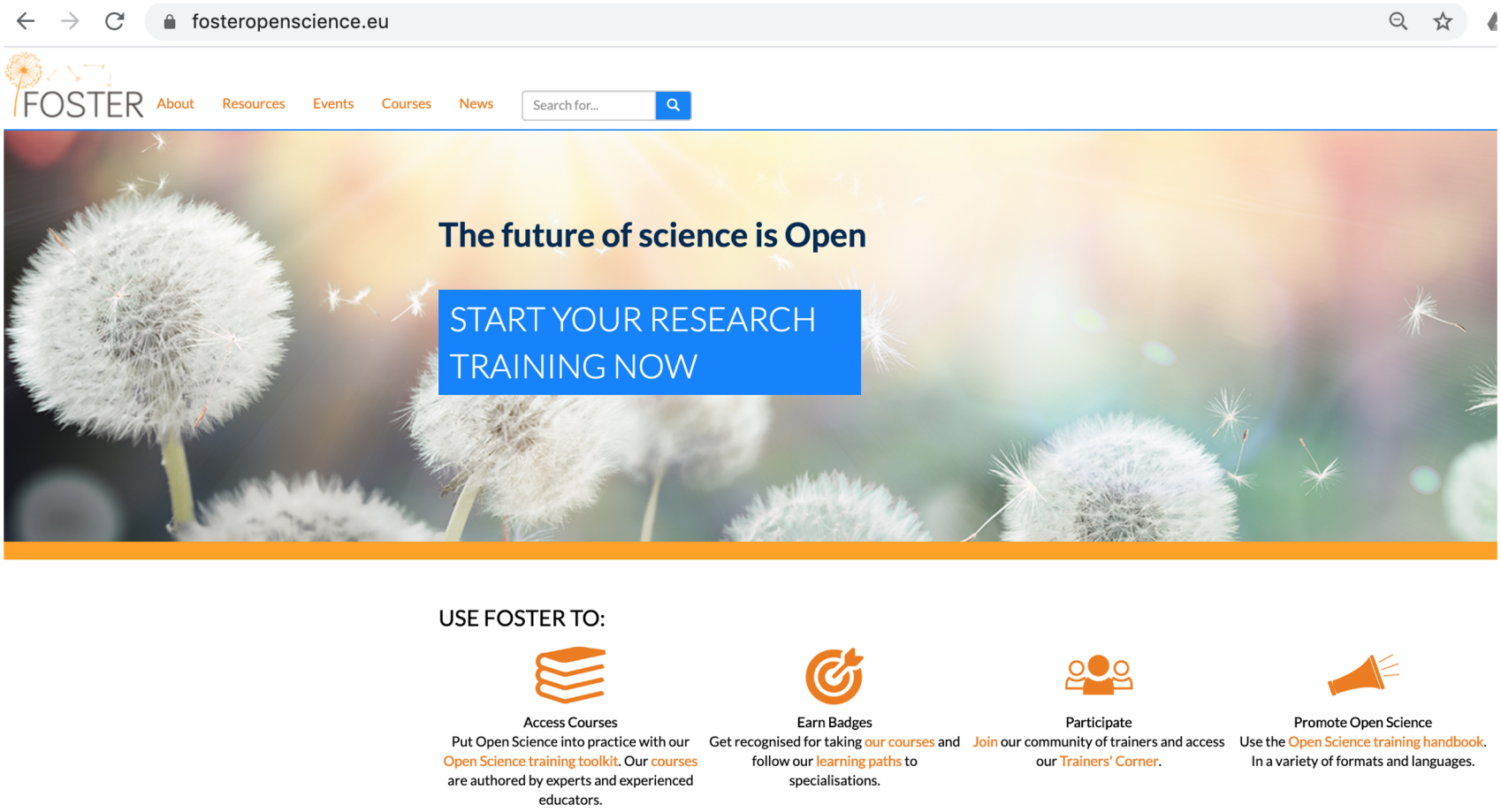
(aims to implement the FAIR data principles)

Detailed description of FAIR principles https://www.go-fair.org/fair-principles/
FAIR resources https://www.go-fair.org/resources/

FAIRsharing - fairsharing.org

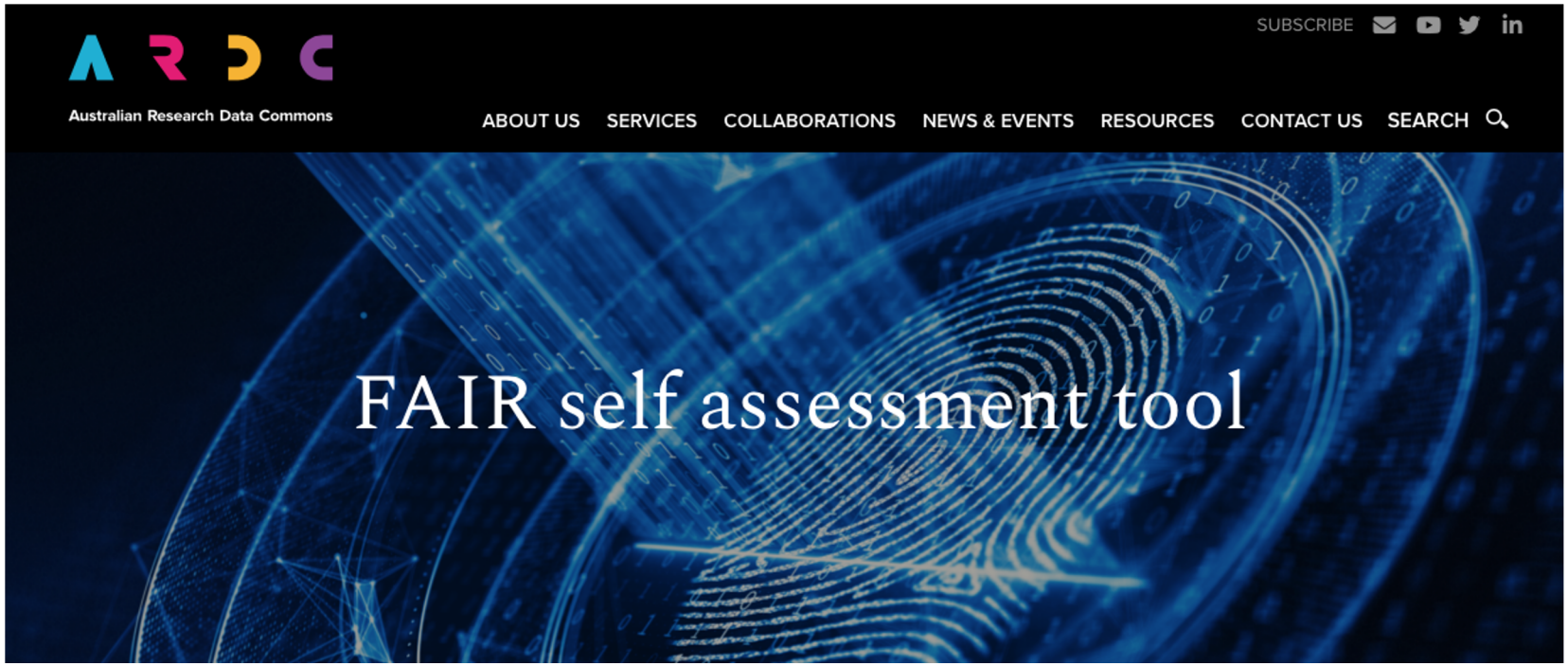
(A curated , informative and educational resource on data and metadata standards inter-related to databases and data policies)

FOSTER: e-learning platform on Open Science



FAIR services for existing data (After research)

FAIR data assessment services:



https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/

FAIR services in ongoing research

Training on data platforms and repositories

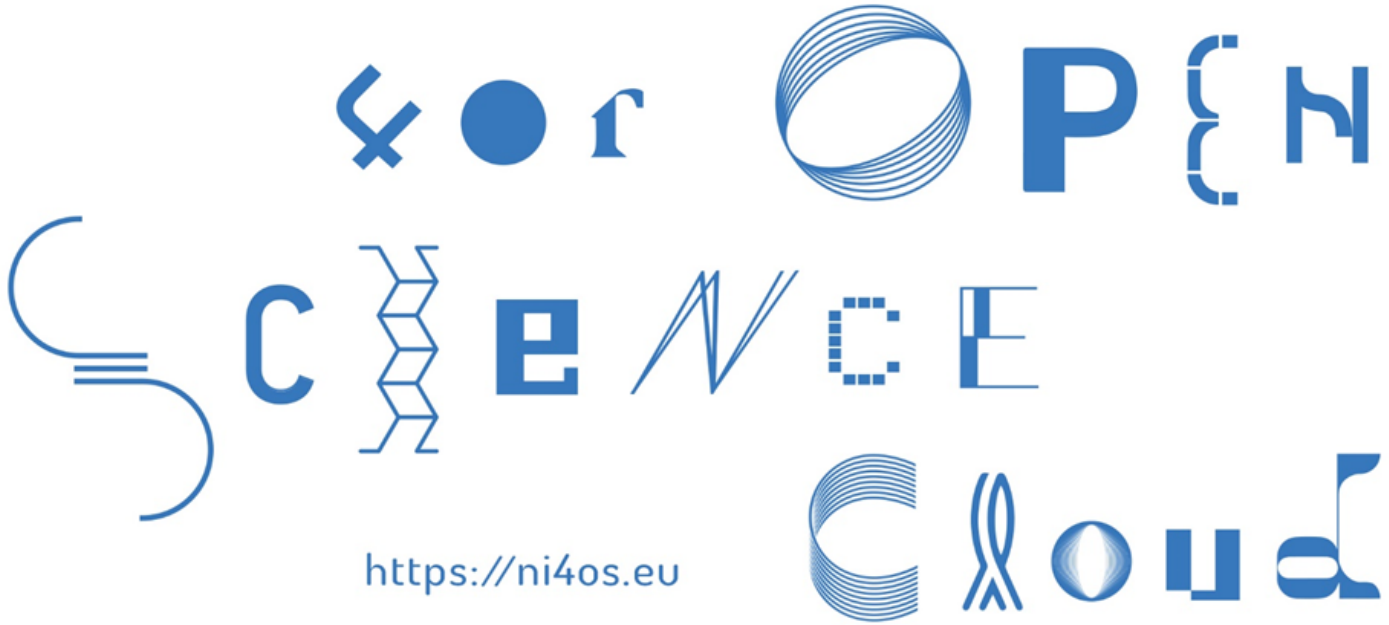
Choosing the right platform / repository very much determines the FAIRness of the stored data!

Exercises / training around re3data.org are often clarifying the issues that require attention (metadata, certification, etc.) but **not readily available**...

https://www.re3data.org

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Thanks!



Vladimir Trajkovikj (trvlado@finki.ukim.mk)

@NI4OS_eu @NI4OS