



**Research Libraries, Researchers & the EOSC:
Central-Eastern European Landscape
REPORT FROM WORKSHOP
WRITTEN BY GERGELY SIPOS**

Introduction	1
Participant analysis	2
Workshop structure and atmosphere	2
Breakout group approach analysis	3
Summary of breakout outcomes	4
Appendix - Event notes	5

Introduction

The workshop was held on 20 January 2021, between 14:30 - 16:15 CET on Zoom. The event was the second in a series of five workshops organized by Scientific Knowledge Services (SKS) 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 2nd workshop focused on the Eastern European landscape. The target audience were researchers and research library staff.

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/w2/>

Participant analysis

Excluding the speakers and organiser staff **44 people registered to the event from 36 institutes from 13 countries** with the following distribution: Austria (1), Bulgaria (1), Croatia (3), Czech Republic (2), Hungary (3), Italy (1), Lithuania (3), Netherlands (4), Poland (8), Republic of Moldova (1), Romania (11), Slovenia (4), UK (2).

The **actual participation was 34 people from 31 institutes from 12 countries (77% turnout rate)** with the following distribution: Austria (1), Croatia (2), Czech Republic (2), Hungary (1), Italy (1), Lithuania (2), Netherlands (3), Poland (7), Republic of Moldova (1), Romania (8), Slovenia (4), UK (2).

The gender distribution¹ of the 34 attendees was: **22 female (65%) and 12 male (35%)**.

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

The structure of the event matched well with the desired goals and constraints:

- Be educational to establish engagement with the target group.
- Gather as much feedback and input from the target group as possible.
- Run a relatively short event without break to maximise participation and avoid audience loss. (Which still happened to ~13 participants when switching from the plenary room to the breakout rooms.)

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



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 ask for clarifications on certain points that were presented before starting the discussion. This also could have energised the audience a bit more before the breakout.

Breakout group approach analysis

There were 9 participants in the **BLUE** room, including the facilitator, the rapporteur, a SKS support person (Damla Bal), and the Workshop rapporteur (Gergely Sipos). Participants in this room were mostly from 'libraries', there were only two persons who contributed with input about 'researchers'.

There were only 5 participants in the **YELLOW** room, including the facilitator, rapporteur, the Grand Reporter and a LIBER support staff. The real contributions came from 3 persons, two with a researcher, and one with a librarian background.

There were 10 participants in the **GREEN** room including the facilitator, rapporteur (who was also the LIBER support person). The audience was a good mix of researchers and research libraries, many participants having both roles at the same time.

In all three rooms the **facilitators followed a similar style**: triggering discussion on each Jamboard page by contributing the first 1-2 sticky notes, then opening the floor for more contributions and allowing participants to explain their sticky note items. **This approach worked well in all the three groups** with the participants quickly catching up with the facilitator and joining the discussion. Having the **speakers as facilitators was a good choice**: they could refer back to certain points in their presentation when they introduced a question or a specific sticky note, making the discussion more dynamic, the two parts of the event better connected.

The **technical setup (Zoom and Jamboard) was suitable and smoothly working** for this kind of group exercise.

Noticeable the GREEN room was the most dynamic - because of the most people involved. While the other two groups also had active participants the input in the YELLOW room was rather restricted (originating from 3 persons), and **the YELLOW group would have benefited from a few more participants**.

Overall, **the size of all the groups worked well for this type of interactive way of working**, and given the rather short length of the workshop, the structure allowed the gathering of information from diverse stakeholders.

The **benefits of the chosen facilitation approach** were

- Participants could elaborate what they meant with their short notes.
- Allowed participants to ask questions and clarifications from each other.
- Enabled participants to better get to know each other, including their background.

The observable **disadvantages of this approach** were



- Relatively few sticky notes have been added to each Jamboard page (e.g. compare them with the Jamboards of the the Technical Universities groups)
- Some of the sticky notes were discussed, while others were not (the participants chose whether they want to talk about their note or not)
- The discussed notes were more likely to be included in the oral report of the group rapporteur (while these notes may not represent the 'most important' points)
- The input in the YELLOW group was from 3 persons only.

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?

- Find data and services irrespective to where those are physically stored, a single framework for simpler discovery.
- An interface for local (i.e. institutional, national) contributions to enter the European/global scene.
- A unified set of data standards that libraries can follow/implement to share scientific data and other resources in a compliant way.

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

- Any institutional/national/regional contribution that facilitates Open Science can be relevant - although, it was noted that currently it is hard to connect such contribution with EOSC, unless the contributor is closely following EOSC developments.
- Researchers often feel that their data are not interesting for others, their data are 'private property'. Policies, incentives, measures in proposals are needed to change this mentality.
- Training for scientists (by libraries), as well as successful examples about the reuse of long-tail research data via EOSC could motivate researchers.
- It is still a big open question whether researchers will use EOSC directly (B-2-C relationship), or EOSC will empower local institutes (e.g. libraries in a B-2-B relationship) so they can provide better services to users locally. Whatever will be the focus, local support will remain key for success in research.

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

- A prerequisite of engagement is to highlight the benefits for the stakeholders (i.e. for libraries in our case).



- Webinars are powerful engagement methods. The EOSC Association should have regular update Webinars about EOSC.
- Staying informed and connected seems the easiest through the national institute that are in EOSC. However these national representatives, who participate in the highest level of EOSC, should be well connected to their national Open Science landscape and transfer/disseminate the updates to them.
- Increase researcher engagement by involving more researchers in the EC funded EOSC projects.
- Establish EOSC-related working groups within the national 'Learned societies'.

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

- Data about EOSC providers, including scale of their use, quality assessment, statistics about services.
- Data about the use of EOSC for interdisciplinary purposes.
- 'Bad examples' are also powerful. Such as the data management horror stories from the University Library of the Vrije Universiteit Amsterdam.

Appendix - Event notes

Giannis Tsakonas introduced the event, its goals and its structure, then invited the 3 speakers:

- Judit Fazekas-Paragh
- Ana Slavec
- András Holl

14:35 - 14:45 Presentation Judit Fazekas-Paragh, Head of Education and Research Support Department, University of Debrecen University and National Library OpenAIRE NOAD - Hungary: How libraries support open science?

Ms Fazekas-Paragh started the presentation with a historical outlook on how libraries contribute for centuries to the 4 main stages of data management: long-term preservation, curation, publication, dissemination. By extending this expertise to the XXIst century, and taking up additional capabilities, libraries can continue to be pillars of Open Science. Particular expertise that can be taken are Repository manager; Data curator; Data steward, Data librarian. All these roles together formulate 'scholarly communication'.

Scholarly communication also requires requires infrastructures (repositories, data repositories publication platforms), and training for their users (about tools, services, Research integrity, FAIR data, EOSC). These can help us achieve transparency and timely publication which is of extreme importance as we have seen because of COVID.



A regional project was launched with European Commission funding in September 2019 to establish infrastructures and training in South East Europe: NI4OS-Europe (National Initiatives for Open Science in Europe). NI4OS-Europe started with mapping the main stakeholders in its member states: Greece, Cyprus, Bulgaria, Croatia, Hungary, Romania, Slovenia, Serbia, Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, Moldova, Armenia and Georgia. The mapping was performed by role of the stakeholder, such as funder, facilitator creator of Open Science objects, etc. (The complete mapping is available on the NI4OS-Europe website). The second most important work area of NI4OS-Europe is the creation of national Open Science Community initiatives which serve as models for each other and for countries outside the NI4OS-Europe region.

Ms Fazekas-Paragh then introduced the experience of one of the NI4OS-Europe member states, Hungary. The libraries that facilitate Open Science in Hungary came together into a Working Group, called HUNOR (Hungarian Open Repositories). The highlight of HUNOR's recent contribution was a meetup series that began in autumn 2020 and introduced to the broader Open Science community the basics of openness and FAIR data principles; Various tools and services through real-life experiences; The benefits of DMPs. The effort is funded by the National Research Development and Innovation Office, the main research funder in Hungary.

14:45 - 14:55 Presentation Ana Slavec, researcher at the InnoRenew CoE, Slovenia: How could libraries support researchers in making data FAIR?

Ms Slavec started her presentation with a very short reminder of what FAIR stands for: Findability, Accessibility, Interoperability, Reusability of research data, tools, publications and other artefacts. One of the EOSC Working Groups was focusing on FAIR and produced 4 publications so far:

- Six Recommendations for Implementation of FAIR Practice
- A Persistent Identifier (IPD) policy for the European Open Science Cloud
- Recommendations on FAIR Metrics for EOSC
- Recommendations on certifying services required to enable FAIR within EOSC

From these documents the Working Group identified the four main challenges to tackle for the widespread adoption of FAIR-ness:

1. Currently there is a low awareness of Open Science and FAIR data concepts within research communities.
2. Researchers still lack training opportunities to learn about FAIR approaches, methods and tools.
3. There is big difference between the disciplines in their maturity of FAIR adoption, and there are no global rules that every discipline could adopt to take the next step. Each community needs to agree internally on data formats, vocabularies, metadata standards, data sharing procedures.
4. FAIR data sharing is not recognised and rewarded, therefore real incentives are often lacking for the implementation.

Four main challenges

to tackle for the widespread adoption of FAIR-ness:

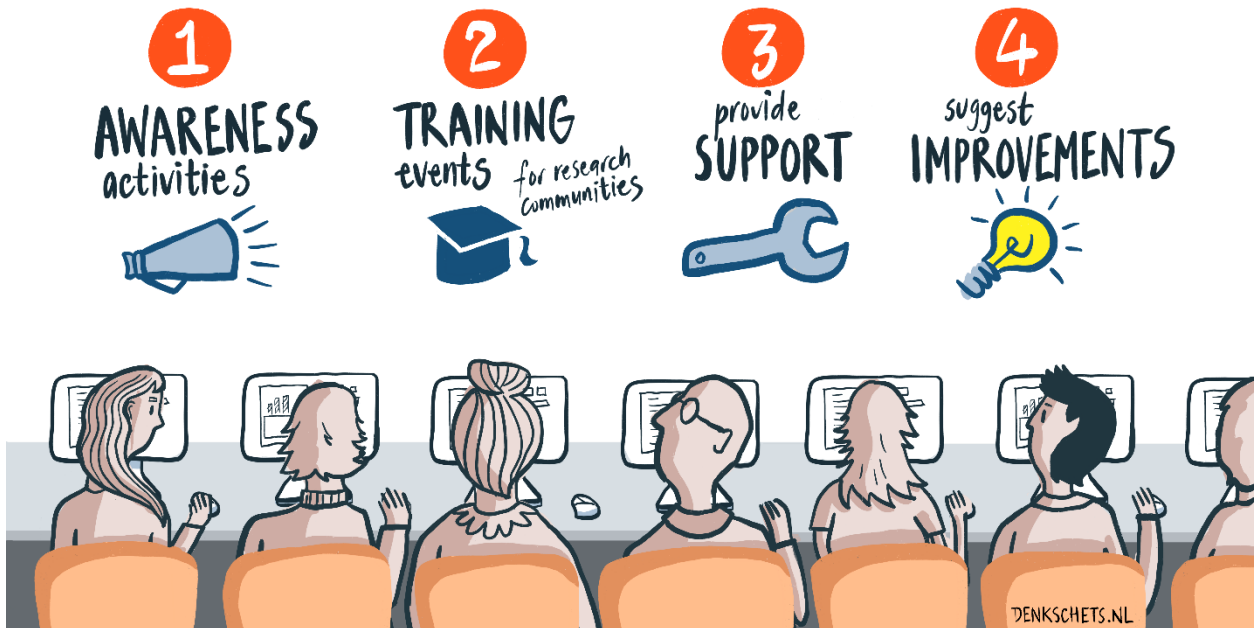


Figure 1: Four main challenges to tackle for the widespread adoption of FAIR-ness (Denkschets, 2021)

Ms Slavec then stated her opinion on how libraries can help research communities tackle these challenges. She suggest libraries to

1. Perform awareness raising activities to promote FAIR principles.
2. Organise training events targeting the research communities.
3. Provide (community-specific) support in making research data FAIR
4. Suggest improvements to current research assessment practices

In the final part of her presentation Ms Slavec highlighted various resources and examples that can help libraries implement these four activities:

- First she introduced the 'Role landscape' that was prepared by the 'EOSC Skills & Training Working Group'. The landscape can help libraries understand the stakeholders that participate in EOSC, and conduct training sessions for specific groups of the landscape.
- Second she introduced the RDA outcome that provides '23 practical, free and online resources that can be used to incorporate research and data management into the practice of librarianship'.
- Third she showed the example of an international Summer School that was organised in 2019 to train Data Stewards. Data Steward is a role that librarians could take on in the



Open Science era. The school was a pilot version with only 5 attendees (one of which was the speaker). The event was successful, but unfortunately could not grow into a full size event in 2020 because of COVID. We all hope that can be organised again in the future.

14:55 - 15:05 Presentation András Holl, Deputy Director, Library and Information Centre of the Hungarian Academy of Sciences, Budapest, Hungary: Libraries & research data

Mr Holl began his talk with the classification of research data and describe the fundamental difference in the support they need:

- The 'big science' research data: We often think this is the only type. It's produced by big experiments that are well funded, centrally coordinated, has data management pipelines and therefore are in a relatively easy position to make their data FAIR. An example is the data from the Kepler satellite stored at Mikulski Archives for Space Telescopes.
- The 'long tail' of research data: Data produced by individual scientists or small teams, in fields and in institutions that are not well funded. Their data is also important, but have different characteristics: they are very diverse in formats, metadata, storage, representation, etc.

Mr Holl argued that libraries have long-standing experience in managing lots of small datasets. Libraries make documents discoverable using metadata. Libraries provide identifiers and classification systems. Libraries have a long history of helping researchers. Therefore, the management of long tail data can be a suitable task for librarians in the EOSC era too.

Talking to the researchers, Andras suggests to 'Ask your librarian for advice!'. Libraries (also in OpenAIRE) operate 'Open Access virtual desks' where researchers can get advice on FAIR matters.

In the next part of the presentation Mr Holl showed the goal of making data FAIR: Then it can be linked to scientific publications and therefore discovered within a proper context. An example was shown with an arXiv paper linking to SIMBAD and MAST data products. The XML description of this linkage was also displayed. The linkage comes with another benefit: the reviewers of the scientific paper, who are domain experts, can assess the dataset too. This would ensure quality control which cannot be provided by librarians who are not domain experts.

Metadata are like toothbrushes...



Figure 2: Metadata are like toothbrushes (Denkschets, 2021)

Mr Holl then continued with a few additional advice for researchers:

- Libraries are unable to handle too big, too complex and too specific data. Keep it simple and use common formats (TXT, XML, JSON, PDF, etc.).
- Libraries can advise on licenses too, best fitting the sensitivity and other relevant aspects of the data.
- 'Grey literature' is a significant segment of science project outcomes (technical documentations, DMPs, project reports). Archiving those is an important task, and one that is very suitable for libraries.
- Keep in mind that FAIR is a journey. There is no such thing as perfectly FAIR data. FAIRness can be always improved.

In the concluding part Mr Holl warned the librarian community to be careful with FAIRness. If we set the threshold of FAIRness too high, we would lose datasets and communities. If it is too low, we would lose functionality. If libraries make data discoverable and accessible then it can be already a huge improvement and pragmatic approach. Commercial publishers try to dictate the terms of archival but these seem to differ rather significantly from those desired by libraries. This disagreement was expressed by the library community through COAR's statement. (It was not



explained further by the speaker, but there is a link provided on the slides: <https://www.coar-repositories.org/news-updates/input-to-data-repository-selection-criteria-that-matter/>)

15:05 - 15:45 Breakout Sessions

After the presentations, the participants were divided into three groups – three breakout sessions, named BLUE, YELLOW and GREEN. 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.

BLUE ROOM: facilitator Ana Slavec, rapporteur Mitja V. Iskrčić

There were 9 participants in the room, including the facilitator, the rapporteur, a LIBER support person (Damla Bal), and the Workshop rapporteur (Gergely Sipos). Participants in this room were mostly from 'libraries', there were only two persons who contributed with input about 'researchers'.

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

For research libraries the EOSC would bring the biggest benefit as a virtual environment where librarians and the people supported by them can find data and services irrespective to where those are physically stored. It was noted that having already an initial set of mature and relevant entries in EOSC would greatly help us demonstrate the concept and 'sell' the vision to additional providers.

Concerning researchers, the biggest value mentioned in EOSC is having a system that brings research data and tools into a single framework for simpler discovery. Access is further simplified through harmonised access conditions, rules, policies. EOSC can be like a 'Google for research'.

Another value that was mentioned is seamless integration among providers, and acting as an interface for 'local contributions' to enter the European/global scene.

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



Connecting with EOSC as a national infrastructure/individual is hard, unless the person/institute is involved in the EOSC design and developments from the start. Training on how to engage with EOSC can be definitely useful. Any training can be useful in this context, starting from Open Science and FAIR practices to technical level on the use of specific tools and services. Research libraries can take a leading role here. An example was mentioned from a technical university in Romania: “We introduce this year training on Data Management Planning, using our own repositories, FAIR data and tools during the course. We hope this will be a successful project.”

The discussion then continued with an observation that change of researchers mentality is also needed, because researchers often feel that their data is not interesting for others, or that their data is ‘private property’.

Successful examples of reuse of data from the long-tail can help. OpenAIRE recognised this need and started collecting examples of reuse of long-tail research data at <https://www.openaire.eu/data-reuse-use-cases>

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

The first contribution emphasised that Webinars are powerful engagement methods (about Open Science and about scientific repositories). The EOSC Association should have regular update Webinars about EOSC to all the stakeholders.

The role of National representatives came up, and it was pointed out that it’s very important that the national representatives who participate in the highest level of EOSC board should be well connected to their national Open Science landscape and transfer/disseminate the updates about EOSC within their nation. Libraries could act as ‘filters and multipliers’ of EOSC information towards researchers in this European to national knowledge transfer.

Convincing researchers about the benefits of Open Science is hard and long. Libraries should talk also to funders to make them define appropriate policies, e.g. to incentivise researchers in implementing Open Science.

Another suggestion to increase researcher engagement was to raise the number of researchers that are funded in EOSC-related projects.

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

Every EOSC service should have a user group which is open and where feedback can be captured for providers. For example, in Slovenia and Ukraine there are national communities of EOSC stakeholders.

Would be important to present available data about EOSC providers and use, including quality assessment, statistics about services.

We need more data on interdisciplinary research and data about the use of EOSC for interdisciplinary purposes.



Sometimes the 'bad examples' are also powerful. Such as the data management horror stories that were collected by the University Library of the Vrije Universiteit Amsterdam (<https://ub.vu.nl/nl/ub-voor-onderzoekers/research-data-support/data-horror-stories/index.aspx>)

YELLOW ROOM: facilitator Judit Fazekas-Paragh, rapporteur Cristina Lincaru

There were only 5 participants in this room, including the facilitator, rapporteur, the Grand Reporter and a LIBER support staff. The real contributions came from 3 persons, two with a researcher, and one with a librarian background.

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

Having a unified set of data standards is important. For example 'open science projects' today kind of use different standards. This landscape should be simplified.

Reusability is another important aspect for researchers. EOSC can bring two values: Make the researcher's data more reused, and empower the researcher to more easily reuse other researchers' data. Other items that have been added but were not discussed:

For researchers:

- Open Science communication
- Access to latest research results
- Common platform for sharing results from EU projects
- Avoid duplication of research topics
- Improve networking

For librarians:

- One common repository reduces the experiences need in assisting researchers

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

The group discussed the meaning of 'open data' in EOSC, because certain researchers reject 'open data' (researchers in law were mentioned as an example). Judit responded that 'open' refers to metadata and not to the data itself. Open data is an element to FAIR data, but does not mean 'accessible to anyone'.

Open Science requires a radical change in research in both top-down and bottom-up to facilitate the adoption and implementation of new practices. Judit mentioned that in Hungary both top-down and bottom-up movements started in parallel: The Ministry (The Funder of research) introduced Open Science measures in proposals, while the Open Science practitioners started education for researchers on how to implement open science practices.

Other items that have been added but were not discussed:

Another point that was discussed is the need to 'change the strategy in research'.

For researchers:

- Researchers should be assured about the safety of their data (licenses, reuse)



- Researcher engagement (which could be achieved by modulating reward system)
- Competition framework for research projects
- New academic metric
- Librarian and Researchers Partnership

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

Participants agreed that a prerequisite to engagement is to highlight the benefits for the stakeholders to involve them.

Due to the time constraints the answers that came together on the Jamboard were not discussed:

Researchers:

- Public consultations through questionnaires
- Co-creation of policies
- Clear goals and steps to achieve them
- Collaboration and constant communication
- Projects and networks

Librarians:

- Money for infrastructure and manpower

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

Due to the time constraints the answers that came together on the Jamboard were not discussed:

Researchers:

- Short feedback form for every EOSC visitor/user to identify their satisfaction level
- Internal kind of social network for all interested stakeholders
- Real-time dialogue (e.g. real-time support system)
- Regular gatherings - online or face-to-face
- Simple procedures of communication in real time for big number of stakeholders

Librarians:

- Encourage policy makers for improving reward systems

GREEN ROOM: facilitator András Holl, rapporteur Tiberius Ignat

There were 10 participants in this room including the facilitator, rapporteur (who was also the LIBER support person). The audience was a good mix of researchers and research libraries, many participants having both roles at the same time.



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

The participants agreed that EOSC brings together different stakeholders, and EOSC is still in the making. Currently 'the average researcher' knows little about its value/benefits. Once assembled, the ultimate value will be delivered to researchers. However, easy discovery of data, discovery of standards to use for data production will be important value for both researchers and libraries. Common standards can help libraries be more successful in their current role, which is supporting researchers.

The participants contributed to the Jamboard with additional points (but those were not discussed):

For researchers:

- Build the necessary Service Level Agreements
- Accelerate standards in sharing data
- Enhance the process for collective intelligence

For research libraries:

- Long term preservation of data
- Support the Open Science transformation
- Expand European Open Science effort to global partnerships

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

András kickstarted the discussion by restating one of the first messages from his presentation: Each and every dataset counts, EOSC should care about the long-tail data. EOSC should ensure that the Rule of Participation is inclusive enough to be able to support each and every research dataset.

Hilary Hanahoe mentioned that she believes it is still a big open question whether researchers will use EOSC directly, or EOSC will empower local institutes to provide better services to local users. András' opinion is that both cases will be prominent in EOSC. Others agreed but emphasised that local support will remain key for success in research.

One comment raised the importance to have standards in EOSC that enable data sharing across organisations. Others argued that researchers are more concerned about sharing data than the reuse/integration of others' data.

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

András mentioned that he was involved in EOSC Working Groups. Others mentioned that the easiest way is through the national institute that is in EOSC. But most of the participants even in this group do not know which are their national institutes. András provided information where this can be found on the EOSC Portal (under About / Governance).

An idea came up to establish EOSC-related working groups within the national 'Learned societies', which was found to be a useful suggestion by the participants.



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

The discussion focussed on how to reach researchers and promote EOSC to them (as this is a prerequisite of gathering feedback from them) A number of ideas have been discussed:

- Put a banner on Zenodo to spread information about EOSC - because many researchers already know and visit Zenodo.
- Promoting EOSC is difficult because it's still fuzzy. Maybe some form of infographics would be more suitable.
- There are many EOSC websites (EOSC Portal, EOSC Secretariat, Project websites. Having one central place would be much better.

(The discussion stopped at this point because the breakout time was over.)

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

16:05 - 16:15 Final remarks

16:15 Workshop Closes



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