Placing research software into Open Science - Initial results from an RDA Sweden and EOSC Nordic collaboration

Arin Savran (SND & University of Gothenburg)
Monica Lassi (EOSC Nordic & Lund University
Maggie Hellström (RDA Sweden & Lund University)

RDA Sweden & SND webinar 2020-09-25

Housekeeping rules

- The webinar will be recorded and shared afterwards with the slides!
- Keep your microphone muted & camera switched off unless you are speaking
- Feel free to add questions to the chat, or raise your hand during the Q&A session
- We request that you display your full name in the participants list!

Welcome to the webinar!

Max Petzold

Director of SND

Welcome - Programme

- 13:00 - 13:10: Welcome and some words on RDA & RDA Sweden *(Max Petzold, Maggie Hellström)*
- 13:10 - 13:15: Introduction to EOSC Nordic (*Monica Lassi*)
- 13:15 – 13:35: Placing research software into Open Science – report from a collaboration between RDA Sweden and EOSC Nordic (Arin Savran, *Maggie Hellström, Monica Lassi*)
- 13:35 - 14:00: Q&A and discussion

RDA global and RDA Sweden

- a quick introduction

Maggie Hellström, RDA Sweden project coordinator
Lund University, ICOS Carbon Portal and SND

RDA Sweden research software webinar, 2020-09-25

Research Data Alliance (RDA)

RESEARCH DATA ALLIANCE

- RDA was launched as a **community-driven initiative** in 2013 by the European Commission, the US National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation
- RDA is building the **social and technical bridges** that **enable open sharing of data** to achieve its vision of researchers and innovators openly sharing data across technologies, disciplines, and countries **to address the grand challenges of society**.

- Read more at https://rd-alliance.org/

What does RDA do?

RESEARCH DATA ALLIANCE

- Activities centre around self-formed, volunteer and focused Working Groups and exploratory Interest Groups
- Aims are to
  - exchange knowledge
  - share discoveries,
  - discuss barriers and potential solutions,
  - explore and define policies and
  - test and harmonise standards to enhance and facilitate global data sharing & re-use.

37 working groups
& 57 interest groups

https://www.rd-alliance.org/groups

RDA in numbers
(August 2020)

RESEARCH DATA ALLIANCE

- **10997 individual members** from **145 countries** (69% Academia & Research, 14 % Public Administration, 11 % Enterprise & Industry)
- **52 organisational members** & **11 affiliate members**
- **94 groups** working on **global data interoperability challenges** (37 working groups & 57 interest groups)
- **50 flagship outputs** of which 8 ICT Technical Specifications*)
- **100+ adoption cases** across multiple disciplines, organisations & countries

*) information and communication technology (ICT) technical specifications defined by the European Commission as eligible for referencing in public procurement processes.

https://www.rd-alliance.org/sites/default/files/attachment/RDA-in-a-nutshell-August-2020.pptx

RDA recommendations & outputs

- **Recommendations:** RDA's flagship outputs and the equivalent of "specifications" or "standards" that other organisations create and endorse.
- **Supporting Outputs:** RDA outputs that originate from the work of RDA groups, but not always adoptable by other organisations.
- **Other Outputs:** Workshop reports, published articles, survey results, etc. produced by RDA groups. (Not formally endorsed.)

https://www.rd-alliance.org/recommendations-outputs

Joining RDA

- Any individual can join the RDA **for free** https://www.rd-alliance.org/user/register
- **Only requirement** is to **agree** to the RDA's **guiding principles:** Openness, Consensus, Balance, Harmonization, Community-driven, Non-profit and Technology-neutral
- Once a member, you can
  - Sign up for as many groups as you like
  - Comment on all outputs & RDA web content
  - Participate in RDA plenaries & other meetings
  - Vote for candidates to the RDA Council & Technical Advisory Board

RDA plenary meetings

Plenary 16
@cyberspace: 9-12 November, 2020 (registration will open soon!)

Plenary 17:
@Edinburgh, UK: 20-22 April, 2021 + virtual

Plenary 18:
@Seoul, South Korea: 8–11 November 2021 (part Of International Data Week 2021)

Regional & national groups

- RDA encourages the initiative and support of its **members** across the globe to animate their **communities on a national level**
- National and regional nodes provide platforms for **exchanging information** (in the local language) about the RDA, its activities and outputs
- In Europe, the H2020 project “RDA Europe” is supporting the formation of **national nodes** (currently **22**)

https://www.rd-alliance.org/groups/national-groups

RDA-SE organization & budget

- National node of RDA global, via RDA Europe
- Swedish National Data service (SND) is host institution
  - Core team at SND Gothenburg + Lund
  - Working groups, advisory group
- Application in April 2019 successful!
- “Seed money” (-> ~~May 2020~~) (now prolonged -> Sept 2020!)

RDA-SE aims

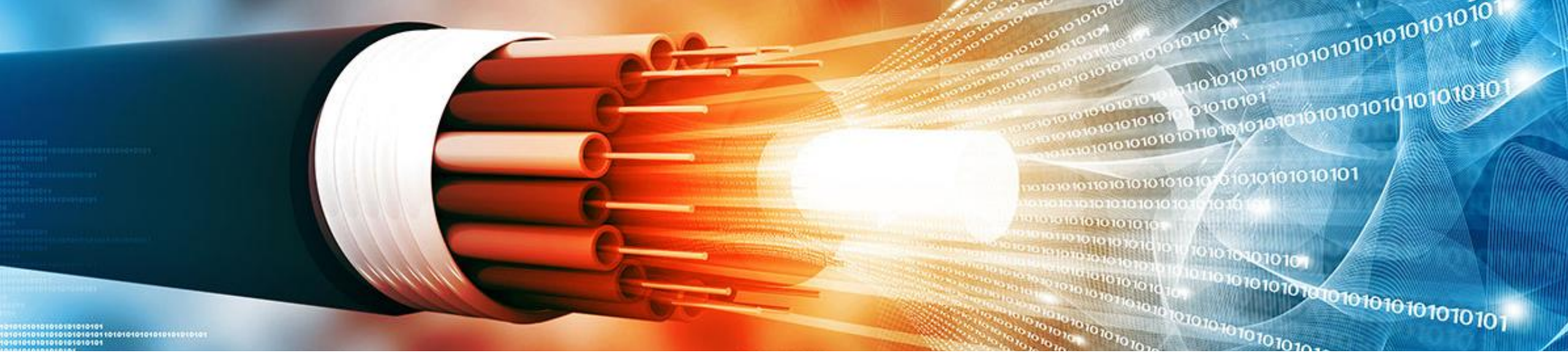
- Spread info about RDA in Sweden
- Increase Sweden-based individuals' membership in RDA
- RDA working & interest groups
  - identify interests, increase active involvement
  - inform, encourage and facilitate adoption of Outputs & Recommendations
- Be a platform for dialogue on research data issues
  - engage with broad range of stakeholders & funders
  - map out the "research data landscape" of Sweden
- Engage & network with other national nodes

RDA-SE activities

- Set up a web site https://rd-alliance.org/groups/rda-sweden
- Information events at SND consortium & network institutes
- Have info points at meetings & conferences (e.g. IFFIS)
- Join network of Nordic RDA nodes
- Promote adoption of RDA recommendations & outputs (**webinars**)
- Project with EOSC Nordic on research software
- Organize a F2F workshop (spring 2020) <- perhaps in 2021?
- Meetings with funding agencies, “academies” & professional organisations <- during autumn 2020

RDA-SE future?

- The support from RDA Europé ends on September 30, 2020.
- SND will continue to host and sponsor RDA Sweden activities
- Everyone is welcome to join us!
  1. Become a RDA member (https://www.rd-alliance.org/user/register)
  2. Sign up for RDA Sweden (https://rd-alliance.org/groups/rda-sweden)
- Comments or ideas for RDA Sweden activities? Get in touch!
  - Maggie Hellström (margareta.hellstrom@nateko.lu.se)
  - Max Petzold (max.petzold@snd.gu.se)



EOSC NORDIC

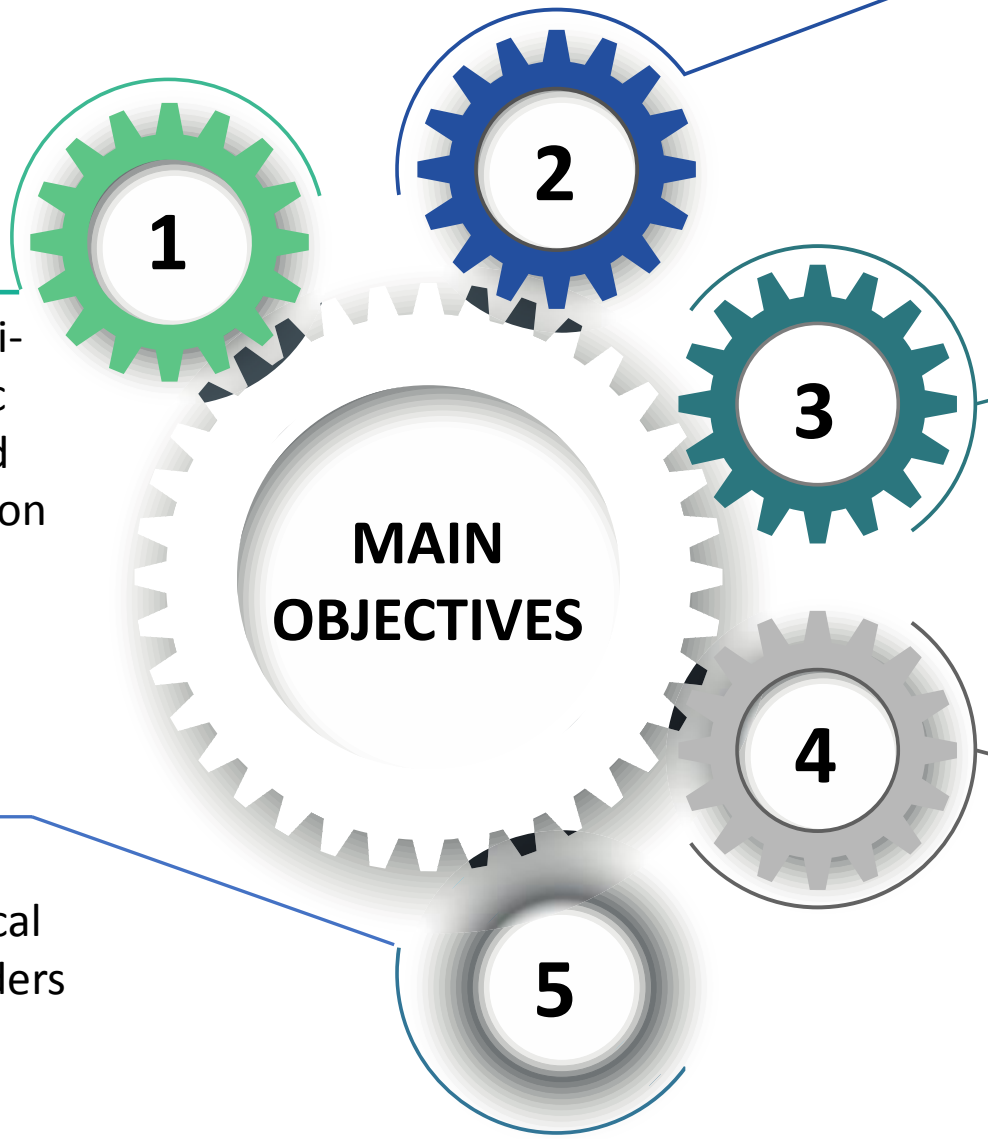
Developing and promoting incentives for FAIR data

Monica Lassi, LUNARC, Lund University, (SNIC/LU), Sweden (EOSC Nordic WP4, task lead for T4.3)

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

neic
NORDIC E-INFRASTRUCTURE COLLABORATION

EOSC NORDIC



OBJECTIVE 1

Support coordination, harmonisation and alignment of Nordic and Baltic national policies and practices related to the provision of horizontal research data services with EOSC

OBJECTIVE 2

Increase the discoverability of Nordic & Baltic services. Extend and expand their use by making them accessible through the EOSC portal

OBJECTIVE 3

Promote and support the uptake of FAIR data practices and certification schemas across the Nordics

OBJECTIVE 4

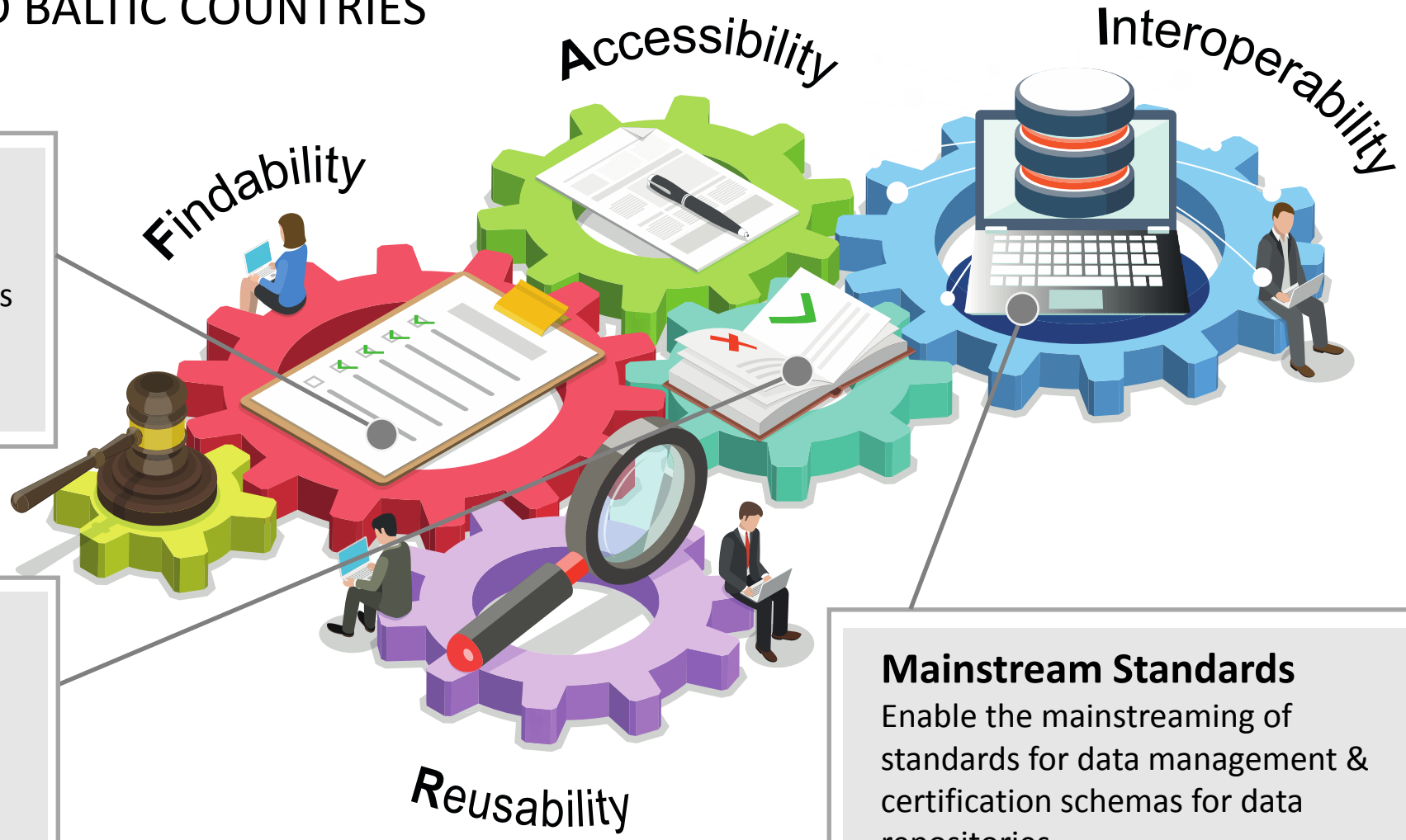
Accelerate the progress and attractiveness of EOSC by piloting & delivering innovative solutions developed and tested in a useful and functional cross-border environment

OBJECTIVE 5

Provide a Knowledge Hub to deliver training and technical support to new service providers and communities willing to engage with EOSC during and after the project lifetime

DEVELOPING FAIR DATA PRACTICES

ACROSS THE NORDIC AND BALTIC COUNTRIES



Investigate Practices

Investigate and inform about the state of FAIR practices in the Nordics and the Baltics, looking at national policies and practices.

Develop Incentives

Develop and promote incentives for the uptake of FAIR data practices across national scientific communities

Mainstream Standards

Enable the mainstreaming of standards for data management & certification schemas for data repositories

Incentives for researchers to make data FAIR

- T4.3.1 Mapping of policy implemented incentives within the Nordic and Baltic members states

- T4.3.2 Identifying effective incentives that increase uptake of FAIR practices

- T4.3.3 Dialog with stakeholders and policy makers

More on EOSC Nordic, including our newly launched Knowledge Hub: https://www.eosc-nordic.eu

Placing research software into Open Science - Initial results from an RDA Sweden and EOSC Nordic collaboration

Arin Savran (SND & University of Gothenburg)

Monica Lassi (EOSC Nordic & Lund University

Maggie Hellström (RDA Sweden & Lund University)

Open Science...

"...is the practice of research, education and knowledge exchange in such a way that **others can collaborate and contribute**,

where research publications, data, lab notes and other scholarly processes and works are **properly and ethically managed and evaluated** and,

unless restricted for justifiable reasons, are **freely available to all levels of society** under terms that **enable reuse, redistribution and reproduction** of the work and its underlying data and methods."

FOSTER project (fosteropenscience.eu)

Image: Digitalbevarink.dk under a CC BY 2.5 Denmark license

Open Science =

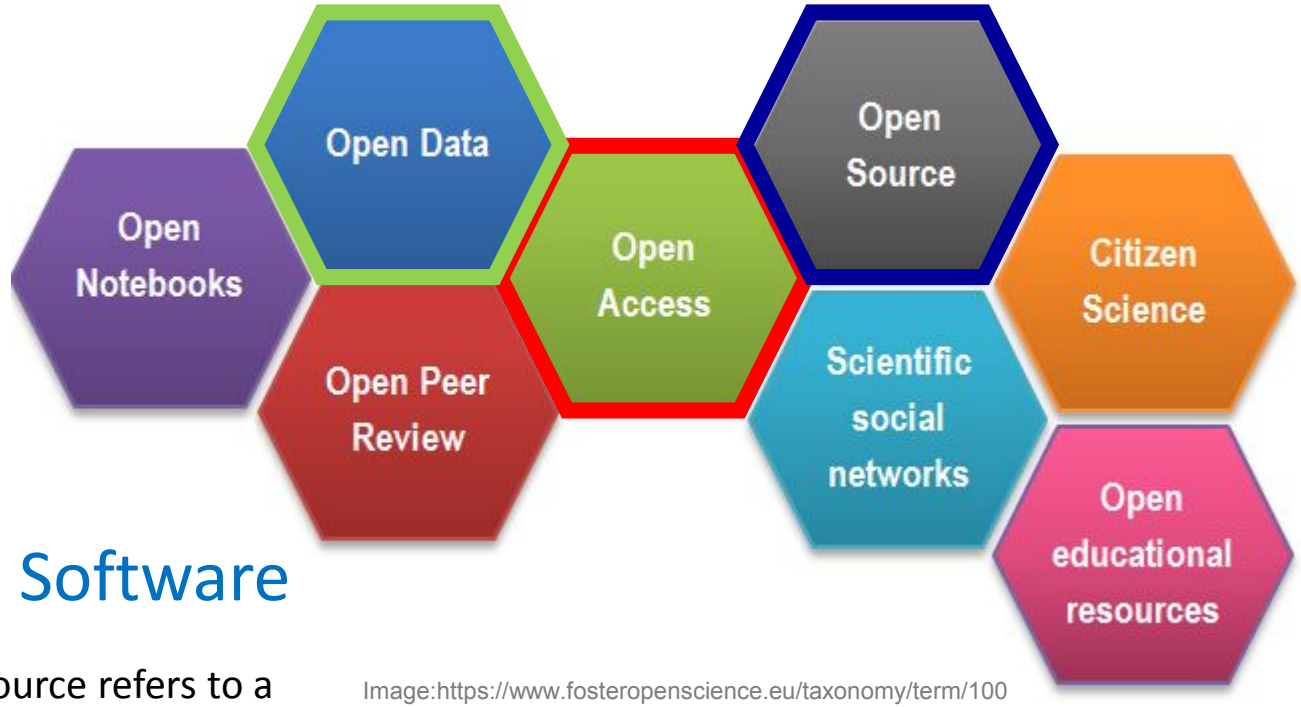


Image:https://www.fosteropenscience.eu/taxonomy/term/100

Open Access + Open Data + Open Software

Open Access refers to online, free of cost access to peer reviewed scientific content with limited copyright and licensing restrictions.

Open Data are online, free of cost, accessible data that can be used, reused and distributed provided that the data source is attributed.

Open Source refers to a program or software in which the source code is available to everyone to use and/or modify from its original design free of charge.

FAIR software?!

- FAIR stands for Findable, Accessible, Interoperable, Reusable
- Originally formulated for data & metadata
- Can (research) software also be FAIR?
  - Ongoing work suggests yes, after some reformulation
- FAIR is not equal to Open or Free
- Says nothing about quality or fitness-for-purpose

Image: Microsoft “screen bean”, free for non-commercial use

Software in the research data lifecycle

A typical research project has a number of steps from idea to publication

(Not always so nicely ordered, though)

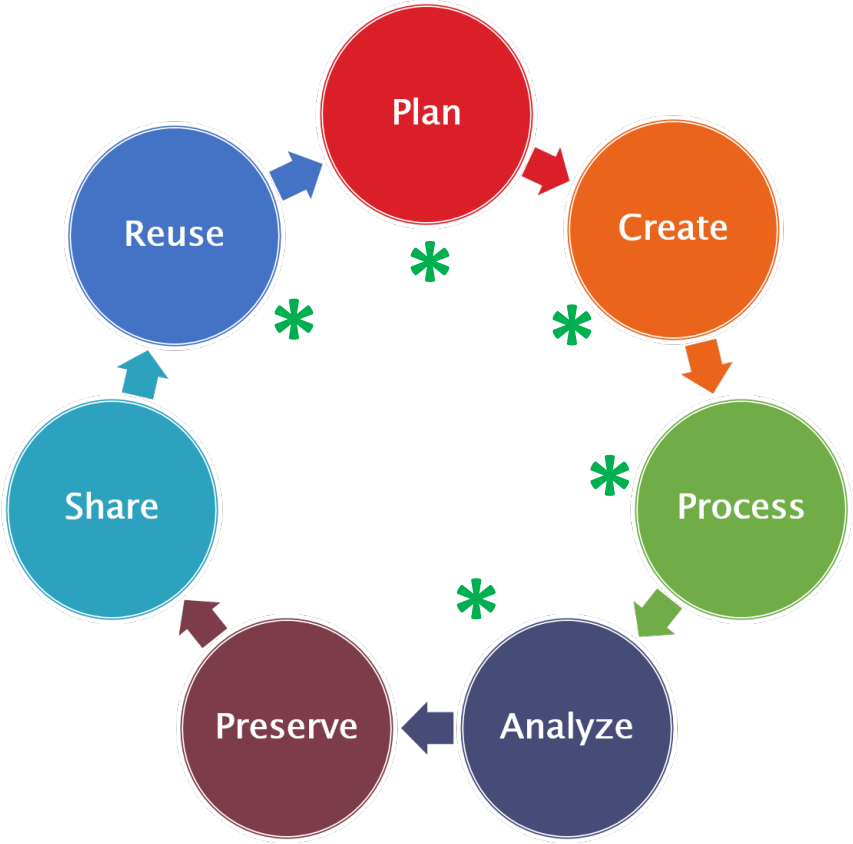


Image credit: University of Ottawa library, https://biblio.uottawa.ca/

"Research software" of different types

- Spreadsheets
- Scripts (R, Matlab,…)
- Professorware
- Numerical models
- …

RDA activities on software

A number of Research Data Alliance groups are, or have been, active on related questions, including:

- RDA/FORCE11 Software Source Code Identification WG
  - Recent output: A. Allen et al.(2020). Use cases and identifier schemes for persistent software source code identification (V1.0). Research Data Alliance. https://doi.org/10.15497/RDA00053
- CURE-FAIR Working Group
- FAIR for research Software (FAIR4RS) WG

About our project

- Understand how software fits into the Open Science umbrella
- Review the literature on Open Software/Open Source Software/research software
- Develop discussion questions
- Gather views from the Nordic and Baltic RDA nodes & other interested parties

Terminology zoo

"Software for research"

Open Science

"Research software"

Open Software

Open Data

Open Access

"Open source software (OSS)"

Code vs. Software

Licenses

"Citable software"

"FAIR software"

Software repositories

Versioning

Software catalogues

Literature review A

- Confusion & no consensus on major terms/concepts
- Somewhat romanticized views of repositories
- Inter-citation but lack of deeper engagement with ideas proposed
- Various gaps exist
  - Researcher incentives?
  - Confidentiality/anonymity
  - Creation/Funding of national observatories or repositories
  - Taxonomies or systems for citation lacking
  - Achievements then?

Literature review B

- Lack of universal or regional policies and laws
- Computing & software engineers largely dominant
- Intra-European and global inequalities re Open Science overall
- European and North Americans shaping the agenda

Discussion questions for survey (selected)

- How would you estimate the level of support among researchers for Open Science and FAIR in your country?
- Are the terms “Open Software”, “research software”, “software for research” used in your country?
- Is (open) software for research part of the national data management discussions in your country?
- Are there any repositories or observatories for research software in your country? Who operates them, and what stage of development are they in?
- In your country, what are the main practical, legal, and/or bureaucratic challenges faced with regards to archiving, storing, documenting, safeguarding access, etc. to research software?

Discussion meeting with Nordic RDA nodes - A

- Software simply not talked about like data. FAIR goals taken with a pinch of salt
  - Researcher problem or policy problem?
- Terminology differs (e.g. algorithms, code, scripts, proprietary and non-proprietary software…)
- Increasing discussions about what researchers need for storage & licensing

Discussion meeting with Nordic RDA nodes - B

- National policies in progress with various targets set, but repositories not always advertised well
- GitHub dominates but repositories exist (with country-specific differences)
- Challenges include funding, finding right expertise, researcher incentives

Conclusions

- Open Software still the odd one out in Open Science.
- Literature still needs development & cross-disciplinary debates.
- Lit review largely consistent with RDA Nordic discussion – meeting addressed some gaps.
- General lack of knowledge on a systemic level on how to curate, store, & access software/code in a FAIR & open way.
- EOSC Nordic survey must ask deeper questions pertaining to issues discussed, incl. incentives, research process, publishing practice, domain specific questions, FAIR.

Thanks for your attention!

- The project report will be released on September 30
- If you have comments or questions, ask now or get in touch afterwards
  - Arin Savran (arin.savran@gu.se)
  - Maggie Hellström (margareta.hellstrom@nateko.lu.se)
  - Monica Lassi (monica.lassi@lunarc.lu.se)

Q & A

- **Questions from you!**

And then, if we have time…

- Is Open Software as important for Open Science as Open Access and Open Data?
- Do we need special policies for research software?
- Who should take the lead: funders, data centers, research institutions – or SND?
- Should we start a working group or set up a discussion forum?