Notes from the EOSCpilot Governance Development Forum in Helsinki, 9 May 2017

Introduction

Per Öster, EOSCpilot WP2 leader, CSC – IT Center for Science welcomed the participants and explained the main aim of the EOSCpilot Governance Development Forum (EGDF), which is to enable stakeholders to contribute to the European Open Science Cloud (EOSC) governance development. He highlighted in particular the science cases and interoperability as motivators for EOSC and EOSCpilot - there are national facilities and resources that need to be visible and used on the European level.

This EGDF "pop-up workshop", organized in conjunction with the 6th ERIC network meeting, focused specifically on research infrastructure´s needs and expectations on EOSC. Two panels facilitated the discussion. The panelists represented a number of research infrastructures from various research domains.

Per opened up for a virtual follow-up meeting in the fall to continue the discussion on this topic. More material and input from stakeholders to the EOSC governance development can be found on the website: eoscpilot.eu (click "about" to find the EGDF space).

First Panel discussion: Research Infrastructures expectations on the European Open Science Cloud.

Panel chair: Per Öster, EOSCpilot WP2 leader, CSC – IT Center for Science
Panelists: Franciska de Jong, CLARIN ERIC, Christine Kubiak, ECRIN ERIC, Niklas Blomberg, Elixir ERIC, Volker Röhling, ECCSEL, Francisco Colomer, JIVE, Sanna Sorvari, ACTRIS ESFRI.

Panel topics:
• Which kind of services would you expect the EOSC to provide?
• From the point of view of the users of your infrastructure / researchers that you Research Infrastructure represents, which kind of benefits do you expect EOSC providing them?
• What is characteristic of your Research Infrastructure that should be taken into account when designing EOSC?

Per Öster initiated the discussion by mentioning that the scope of EOSC is still quite unclear, and that research infrastructures' concrete expectations on EOSC - what issues EOSC can solve, and what services to provide - will play an important role in helping defining what EOSC should become.

Francisco Colomer, JIVE, described that there is a transition going on in their research infrastructure from developing everything themselves for the astronomy community (tools etc.), to start outsourcing part of the work to be able to focus on their core services. EOSC could share part of this work, as **cloud services for the astronomy community are much needed.**

For the life science domain, the main benefit of EOSC is expertise, according to Niklas Blomberg, Elixir ERIC. More of the non-technical expertise is needed, for instance in developing ways to achieve mutual recognition, by simple standards, codes of conduct etc., with a set of shared basic rules. EOSC could be useful for this kind of work.

Sanna Sorvari, ATRIS ESFRI, pointed out that they are interested in checking solutions (technical and non-technical) outside their own community, for instance on interoperability and workflows of data, where they can benefit a lot from collaborating with e-Infrastructures. **EOSC could help in getting**

e-Infrastructures more sustainable, for research infrastructures to get a similar counterpart, since one major hurdle is that they have very different funding schemes. Per Öster noted that sustainability can be quite stable for some e-infrastructures, like CSC in Finland for instance and similar organisations in other countries, but the question is how the e-Infrastructures' activities are directed to the actions needed. Niklas Blomberg suggested to use the expression "**Life cycle management" instead of sustainability, to highlight that some things should not exist forever.**

Volker Röhling, ECCSEL, stated that they are in the early stage of cloud storage and data reuse processes (related to CO2 capture, transport and storage technologies research), with a big range among the partners. He concluded that **some research domains currently experience a paradigm shift where researchers start moving over to cloud storage**. **EOSC could assist in provide some training how to make this shift**, for instance in curation of data, how to make data useful for a longer period.

CLARIN ERIC has a big diversity of end-users for their language resources and other services, which was highlighted by Franciska de Jong. **EOSC could assist in very clear, user-friendly tutorials, help-desk support for the various services in order to attract this diversity of users**. If there were to be only training and no on-going support, there is a risk to miss a large part of the CLARIN ERIC community. The diversity not only in different levels of skills but the multilinguality itself can multiply the number of questions user can put to services.

Christine Kubiak, ECRIN ERIC, described the concerns around **medical data, which needs specific consent**, since it is collected for clinical research. **Interoperability between countries and metadata, where there are currently no common standards for this kind of data are two main issues where EOSC could play a role**. Another issue is the confidential data, which needs high security. Due to these aspects, reuse of data poses quite special questions and concerns.

The reuse of data
Per Öster asked the panelist from a about the reuse of data from a general perspective. **How much is data really reused**, **both within a community and across domains**? How to facilitate the reuse?

Sanna Sorvari stated that within their community, there is reuse of data but she wondered if there may be a bit too much "hype" around cross-domain data reuse, since she has not seen many examples of this. **Sanna concluded that if reuse of data across domains is to be the main aim of EOSC, she does not think it is going to be achieved. She recommended for EOSC to facilitate for interoperability first** and for this there is a need to discuss more on who are the users of EOSC.

Francisco Colomer mentioned that they have used meteorological data within astronomy as an example of successful data reuse across domains. **The limitations are that the astronomy data remains at the producer or at the telescope, and there is a risk that the data disappears, if it is depending on the resources of the provider. EOSC could help in making this data available in a more guaranteed way.** A pilot could easily be done in astronomy, since there are no privacy issues. The challenge lies more in the volume of data - computation and archiving are the most common problems in the astronomy community.

A person from the audience asked **how private sector can use data provided by research infrastructures**. Niklas Blomberg stated that **open science data is an important tool for local innovation in Europe**. **Private sector need to be educated in how research data can be of use - modern knowledge economy.** Researchers should be trained to be able to approach companies and work with these issues. Sensitive access controlled data does not mean it has to be totally closed.

Role of cluster projects in EOSC governance set-up
Per Öster asked the panelists their views on **how to use cluster projects in EOSC governance and set-up.**

Franciska de Jong described cluster projects as a good way to "simulate multidisciplinary collaboration", and reminded that **joint problems and solutions in cluster projects can be used in EOSC.** Niklas

Blomberg concluded that **organisational interoperability can be reused from cluster projects to the EOSC context**. In cluster projects it has been noticed that rather small issues can become big blockers for the users, for instance different access schemes. This has been addressed organisational interoperability in cluster projects. EOSC is similar but on the macro level.

Sanna Sorvari reminded that cluster project is not only a useful too towards EOSC, it is also about community building for their own domain. But **cluster projects can act as important dialogue partners for EOSC as well as providing practical example for the governance** (a board to agree on common strategies for instance). Sanna also called for **research infrastrucutres to have a natural role in EOSC governance** as they are not just users, but also providers. **Many other panelists also underlined research infrastructures dual role as both users and providers in EOSC.**

Second Panel discussion: Optimal governance framework for EOSC from Research infrastructures' point of view

Panel chair: Per Öster, EOSCpilot WP2 leader, CSC – IT Center for Science
Panelists: Ron Dekker, CESSDA ERIC, Jan-Eric Litton, BBMRI ERIC, Fruzsina Molnar-Gabor, EMBL, Jonathan Taylor, ESS ERIC, Bahne Stechmann, EU-OPENSCREEN ERIC

Panel topics:
• The optimal governance framework for EOSC from the Research Infrastructure point of view
• The characteristics of different Research Infrastructures that should be taken into account in the governance of EOSC
• Important aspects of the governance in EOSC?
• What the governance of EOSC should consist of
• Who should form the governance of EOSC
• What kind of funding streams (instruments) are foreseen to fund EOSC. How Research Infrastructures should contribute to this or should they?

Per Öster asked the panelists to tell about **experiences when forming ERICs and research infrastructures that could be useful when developing EOSC governance.**

Ron Dekker, CESSDA ERIC, explained that when forming the CESSDA consortium, they had to convince countries (national service providers) that they had an added value of being a consortium at the European level. This was hard and required a lot of work. He concluded that **EOSC should also work on making the added value clear and visible for different stakeholders.**

Fruzsina Molnar-Gabor, EMBL, mentioned that the harmonization expected to come with General Data Protection Regulation (GDPR) will not come. Instead, implementation will vary on the national level. **EOSC could have a role in showing the direction for how GDPR could be implemented.** Jan-Eric Litton, BBMRI ERIC, agreed that countries are on extremely different levels on preparing for GDPR implementation, and that there is a need to help the countries calibrate the unknown in the GDPR regulation.

Jonathan Taylor, ESS ERIC concluded that their governance has changed a lot over the years, and since EOSC Governance is even more complex, it will most probably change over time as well. He suggested for **EOSC governance to have some flexibility but to be based on some core values of what EOSC is about to deliver.** According to Jonathan, the **core of EOSC is the data and the end product is science**. Per Öster agreed that **governance of EOSC needs evolutionary thinking.**

Bahne Stechmann*, EU-OPENSCREEN ERIC* suggested for **EOSC to lead the discussion with funders and encourage funders to make it mandatory for the data to be available.**

Jan-Eric Litton described scattered landscape in Europe as a challenge when forming BBMRI-ERIC, with different ministries responsible etc. Still he believed that there starts to be an understanding that it is important to make the cloud environment part of personal data. The aim should be to make **cross border exchange of personal data possible in Europe**. Per Öster stated that **it is the science cases that can make the dialogue happen between different ministries, and EOSC should try to give the science cases that can initiate such a dialogue.**

Ron Dekker, CESSDA ERIC, called for **the science community to be the target group of EOSC** and suggested rather to go for scientists than for countries when setting up EOSC governance. He suggested not only focusing on the policies, but also rather going for many pilots that may not all succeed when setting up EOSC governance.

Bahne Stechmann told about an experience in *EU-OPENSCREEN ERIC* he believed is true for almost all ERICs: in the preparatory phase, they discussed with scientists for three years. After that, they went to the funders, and had the same discussion for five more years. The conclusion is that **EOSC needs to have the funders onboard from the beginning, and to be able to convince funders EOSC needs to find good arguments for the added value.**

Niklas Blomberg mentioned that ESFRIS are experienced in negotiation with Ministries versus science communities. He also stated that **there is no need to create new legal structures to solve data issues and similar, but instead for EOSC governance to leverage on what is already there.**

Jonathan Taylor viewed **EOSC added value as using data more effectively**. Per Öster asked if ESS ERIC take their multiple users into account in the governance of their facilities and Jonatan explained that they are taken into account in the operating strategy but not in the governance as such.

Some final remarks was that **EOSC governance should be as simple as possible**, and that it should reflect what EOSC actually wants, actively involving main stakeholders. **ESOC should follow the principle of subsidiarity and not redo what ERICs are already successfully doing**. There was also a suggestion for EC to take the responsibility of financing the first 2-3 years of EOSC for a successful start-up period.