Research Libraries, Researchers & the EOSC: Technical Universities

REPORT FROM WORKSHOP
WRITTEN BY GERGELY SIPOS

Introduction

The workshop was held on 27 January 2021, between 14:30 - 16:15 CET on Zoom. The event was the fifth and last in a series of workshops organized by Scientific Knowledge Services and LIBER with the aim of answering the following questions:

1. What is the value of EOSC for researchers and research libraries, based on the goals/work of the EOSC?
2. What is the input needed from these stakeholders towards the EOSC?
3. How can these stakeholder groups be actively involved in EOSC activities and what do they need to get involved?
4. What feedback mechanism could be built to continuously inform EOSC, in its quest to remain an agile infrastructure?

This 5th workshop focused on Technical Universities. The target audience were technical universities.

The workshop programme and information about the speakers, moderator and rapporteur can be found here: https://www.knowledge.services/research/research-libraries-researchers-and-the-eosc/w5/

Participant analysis

Excluding the speakers and organiser staff **28 people registered to the event from 22 institutes from 12 countries** with the following distribution: Austria (2), Estonia (1), Greece (1), Ireland (1), Italy (1), Lithuania (1), Netherlands (5), Poland (4), Romania (4), Spain (5), Sweden (1), UK (2).
The **actual participation was 19 people from 16 institutes from 10 countries (68% turnup rate)** with the following distribution: Austria (2), Estonia (1), Italy (1), Lithuania (1), Netherlands (3), Poland (3), Romania (3), Spain (3), Sweden (1), UK (1).

The gender distribution[1] of the 34 attendees was: **12 female (63%), 6 male (32%) and 1 unknown (5% - did not provide a real person name on Zoom)**.

It is probably worth noting that the workshop date overlapped with a major European event hosted by Science Europe, *Launch Event of Updated Practical Guide on Aligning Research Data Management Across Europe*.

Workshop structure and atmosphere

The workshop included 3x10' long presentations, then a 40' long break-out session in 3 parallel Zoom rooms, answering the same 4 questions, followed with a short reporting back session in the plenary room from the 3 breakout groups.

The rapporter found the event professionally organised with no technical or organisational issues during the workshop.

While 19 persons attended the presentations, only 11 remained for the breakout groups, resulting in only 2 persons in one of the breakout rooms. 2 break-out groups with 5/6 people would have been a better arrangement.

An important observation on this event is that the breakout questions were designed about 'Researchers' and 'Libraries', while the workshop focus was on Technical universities. During the reporting back session one asked the question of 'what is the relevance of EOSC to technical

[1] Assessment was made by the rapporteur based on the participants' first names. Participants were not asked to state their gender during registration/participation.

universities then?'. This question triggered a separate discussion which directly responded to the title of this workshop and clarified the specificities of research universities.

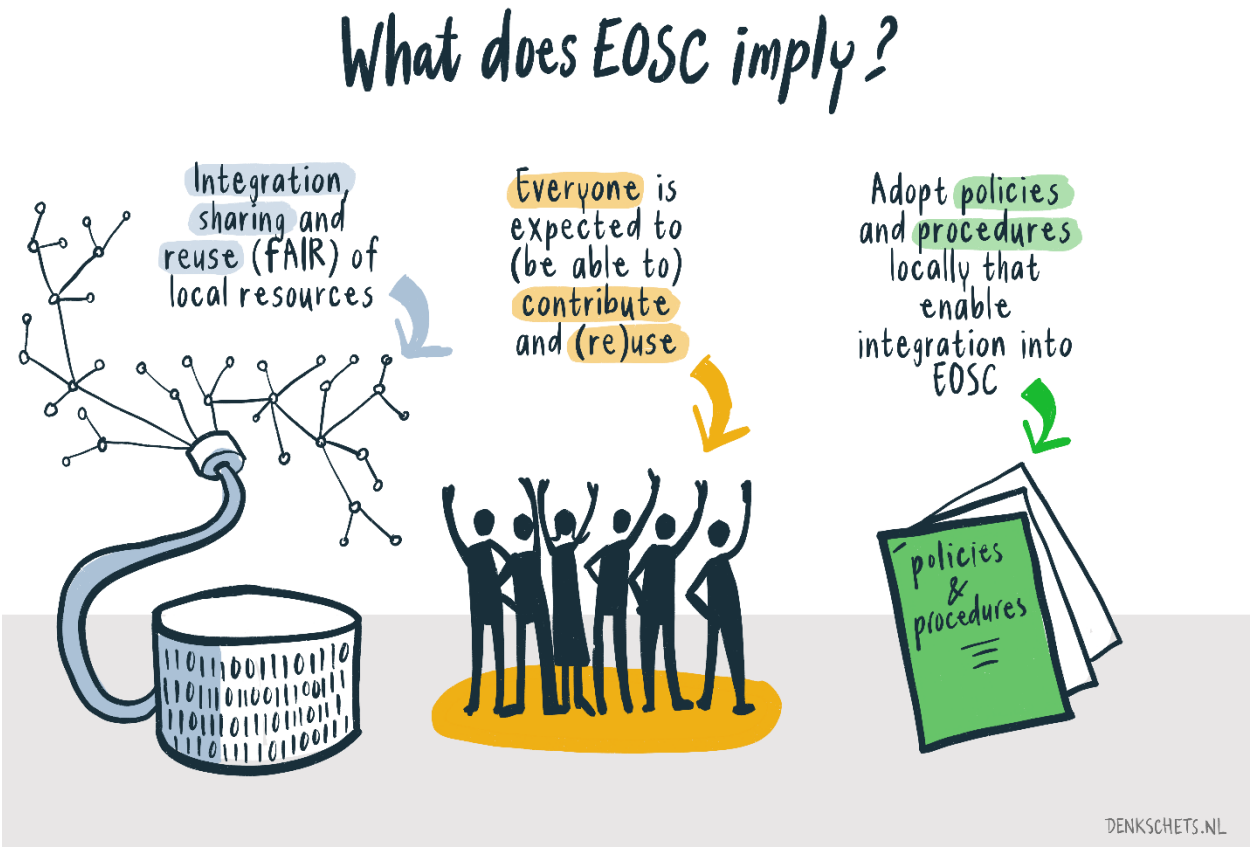

Figure 1: What does EOSC imply? (Denkschets, 2021)

The main points from this discussion:

- Technical universities often focus on technology transfer into industry, instead of fundamental research (this is the high TRL according to the EC definition).
- The diversity of data types used at technical universities is broad - probably broader than in other research institutes. Moreover, there is huge fragmentation of approaches here - faculties do not adopt the same tools and standards. (This was also one of the key points in Federica Cappelluti's presentation.)
- There is a cultural difference between technical and humanities/law faculties: Researchers at technical and applied science faculties think that 'we know everything about research software and data', and there are a lot of solutions already in use at those faculties for RDM. EOSC can benefit these different faculties in different ways and needs different approaches. Applied science faculties do not want EOSC 'to lecture them' about what to use. Law/humanities may be more open to lectures, training, and new tools.

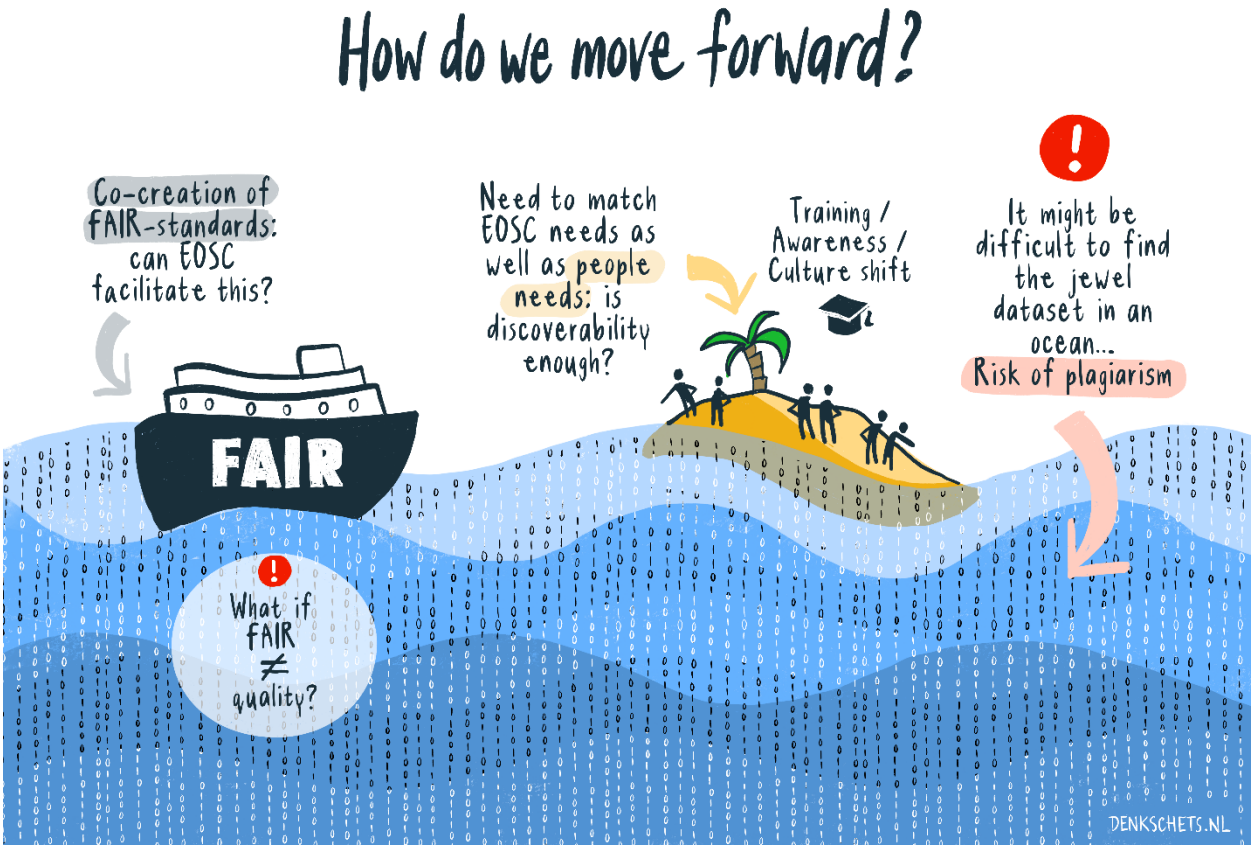

Figure 2: How do we move forward? (Denkschets, 2021)

One change to consider for future events is the inclusion of Q&A time between the plenary and the breakout sessions to allow the audience to comment on or ask about certain points that were presented. Such a Q&A slot could also energise the audience a bit more for the breakout.

Breakout group approach analysis

The style of the three breakout rooms were quite different from each other:

- The facilitator of the first room (with 5) participants chose the style of discussing sticky notes that were added to the Jamboard, resulting in more discussions and interactions, but a lower number of post-it compared to other groups.
- The facilitator of the second room (with 2 participants) was basically a dialogue between the two persons. They added only a few sticky notes and only to the first page of Jamboard.

- The facilitator of the third room (with 4 participants) used the time for letting participants add stickers to Jamboard instead of discussion. There was nearly zero discussion of the content added on stickers - resulting in a lot of stickers but no deeper analysis.

Summary of breakout outcomes

This section provides **the main, recurring answers points for the 4 questions from the 3 break-out groups**.

Q1: What is the value of EOSC for researchers and research libraries, based on the goals/work of the EOSC?

- Accessing standards and best practices, sharing knowledge with others who have common goals.
- Have a framework to learn about good practices and practical implementation of open science and to contribute as a trainer.
- A key selling point of EOSC can be research reproducibility. Demonstrations of such use case examples would be beneficial.

Q2: What is the input needed from these stakeholders towards the EOSC?

- The structure of EOSC is still unclear, clearer requirements and structure is needed before stakeholders could contribute.
- Data sharing often requires special knowledge of legal constraints, possibilities. This is something that not even librarians have and today they contract experts for such cases.
- A system like EOSC needs metadata and semantic standards that makes sharing cross-disciplinary. It was noted that reaching such standards is very difficult.
- Provisioning of infrastructure - Libraries either must become data centres, or contract with data centres if they wish to offer science data archival services.

Q3: How can these stakeholder groups be actively involved in EOSC activities and what do they need to get involved?

- There are significant differences in how countries want to engage with EOSC. (E.g. The Dutch approach presented by Laurents Sesink uses a hierarchical model with only certain Competence Centres to engage with EOSC. Other countries are assuming a flat model within the country allowing any institute to engage with EOSC.)
- New methods for recognition of 'data sharing', 'FAIR data production' are needed to change the researchers' mindset.
- The main EOSC actors (WGs, national members, projects) should actively communicate their achievements to libraries and other stakeholders.
- Start pilot projects that engage with as many disciplines as possible.

Q4: What feedback mechanism could be built to continuously inform EOSC, in its quest to remain an agile infrastructure?

- Working groups, feedback groups.
- A helpdesk, with a ticket dispatching system that forward tickets to the most appropriate place in EOSC.
- An online forum and FAQ for users.
- Annual surveys to users and providers.

Appendix - Event notes

Birit Schmit introduced the event, its goals and its structure, then invited the 3 speakers:

- Raphael Ritz, Head of Data Division at the Max-Planck Computing and Data Facility (MPCDF) – the central IT service unit of the Max Planck Society (MPG) in Germany
- Dr Federica Cappelluti, Associate Professor of Electronic Engineering at Politecnico di Torino
- Laurents Sesink, Head Centre for Digital Scholarship, Leiden University

14:35 - 14:45 Presentation: Dr Raphael Ritz: **<u>What to expect from EOSC As Researcher, Research Library, or Computing Centre</u>**

- He stated that the talk expresses personal opinion based on the involvement in various WGs (Arch) and projects
- Participant in the architecture WG. The WG delivered
    - A number of documents (slide 2)
    - Another important outcome was contribution to SRIA (slide 3). To what extend this feedback will be used by the commissions is still to be seen
    - Proposal for 'Minimal Viable EOSC' (MVE) (slide 4-5)
- What does this all mean?
    - Noone expects that people follow closely all these deliverables. But everyone is expected to be able to contribute and (re)use
    - The main point/concept about EOSC which is shared with the commission: "Integration layer over existing resources" - at institutions, universities, compute centres. They should be able to connect/contribute to EOSC to support the FAIR principles.
- Researchers: (slide 7)
    - Consider sharing publications
    - Take metadata seriously! (you will benefit from it, others will benefit from it)
    - Don't be afraid of licenses and use them early on (the longer you wait to more difficult it becomes)

- Seek support from your environment (libraries, compute centres, research coordinators, …)
- Libraries
  - Embrace data stewardship and curation
  - Get ready for the cloud (AAI, PIDs, FAIR)
  - Join forces with your computing department - or become one (seriously!): Bring compute capacity where the data is. If libraries want to be the place where data is preserved then they should build up (or closely partner with) compute capacity.
- Compute centres:
  - Become part of the EOSC federation (AAI, etc.)
    - Enable your users to access resources elsewhere
    - Enable others to access your resources
  - Make (parts of) your infra broadly accessible (e.g. with Virtual Access proposed by EC; using interfaces like Notebooks and MyBinder)
  - Think also "eduroam" - It's based on the promise that you as a provider serve other people with the hope that your people are served by others too when they are abroad → Resource sharing)

14:45 - 14:55 Presentation: Dr Federica Cappelluti: <u>EOSC and Technical universities: what can they bring to each other?</u>

- She is a researcher and also participant in the CESAER collaboration
- Politecnico di Torino is one of the two largest technical universities of Italy (36k students; 1500 PhD students and postdocs)
- Complex world of Engineering/Technology universities RDM. Challenges:
  - Fragmentation: how much data is stored across PoliTo?
  - Scale: Back-up our big data? What to keep, what to delete?
  - HOW CAN WE LEVERAGE EOSC TO ANSWER THESE QUESTIONS?
- Diversity is a big challenge. Different data means need for discipline specific standards
- Awareness about FAIR-ness: She presented statistics about the awareness of FAIR-ness in different disciplines. Common observation: Majority of researchers did not hear about FAIR.
- Innovation: Important to highlight that FAIR does not mean 'open for all'. Commercialisation, industry collaboration is possible.
  - Win-win keywords: Reproducibility, verifiability
- Community standards: No coordinated approach. Few examples were mentioned.
- We must recognise that FAIRification is difficult. We cannot be bottom-up only. Cannot be academic only.
- How to move forward?
  - Co-creation of FAIR standards: Can EOSC facilitate this? (with academia and industry)

- Librarians to participate
- Need to match EOSC and people needs: Is discovery enough? Shall we have more capabilities?
- Training is a must

14:55 - 15:05 Presentation: Laurents Sesink: <u>How can Digital Competence Centers bridge the gap between research intensive universities and the EOSC</u>

- Leiden university is broad in disciplines, not only science
- One of the common topics: Open Science. We address this across 4 lines of actions: Towards Open Science; Towards FAIR Data; Sustainable Research Software; Citizen Science
- If we zoom into infrastructure then we need this at 3 levels: local, national, international. EOSC plays an important role at the international level.
- Leiden Uni is in LERU. - LERU Statement in Dec 2020 was quoted (S4): "*As for the European Open Science Cloud (EOSC), a tremendous amount of work remains to be done culturally and practically to enable the EOSC to engage with individual research-intensive universities.*"
- The Dutch approach to EOSC:
    - Digital CCs (DCCs) that bundle expertise in the field of FAIR Data Stewardship, software and computing.
    - NWO is going to invest so universities can setup DCCs: how to make data fair, sustainable research SW, Research ICT.
        - Leiden Uni also got funding for a DCC
    - Support in a fragmented landscape: connecting local, national and international DCCs to best support the researchers. EOSC is a DCC at the international level.
- S7: The Leiden DCC is not from scratch. It links and enhances existing elements. Exact objectives were mentioned from the slide
- Governance is key, was read out from slide 8.
- Local DCCs cooperate with other local DCCs, with National DCC (e.g. SURF/DANS), and with International DCC (EOSC)
    - The cooperation is different and therefore requires different approval. E.g. Executive board needs to be involved for EOSC cooperation
- Recommendations for libraries & EOSC (S10): <Were read out as they are>
    - Libraries must take on an active role in the feedback loop going from leading practices to implementation at faculty and institutional level and vice versa. By collaborating closely with EOSC we help to co-develop leading practices and implementation guidelines for FAIR data, sustainable research software, Open Access and citizens science in a manner that is informed by first-hand experiences from researchers and support staff working in research intensive universities.
    - To be able to take an active role in the co-development of services and to further the development of the data stewardship- and sustainable research software

profession libraries should invest in building up a number basic skills in the fields Information Technology and Computer Science, next to innovation skills and collaborative skills.
- In addition to this, librarians must be given the opportunity in form of time and funding to take part in leading practices activities to develop the data stewardship - and sustainable research software profession further.
- Libraries can add value on top of EOSC services for FAIR data, sustainable research software and research IT by delivering expertise on reproducible research, registration and (alt)metrics for sharing research software and legal advice on Text and Data mining.
- Libraries need to advise the executive board of research intensive universities to cooperate at a strategic level with EOSC.
- EOSC should by design be inclusive and develop funding schemes to enable libraries to actively participate in activities and projects.

<There was no Q&A after the presentations>

15:05 - 15:45 <u>Breakout Sessions</u>

After the presentations, the participants were divided into three groups – three breakout sessions, named MUSIC, FILMS and BOOKS. Each group had its own facilitator and rapporteur. They discussed the same four questions already listed in the introductory part of this report, and noted their comments using the Google application called Jamboard. The sessions lasted for 40 minutes, after which the participants reconvened in the main room and the groups' discussion results were presented.

The notes below capture the discussions that happened in the break out rooms, underlining the key outcomes of the discussion for each question in each room.

<u>BOOKS ROOM: facilitator Laurents Sesink, rapporteur Anna Rovira</u>

The group had 5 participants incl. The facilitator and rapporteur.

(Approx 10 minutes was spent on finding a rapporteur, and training people about the use of Jamboards.)

In this room the facilitator chose the style of discussing sticky notes that were added to the Jamboard, resulting in more discussions and interactions, but a lower number of post-its compared to other groups.

The main topics of discussion were:

- What is EOSC really, what value can it deliver for an organisation, shall we invest into it and we do what else shall we give up because EOSC brings better value-for-money?
- Libraries manage data for a long time, their knowledge can be updated to the EOSC landscape. The Train-the-trainer structure run by LIBER, FOSTER can be used for EOSC topics.
- Data sharing often requires special knowledge of legal constraints, possibilities. This is something that not even librarians have and they contract experts.
- New methods for recognition of 'data sharing', 'FAIR data production' are needed to change the researchers' mindset.

MUSIC ROOM: facilitator Rafael Reetz, rapporteur Paula Martinez Lavanchy

The group had 4 participants incl. the facilitator and rapporteur, excluding LIBER people, workshop rapporteur and grand rapporteur.

The facilitator in this group used the time for adding stickers to Jamboard instead of discussion. There was nearly zero discussion of the content added on stickers - resulting in a lot of stickers but no deeper analysis.

FILMS ROOM: facilitator and rapporteur Federica Cappelluti

The group had 2 participants incl. the facilitator-rapporteur.

(Approx 10 minutes was spent on finding a rapporteur, and training people about the use of Jamboards.)

The room was basically a dialogue between the two persons. They added only a few sticky notes and only to the first page of Jamboard.

15:45 - 16:05 The results of discussion groups were presented to all participants by the rapporteurs

Books group by Anna Rovira:

- EOSC should consider not only researchers, but also the voice of libraries, also of the governance of the institutions. Also there are other networks in Europe. Broad landscape.
- We appreciate that EOSC wants to hear about us.
- Important that researchers are involved in the development phase not only when the ‘end product’ is ready.

- Important to hear the voice of libraries. They are the one who e.g. give advice to researchers on legal issues, on research integrity.
- How to engage: Build up recognition; Build up funding

Music group by Paula Martinez Lavanchy:

- Q1:
    - Availability of resources is common to both groups
    - For Libraries: Coordination and framework for services
    - For Researchers: Access to data, access to infrastructure, opportunity to collaborate through the services
- Q2:
    - Common: Working together
    - Libraries: Training for researchers, Input to policies, Disciplinary standards
- Q3:
    - Most comments were around training and competence development
    - Clear role of participation
- Q4:
    - Common: User forum
    - FAQ forums
    - Usage stats and analyses

Movies group Federica Cappelluti:

- Q1:
    - Pushing forward the recognition of OS
- Q2:
    - Metadata, semantics that makes EOSC cross-disciplinary. Difficult task!
    - Getting both stakeholders in full comprehension of the overall landscape.
- Q3:
    - To understand the overall workflow, but be able to integrate country-specific, regional differences. (e.g. countries move with different speed on the recognition discussions)
- Q4 (was analysed very quickly):
    - Use some statistical analysis

The workshop series is funded by the European Union under the EOSC Secretariat project