Notes: EOSCpilot Governance Development Forum workshop in Tallinn, 2 – 3 October, 2017

Author(s)	Saara Kontro, Riina Salmivalli
Date	17.10.2017

1. INTRODUCTION

The EOSC and EOSC governance are being designed at the moment. In the EOSCpilot governance development forum activities the representatives of all stakeholder groups are taken on board in the discussion and definition of how the EOSC should be governed. In the EOSCpilot governance development forum workshop that took place on October 2 – 3, 2017 in Tallinn, in conjunction with the e-Infrastructures Reflection Group Workshop under the auspices of the Estonian EU Presidency of the European Union, national e-IRG delegates, representatives of member states, advisors to the ministry as well as scientists and service providers were invited to take part in the debate and plan of the future EOSC. There were all together 53 participants in the workshop.

In this document the discussions of the two panels that took place during the first workshop day are reported.

2. FIRST PANEL DISCUSSION: MEMBER STATES' ROLE AND ENGAGEMENT IN THE EOSC GOVERNANCE FRAMEWORK

Panel Chair: Leif Laaksonen, CSC - IT Center for Science

Panelists: Toivo Räim, Estonia; Gabriele Von Voigt, Germany; Françoise Genova, France; Hanifeh Khayyeri, Sweden; Ivan Maric, Croatia.

Panel topics:

- What is the position of your country or organization towards the EOSC declaration that has been produced by EC as outcome of the European Open Science Cloud Summit (12 June 2017)?
- Does your country or organization support the declaration? Are there some particular aspects in the declaration that you would like to highlight?

Prior to the panel discussion Matthew Dovey, JISC, EOSCpilot task leader working on the Federated governance framework for the EOSC, presented the latest developments in drafting the EOSC governance framework, which panelists were invited to comment.

All member countries represented in the panel expressed their support in the building of EOSC and see that the EOSC declaration has good topics although not all have official position on it, yet. The importance of Member States and Associated Countries in the governance and decision making of the EOSC was highlighted in the discussion. There is a need to find and clarify a proper funding model for the EOSC. The panelists identified the ERAC Standing Working Group on Open Science and Innovation as a proper vehicle for discussion between the EC and MS on the development of the governance model for the EOSC. The wise use of resources was stressed, not duplicating things already done in other initiatives. There should be a clear focus on interoperability of all layers (policy, organizational, semantic, and technical) and member states should agree on mechanisms for interoperability with the support of European Commission to obtain sustainability of the EOSC. ESFRI forum, e-IRG and RDA were identified as important stakeholders for the EOSC.

Views of different countries

Germany

Germany is keen on EOSC, but there is no information if EOSC declaration has been signed, yet. It is the initiator of the GO FAIR initiative[1] with the Netherlands. France is the third country that has now joined the initiative. It is also very active in implementing FAIR principles. When looking at the European scene usually the governance of infrastructures is handled in big European projects, but member countries still have a strong position as enabler since the investments come from the member states. This should be taken into account in the EOSC as well.

Estonia

In Estonia the Estonian Computational Infrastructure was established in 2011 when the four major IT institutes joined their force together. Structural funds were used for realizing the infrastructure. There is a national open science initiative run by the research council. Estonia supports the building of the EOSC and sees that the EOSC declaration has good topics. Member states have an important role in the EOSC governance and this should be visible in the upcoming governance model. Another pressing issue is to find a

[1] https://www.dtls.nl/fair-data/go-fair/

proper funding model for the EOSC. Estonia has experience of designing large infrastructures and systems, for instance the example of xRoad [2]of Estonia could be used in the design and implementation of the EOSC.

Sweden

Sweden has no official position on the EOSC declaration yet. However the EOSC declaration has topics that are in line with the Swedish government objectives in the area of open science and research data management. The resources should be used wisely when developing the EOSC in order not to duplicate things that are already under work or done in FAIR and other initiatives. The ERAC Standing Working Group on Open Science and Innovation has been identified as a proper vehicle for discussion between the EC and MS on the development of the governance model for the EOSC, but not all member states are aware of this.

France

France agrees with the Declaration sections regarding data culture and FAIR data. The role of the RDA and of other organizations such as the W3C and FORCE11 has to be emphasized. The implementation level raises many issues. In particular, the EOSC should federate resources, but not integrate them in a single framework. It should build on existing resources at the disciplinary, national and local level. Moreover, the Declaration lacks of a clear focus on interoperability.

- Member States could agree on an interoperability mechanisms at different levels (technical, semantic, operational and legal), as described in the European Interoperability Framework.
- These interoperability mechanisms will help to share and reuse the national e-infrastructures at a pan-European level.
- The European Commission should support these interoperability mechanisms to achieve the sustainability of the EOSC.
- Open science includes open access to publications, which should be fully part of the EOSC, and this has components also at the local, national and disciplinary levels.

The governance and funding sections of the Declaration need to be to be revised. For France, the governance of the EOSC should be endorsed by the Member States and Associated Countries through the ERAC SWG on Open Science and Innovation. EOSC governance should be inspired by the governance set in place for research infrastructures in general and in ESFRI more specifically. It should include formal gathering of the scientific users' needs (beyond users involved in EOSC).The ESFRIs should be considered as a major component of the EOSC and the ESFRI Forum should be an important stakeholder of the EOSC.

France considers to structure its national e-infrastructures into a French Open Science Cloud. France wishes to nominate a representative at the ministry level in the Board of Funders of the EOSC. France encourages its scientific communities to be involved in the EOSC through the ESFRIs. Some of the research infrastructures in the national roadmap are or include disciplinary data infrastructures and the way they will interface with EOSC has to be assessed.

Croatia

The Croatian representative expressed his criticism on EOSC and EOSCpilot: There is very little information available before getting the invitation to participate in the panel. There is no critical mass of organizations or people involved and engaged in the development and discussion about EOSC. There is a need to communicate to countries, research communities and institutions. There are nice visions but no details are existing. The governance model should be a representation of countries, not federation of countries. An all-inclusive model should be in place. A much better understanding of governance and business models are needed to be able to discuss about the details of EOSC and its governance. For instance, how to combine 40 different business models of member states on how services are provided to researchers at national level?

[2] X-Road is the backbone of e-Estonia, which allows the nation's various public and private sector e-Service databases to link up and function in harmony. https://e-estonia.com/solutions/interoperability-services/x-road/

In the questions, answers and comments part of the panel discussion the following themes were touched upon:

- It was clarified that in the EOSC governance member states and associated countries have the same position. This will be corrected in the EOSC communication and documentation and where associated countries are not at the moment mentioned in the text in the parts where member states' role is described.
- The communication strategy of the EOSC was discussed. How to inform about and engage researchers at national level with the EOSC? There could be EOSC champions in the same way that digital champions[3] were established to promote and implement the Digital Single Market at member state level. Dissemination should be on focus when the EOSC is more defined: How to tell about things at the moment, when you do not know what to answer when different questions arise? However, inclusiveness is important that all not only 'insiders' will engage to the development of EOSC.
- It was clarified that the stakeholder forum model in the governance framework is based on the RDA forum model. In the EOSC governance framework context there are many different tasks and players involved, for instance one needs to define how the services should interoperate. There is a difficulty in balancing in top-down and down-to-top approaches. Member countries and EC want to have their say as well. There are different actors on different levels: service providers communicating with researchers, EC with policy makers etc.
- It was pointed out that we should build on existing infrastructures and they should become interoperable. It is important to agree on the standards to make these existing parts interoperable. This development cannot be community driven to be sustainable and realizable. Funding comes from the governments, so they are at the driver's seat. A distinction should be made between general access to 1) the infrastructure, 2) services, and 3) access to data. There might be different access and pricing policies involved: Access to the point of use free of cost and additional costs when using different services and accessing data. With respect to the investment we should also consider and define the actors accordingly: 1) Who we expect to fund EOSC, 2) who will use the services and 3) who will pay the access to whom.

[3] https://ec.europa.eu/digital-single-market/en/content/about-digital-champions

3. SECOND PANEL DISCUSSION: PRINCIPLES OF ENGAGEMENT FOR EOSC FROM USERS' AND SERVICE PROVIDERS' POINT OF VIEW

Panel Chair: Andrew Smith, Elixir Europe

Panelists: Yannick Legré, EGI.eu; Ari Asmi University of Helsinki and ICOS ERIC; Damien Lecarpentier, EUDAT; Bob Jones, CERN; Françoise Genova, Strasbourg Astronomical Data Centre CDS

Panel topics:

- How is access to your services provided currently? For users, how do you access the services you require?
- What are the rules/principles around using those services?
- Are there legal/financial/ethical/organisational bottlenecks that would hinder your participation as a service provider in EOSC?
- Do different service types (for example, cloud, training, databases, interoperability) require different Principles of Engagement?
- How can cost-recovery be managed, especially for access to services provided in a different country to where researchers are based?

At the beginning of this session Andrew Smith introduced the EOSCPilot on-going work in the area of Principles of Engagement and some prevailing questions. In the panel discussion it was underlined that the EOSC will be built on existing credible services and infrastructures. There is a clear division of duties among the projects in this field: eInfracentral is designing a framework for EOSC services and EOSC-hub will build interoperability between eInfrastructures and Reseach Infrastructures. It was concluded that the harmonization of all access policies is challenging, but a definition of 2 to 3 access types to which the services could be categorized should be feasible. Furthermore since the EOSC is intended to be inclusive, it should be that also from the service providers' point of view. A minimum set of principles should be set for all providers, and another set of principles for those service providers, who want to provide more sophisticated, EOSC-compliant services as described in the draft EOSC governance framework[4].

Françoise Genova from Strasbourg Astronomical Data Centre shared their experience of researchers needs in her speech. The researchers want to use computing at local, national and European level. The Interface, single-stop-shop with all the interfaces that the EOSC will provide will be very attractive for the researchers. From the user point of view the research community specific services are equally important as computing, both kind of services should be part of the EOSC catalogue of services. Above all, services really have to work to be used.

Ari Asmi representing ENVRI cluster and ICOS ERIC pointed out the importance of provenance to the data. The users of their infrastructure must be ensured that the data is coming from them and not from someone else. He also brought up the challenge of connecting access policies together. The idea of providing generic services through the EOSC sounds a good idea, but from his experience there is a lot of work to make it work. Another question is that are the member states willing to invest in building these generic services and will they be worth the investment?

Bob Jones from CERN shared with the audience the experience of Helix-Nebula[5] and how principles of engagement are set in it. There are two type of services: 1) general open access services (a repository, where you can store your data, limited capacity to each user, access to LHC data, subset of data),and 2) access

[4] https://eoscpilot.eu/sites/default/files/eosc_governance_framework.pptx

[5] The Helix Nebula Initiative is a partnership between industry, space and science to establish a dynamic ecosystem, benefiting from open cloud services for the seamless integration of science into a business environment. http://www.helix-nebula.eu/

limited through federated identity management for the high-end physics community. As the policy for getting access through authorization there is a quota for different research infrastructures that is proportional to their investment. There are documented legal, financial and technical requirements for principles of engagement that are agreed across ten sponsors from different disciplines. One of the definitions is the service agreement. The barriers in the Helix-Nebula case that have been encountered are on the policy level of the network layer. It is not possible to transfer directly data to private providers. An agreement must be reached between all service providers about which kind information is shared about users.

Jones also raised the question of the public service providers in the EOSC: Will public service providers be able to sign service level agreements? Furthermore, when using the EOSC the users should be able to migrate from one service provider to another if they are not happy with the service.

Yannick Legré from EGI.eu pointed out that many public service providers are improving their service offering and some have taken an insurance for the case of not being able to meet the service level agreement. For example EGI.eu has been awarded with two ISO certifications (ISO 9001:2015 and ISO/IEC 20000-1:2011). Damien Lecarpentier from EUDAT debated if we then should take for granted that we exclude public service providers from the EOSC if they are not able to meet service level agreements. Or should exceptions be made or the principles of engagement bet set enough loose to allow more providers to join the EOSC? In the EOSCPilot service development the vision is that the EOSC must be as inclusive as possible: A minimum set of principles will be to set for all service providers, and another set of principles for those service providers who want to provide advanced services.

Asmi suggested that funders must be the drivers of the EOSC development. Otherwise infrastructures may be reluctant to develop their services towards compatibility to the EOSC and make different investment decisions of the available scarse resources. He also underlined that we need to demonstrate how the EOSC will benefit the field. Branding the EOSC for researchers and pay attention to dissemination. We should make the researchers ask for EOSC services!

Genova raised the issue of what is a service in the EOSC? There should be principles of engagement not o*nly* to IT services but also for scientific community and user group specific services. The experience has showed that from the user point of view the community and user group specific services are as important as computing services, such as PRACE. Andrew Smith pointed out to this that the governance framework draft presented earlier gave some perspectives and solutions to this: We need to have definitions of different kind of services, the EOSC compliant and EOSC compatible services, etc.

Genova continued that the complexity of the EOSC is in the diversity. There are many kind of services and many scientific communities, but not all are organized. The core is the interoperability of local, national and European level services and infrastructures. If there is a disciplinary infrastructure, the users should follow their principles of engagement and rules to ensure the FAIRness of data. If there is no such a disciplinary infrastructure or organization, the researchers / users of this field will follow generic rules of data management.

Carmela Asero from the European Commission highlighted at the end of the discussion how cross-disciplinary-usage is more complicated than intra-disciplinary. How can we scale up a service of such a community as Genova represents? It is service provider's responsibility to deal with the principles of engagement, access policies and barriers, and negotiating with funders and other parties concerned. This work should be done for the potential cross-disciplinary users to promote and widen the group of users for the infrastructure.

4. ANNEXES: WRITTEN CONTRIBUTIONS FROM THE PANELISTS

4.1. First panel discussion: Member states' role and engagement in the EOSC governance framework

4.1.1. Françoise Genova, France

Questions: What is the position of your country / organization towards the EOSC declaration that has been produced by EC as outcome of the European Open Science Cloud Summit (12 June 2017)? Does your country / organization support the declaration? Are there some particular aspects in the declaration that you would like to highlight?

- France agrees with the Declaration sections regarding data culture and FAIR data.
- The role of the RDA and of other organizations such as the W3C and FORCE11 has to be emphasized.
- The implementation level raises many issues. In particular, the EOSC should federate resources, but not integrate them in a single framework. It should build on existing resources at the disciplinary, national and local level.
- Moreover, the Declaration lacks of a clear focus on interoperability.
- Member States could agree on an interoperability mechanisms at different levels (technical, semantic, operational and legal), as described in the European Interoperability Framework.
- These interoperability mechanisms will help to share and reuse the national e-infrastructures at a pan-European level.
- The European Commission should support these interoperability mechanisms to achieve the sustainability of the EOSC.
- Open science includes open access to publications, which should be fully part of the EOSC, and this has components also at the local, national and disciplinary levels.
- The governance and funding sections of the Declaration need to be to be revised.
- For France, the governance of the EOSC should be endorsed by the Member States and Associated Countries through the ERAC SWG on Open Science and Innovation.
- EOSC governance should be inspired by the governance set in place for research infrastructures in general and in ESFRI more specifically.
- It should include formal gathering of the scientific users' needs (beyond users involved in EOSC).
- The ESFRIs should be considered as a major component of the EOSC and the ESFRI Forum should be an important stakeholder of the EOSC.
- 2. Views on how member states can or would like to engage both at governmental/policy level and national communities (researchers and infrastructure providers) in the development of the EOSC
- France considers to structure its national e-infrastructures into a French Open Science Cloud.
- France wishes to nominate a representative at the ministry level in the Board of Funders of the EOSC.
- France encourages its scientific communities to be involved in the EOSC through the ESFRIs.
- Some of the research infrastructures in the national roadmap are or include disciplinary data infrastructures and the way they will interface with EOSC has to be assessed.

Questions not covered during the panel but provided to the panelists prior to the workshop:

Question 1: How to arrange cost recovery for trans-national access when a researcher based in another member state would like to use a national e-infrastructure? Bartering? Cloud-credits? Market? EC trans-national, and virtual access funding instruments? Open access policy? …?

Answer 1:

- We do not yet know how to give a generalized access to researcher based in other member states to our national e-infrastructures.
- We feel that this question is raised too early.

- We first have to know what the user needs are that require an access to our national e-infrastructures.
- We then have to assess whether the capacities of our national e-infrastructures are dimensioned to provide a transnational access.
- We finally have to define a scientific framework and a business model to implement this access.
- PRACE is an example of e-infrastructure providing transnational access to national resources.

Question 2: Are open access policies (for instance based on scientific excellence and evaluation like in the case of some research infrastructures and PRACE) possible to be applied to national e-infrastructures?

Answer 2:

- France is preparing an open science policy and an open science action plan, where these questions could find an answer.
- Peer review based access has been successfully experienced by the GENCI national e-infrastructure since its creation in 2010 for the allocation of HPC computing resources. Therefore, such a policy is suitable for some national e-infrastructures.
- However, we also provide access to network backbone services for all higher education and research users, thanks to a direct national support to RENATER, the French NREN.
- Different national e-infrastructures require different business models and associated access policies.

4.1.2. JuanMiguel Gonzalez, LifeWatch ERIC and Spain

Due to force majeur situation Gonzalez was not able to attend the workshop, but provided a written contribution after the panel:

"As already stated during and after the European Open Science Cloud-EOSC Summit held by the European Commission in Brussels on June 12th, 2017 (see http://www.lifewatch.eu/News/881/lifewatch-eric-heartily-endorses-the-european-open-science-cloud-eosc-establishment), LifeWatch ERIC heartily endorses the European Open Science Cloud-EOSC establishment.

As you also know, ERICs (European Research Infrastructure Consortia) foundations are based on Member States agreements, and thus their role is crucial in relation to the definition of the EOSC Governance. Therefore, LifeWatch ERIC is totally aligned with ERAC Standing Working Group on Open Science and Innovation outcomes to this regard.

In any case, EOSC Governance mechanisms must not deeply interfere and thus modify the own ERICs' Governance schemes but only guaranteeing the provision of interoperability supporting tools."

4.2. Second panel discussion: Principles of engagement for EOSC from users' and service providers' point of view

4.2.1. Françoise Genova, Strasbourg Astronomical Data Centre CDS

I took the point of view of Strasbourg astronomical data centre CDS as a service provider, with reference to the astronomical interoperability framework as it is defined by the International Virtual Observatory Alliance IVOA.

1. *As a service provider, how access to services your organization is providing, is organized at the moment? As a user, how access to services the research community you are representing uses, is organized at the moment?*

Strasbourg astronomical data centre CDS provides added-value services freely accessible and widely used by the astronomical international community (~800 000 queries/day). The services are linked (through web links) with observatory archives, academic journals and the ADS bibliographic database, which is maintained by NASA. The services are accessible through their own web interfaces and the CDS Portal, they can also be queried by software and provide information on-the-fly to other services, and they are accessible through the Virtual Observatory (VO).

Astronomical data are free and public (observational data in most cases after a proprietary period which is usually short). Metadata are public with very few minor exceptions (e.g. sensitive targets such as extrasolar planet candidates).

Astronomical data is FAIR. The astronomical community has been developing a data format which includes metadata, FITS, which has been used since 1977 (first paper published in 1981), and a global disciplinary interoperability framework called the Virtual Observatory since ~2000. Standards to enable interoperability of data and tools are defined by the International Virtual Observatory alliance (IVOA, http://ivoa.net). The VO is operational and evolves in particular to include the next astronomical large facilities at best. The coordination of European Virtual Observatory activities (requirements from and uptake by data providers, in particular ESFRI and ESFRI-like infrastructures, and the scientific community, technological developments to adapt the framework to new projects) is currently supported through a Work Package of the ASTERICS Cluster, which includes astronomy and astroparticle infrastructures.

The IVOA registry of resources (Resource metadata http://www.ivoa.net/documents/RM/index.html ; registry interfaces http://www.ivoa.net/Documents/RegistryInterface/20091104/) is based on OAI-PMH with Dublin Core metadata and disciplinary extensions. It is customized by the planetology community and the Virtual Atomic and Molecular Data Centre, and more recently its principles are reused by Material Sciences to build their own registry of resources in a RDA Working Group (https://www.rd-alliance.org/groups/working-group-international-materials-resource-registries.html).

The IVOA Registry is included in EUDAT B2FIND.

2. *Which kind of Principles of Engagement the service / infrastructure you are using / providing has?*

The VO is a global, open and inclusive Virtual Research Environment. Anyone can register a service in the VO or develop a tool able to access VO-enabled data and interoperate with other VO tools. There are currently more than 100 ‘authorities’ which have registered at least one service in the VO registry.

Metadata are open and query-able from the service web interface and the VO. Some of the observatory archives require that users register before getting access to their data, but it is not a general case. CDS services are open.

3. *Which kind of legal/financial/ethical and other organizational aspects do you foresee that could become barriers for participation – either using or providing services - in the EOSC?*

The astronomical international data infrastructure is used by the international research community in its daily work. The VO enables seamless access, adapted to the community needs, to observatory archives of ground and space-based telescopes and value-added databases such as those provided by the CDS. Any loss of the capacity to fulfil the users’ needs would be an issue. Imposing a huge work overhead to data and service providers to be referenced in the EOSC would also be an obstacle. Imposing pre-registration of the users would also be an issue for many of the services.

For a community already trained to the usage of remote resources and providing many data services and tools, the key points are to fulfil user needs. Users are both the science community using data and services and data providers.

Astronomers use local, regional, national and European facilities (PRACE) for computing. Some of the sub-communities use EGI but it did not fulfil the needs of others. Interface with computing is one reason why an astronomer can be interested in EOSC, in particular to “bring computing near data”. The main interoperability issue for us is between data and computing infrastructures, and it must be transparent to the end-user.

4. *Which kind of principles of engagement the different components of the EOSC ie. cloud/compute, databases, training, interoperability, should have?*

'Gems', including resources recognized as important for communities, should be easily included in EOSC. For instance, the CDS belongs to the national research infrastructure Roadmap. 'Included' should not mean to include them in a rigid framework but recognize them as EOSC components (with the one-stop-shop vision) and providing them the possibility to interface with other EOSC components.

5. *How to arrange cost recovery for trans-national access when a researcher based in another member states would like to use a national e-infrastructure? Bartering? Cloud-credits? Market? EC trans-national, and virtual access funding instruments? Open access policy? …?*

In astronomy data is shared globally. There can be ad-hoc support to data provision. A possibly relevant information is that EC trans-national access funding is used to provide trans-national access to national telescopes.

6. *Are open access policies (for instance based on scientific excellence and evaluation like in the case of some research infrastructures and PRACE) possible to be applied to national e-infrastructures?*

In astronomy data is open. A possibly relevant information is that access based on scientific excellence and evaluation is applied to national or multi-national telescopes.