Report

GEDE EOSC Relations VM

Date: 2019/06/28

Present: Peter Wittenburg (Also member of EOSC FAIR WG), Ari Asmi (ENVRI-FAIR), Bert Meerman (GO-FAIR Foundation, Also member of the EOSC secretariat), Carlo Maria Zwölf (VAMDC), Claudia Zoani (MetroFood), Karl Presser (MetroFood), Emiliano Degli Innocenti (CNR, E-RIHS), Franco Niccolucci, Ingemar Häggström (EISCAT), Jan-Willem Boiten (EATRIS), Jacques Demotes (ECRIN), Kay Graf (ESCAPE), Luciano Milanesi (CNR, BBMRI, CNR-BIOmics - ELIXIR), Maggie Hellström (ICOS), Michael Lautenschlager (ENES-DKRZ), Rob Hooft (ELIXIR - Also member of EOSC FAIR WG), Rosie Bolton (SKA), Michel Schouppe (DG RTD), Kilian Schwarz, Suzanne Dumouchel (DARIAH), Zsuzsanna Szeredi (GEDE)

14 other GEDE experts expressed their apologies to not be able to participate due to other appointments already made long beforehand, but to be interested in the results.

In the introduction a brief overview was given about GEDE and the major points for this meeting were described: Shall GEDE become involved in EOSC and if so in which way. Carlo expressed the complexity of the EOSC landscape in small sketches and welcomed Michel Schouppe and Bert Meerman to help clarifying what EOSC is and how it is organised.

Michel Schouppe (DG RTD)

Michel made in his talk a number of important points briefly sketched here.

* EC is aware of the fact that everyone has its own view on what EOSC is.
* EOSC concerns all research communities in Europe and it’s a long term process.
* To describe EOSC it is best to go back to the EOSC Roadmap ( March 2018) reference documents.
  + EOSC is an access channel to a federation of research infrastructures. Allow access to researchers, to curate, annotate, find data and allow reuse (&access) to other’s data.
  + Findability and access will be supported by EOSC portal.
  + EOSC is also a process leading to European policy finally with the intention to make Europe a better place for researchers, contribute to an “internet of science”, transform research from how it is performed today towards open science, i.e. help changing culture.
  + EOSC is the supply side of this OS agenda.
  + EOSC needs to develop data stewardship & push for a data commons.
* Under H2020 already 250 mio € have been invested to advance EOSC. Two additional calls will be coming, i.e. about 20-25 projects were already involved in pushing EOSC. These are all research grants to deliver whatever they can in a project time.
* To achieve operational capacity much more is needed, i.e. its needed to change the modus operandi and need a process of convergence, and this needs steering. This is why the MS requested the Commission to set up a governance structure by 2018 (Council Conclusions of May 2018)
* Governance is certainly not perfect, but it had to put it in place very quickly for an operational EOSC. At the Vienna event, under Austrian presidency last autumn, the EOSC implementation phase 2018-2021 was formally inaugurated.
* EOSC should not be top-down, it really should be driven by communities, although currently it may look differently. EOSC knows that it has to be community driven to come to a sustainable EOSC. We have to built EOSC in a gradual, pragmatic way. Start small and progressively expand. Flexible design at supply and user side will evolve as well as the federated core. We will try to set the main strategic orientation for the future, there will be a review by the EC and MS at the end of 2020 about the state of the implementation achieved.
* Many open questions need to be settled until 2020:
  + What legal entity could operate an operational EOSC?
  + What is a suitable financing model?
  + What are the RoP to govern transactions between users and suppliers, service providers?
  + What is the role of the EOSC Portal?
  + How will service offerings be expanded to future users?
* All these developments need to be aligned with the Digital Single Market Agenda and needs to be seen in the international context.
* There are now two main engines in the governance structure: the Governance Board (GB) with MS and association country representatives and the Executive Board (EB) made up by different communities, having 11 seats. The design of the governance structure happened after a long consultation (2015-2018). Until 2020 we have it as it is and now we need to make it work.
* Since the Vienna event, EC had 6 -7 months to put this governance into motion. The persons appointed in the Boards had to understand their own mandate, how to distribute the tasks, develop interactions between the Boards and with the Secretariat, etc. It took 6 months, for everybody to understand their role and how to interact with others.
* Therefore it is not surprising that so far the EC has been looking inwards, community might see this as a blackbox from the outside, which is understandable, but not the aim of the EC. Now we are trying to interact with others, like the GEDE group and other stakeholders. The EB has the task to propose strategic options for the implementation of EOSC, it will then be up to GB to adopt or reject those proposals.
* Five working groups have been initiated to support the boards and they start with their activities now in June -July. Please, have a look at their tasks and check whether you might want to link with WGs and contribute to their work.
  + **Rules of Participation** led by Juan Bicarragui- to work out a minimal set of commitments. Simple rules that would allow for massive participation in EOSC. Rules have to represent policy objectives: implementation of FAIR, Open access, data regulation. Minimum barriers we want for communities to join the EOSC.
  + **Landscape** led by Jan Hrusak (ESFRI chair)- to map not only technical arrangements at the level of RIs but also to map policies driving to infrastructures, and the RoP that might already be in place. Both at national and organizational level.
  + **Architecture** led by Jean - François Abramatic - to work out technical arrangements the functionalities of EOSC. Catalogues, catalogues of catalogues, single silent systems, tools to authorize users. Tools for federated model.
  + **FAIR** led by Sarah Jones (DCC) - to propose specifications, tools to make data FAIR. Find, certify repositories. Developing trainings on FAIR data management.
  + **Sustainability** led by Rupert Lueck (EMBL) - to find out suitable business opportunities and strategic considerations after 2021.
* Please, check the webpage of EOSC Secretariat. Some experts have been appointed in these WGs, but open consultations will be made on draft outputs of the groups to allow participation of non nominated experts.
* F2F discussions are planned: one in October (RDA Plenary in Helsinki ([22nd of October](https://www.eoscsecretariat.eu/international-research-data-community-contributing-eosc#overlay-context=workspace))), an event 26-28 November in Budapest.

**Q & A:**

ECRIN participated in a number of EOSC projects (Hub, Pilot etc.) addressing the issue of sensitive data and the requirements of GDPR. Inside EOSC we feel that there is a need for addressing the special requirements of such sensitive personal data as it occurs in medical research (BBMRI, ECRIN, etc.) and in the field of human sciences. What is the relation between the Health Research and Innovation Cloud project with EOSC? How to address the issue of personal data in the medical field?

*M. Schouppe: We know about several thematic clouds (medical, energy, transport, environmental sciences, EO and space data). There are strong bottom-up initiatives to federate and structure communities. EOSC can only tap on those community efforts. There will be a strong EOSC when structured communities contribute. We need to build on what is existing such as these thematic clouds. Focussing on sensitive personal data is exactly the kind of example, where EOSC can make a difference. EOSC will not develop a specific subgroup on that topic but if the community already has a well-developed approach it should be communicated and uptaken.*

Currently we have a lot of parallel structures to discuss EOSC related aspects. We have WGs, GEDE, the clusters projects, etc. Often this work is carried out by a few people again and again. I am a little concerned that only these few people effectively bring in their expertise and wishes.

*M. S.: There are now many people concerned with EOSC and many stakeholders. Its clear we need to pay attention to the MS and the funders, because most of the money is not coming from the EU. EOSC is a european federation, European by nature. Regarding ESFRI, in the EB we aim for inclusiveness and two different ERICs are represented (long-tail of science and big data). The chair of ESFRI is also sitting at the table. So ESFRI communities should be well in place (there are only 11 seats). The cluster financed by H2020 are way to reach out to the users and bring information to them. And also GEDE could play a role.*

You see a role of GEDE. We talked with many people during the last days and one point which was constantly mentioned is the fact that there is now a decade of ESFRI experience and that it is not obvious how this experience will be analysed and made visible to the EOSC process? There needs to be a broad and neutral analysis of what has been successful and what not such as it was done in RDA Europe which led to a number of RDA Working Groups? Who is going to make this broad analysis? Should GEDE take this up? Will clusters do this work? And given that there is such an analysis, how then do the link this to the different WGs? Basic question: How to make this massive experience available to the EOSC process?

*M. S..: Differentiate ESFRI clusters and GEDE. ESFRIs look at their own services, users, how can that further converge and if it's major enough, how to make it available in EOSC. GEDE can have an overlook across different disciplines. This is interesting for EOSC, which is interdisciplinary not disciplinary. EOSC can be seen as a community of communities, catalogues of catalogues, cluster of clusters etc. GEDE can have an overlook across different disciplines. Would you agree with this perception? How to link this? Probably could contribute to the work of WGs (Architecture, Landscape etc). Submit position papers by GEDE to the WGs or to individual members in the WGs who might be already participating in GEDE. This is the type of feedback I talked about before on how to help work of WGs from outside. Are these channels sufficient? Do I understand GEDE well?*

GEDE is understood very well. Clusters are one possible way to analyse and aggregate discipline knowledge about components. I am not informed about whether clusters actually do this work? If so, GEDE would not have to do it.

*Ari: Even if clusters would do this collection of information, EOSC might need additional information that will go beyond the boundaries of the individual infrastructures.*

What is the best way to put data in EOSC or to share data usingEOSC? From a practical point of view the questions arise who the technical and administrative contacts within EOSC are? What are the paths? If data is shared via EOSC, is it automatically FAIR?

*M. S..: So far the EOSC governance is not owning any portal. The portal is developed through the EOSCHub project and they will be maintaining the facility until the end of the grant. There will be a follow up call in 2021 (DG Connect) for consolidation of the portal and to expand the prototype. There is an onboarding process for services administered by the EOSCHub project (EGI is coordinator). They already have a process and a RoP in place and it is functioning. But the portal is based on a project approach. It will be the task for the new governance to develop the RoP of EOSC. Once adopted, EOSCHub will most probably take them on board. EOSC is an access channel, providing access to repositories, but it will follow a distributed model and EOSC will only have a platform to get in touch with those repositories. In 2021, we hope to have a legal entity that could take care of the maintenance of central facilities. Until then we rely on the EOSCHub project. The RoP to come will also define FAIRness conditions to be met. Asking for FAIRness is complex and we are trying to get an international consensus in RDA (WG FAIR Maturity Model).*

I do not understand the role of GEDE. Why is the coordination of ESFRIs done by an RDA interest group. Do we have ERICs in this WG? Who is in GEDE?

*P. Wittenburg: Good question, but please postpone it till the end of this session because it is an internal GEDE question.*

**Bert Meerman (EOSC Sec Project) (see also the slides)**

* The EOSCSecretariat project has a supporting function. The projects aims to support the entire governance. Most of the interaction goes via the Executive Board. Let me make two basic comments.
  1. We are talking about a long-term project, only the taking up of the real secretariat will take three years. Most likely this will lead to a legal entity. We are now in the very early phase of a long process.
  2. Implementing FAIR is not as easy as saying FAIR. The effort of being FAIR is mostly underestimated which helps to keep everyone on board, especially researchers. But we need to stress that FAIR has to be for their benefit.

We aim to be as independent, transparent and neutral as possible for the benefit of researchers. Four actions are most important for inspiring the EOSC governance:

* Stakeholder Forum, where we take input from communities. So we clearly build something bottom -up.
* Governance Board with ministries of different MSs.
* EOSC Working Groups (five)
* Executive Board steering project by interdisciplinary force

EOSC Secretariat is a project the details can be seen in the slides.

The EOSC WGs try to get on board stakeholders who can speak on behalf of researchers. WGs are looking for additional people who help in the work. We aim to reach out to the researchers but somehow need to rely on existing initiatives bringing together researchers.

Right now, we are setting up two-way communications with various stakeholders. We also inform about the results of running projects (EOSCHub, InfraEOSC-05-B, FAIRsFAIR). Next few months quite a number of webinars and infoshare actions are planned. In this respect we are not mandating GEDE but want to motivate GEDE to get active in the realm of EOSC. Everyone can apply for funds following a strict structure (see slides).

Within GEDE there are some interesting topic discussion which could be relevant for EOSC such as for Uniform identification, EduGain service extension, DO, Semantic mapping, GDPR implication for data -exchange, Research Workflow automation use case or pilot. All this is work towards machine readability of data. It’s up to GEDE members to decide about contributions, of course. Contributing the work of WGs is of great interest and bringing in best practices.

1st of July webinar (11.00 -12.30) - Building the European Open Science Cloud: A deep dive into the EOSC Working Groups and engagement opportunities with EOSC Secretariat (<https://zoom.us/w/380924513?tk=fn69Cw00VPh4vjZy88cVTCGYsu-FE0bn7c0nZOK8K6E.DQEAAAAAFrRyYRZWcjE2VzFKQVJrV294WjhPVENrVkJ3AA&uuid=WN_9mJMRilBQhuHkBE2Uut4Ow>)

**Q & A:**

The way to participate is to write a request for funding support – correct? To make this more specific, if we would like to run an analysis about what RIs have done, we should write a request.

*B. M.: Yes. There will be different organizations applying for such grants. It will probably take 1 -2 months to the final Board approval. Was this clear on how to apply etc? Feel free to send me question in the next few months.*

Ok clear, but what are the chances of success given the effort that is necessary. Grants are probably about small money.

**GEDE Decision Finding (Peter Wittenburg)**

GEDE is an initiative started by RDA Europe, since RDA could not reach out to all RIs, difficult for people to go to plenaries, participate in the WGs, etc. So we started GEDE to interact on selected topics and to communicate this to RDA and others. In this spirit GEDE already created a broadly supported report on PIDs. GEDE is a voluntary organization[[1]](#footnote-1) and does NOT want to be in competition with clusters. If there would be too much competition, we should cease to exist. During the recent two weeks we had a lot of discussion with members on EOSC (about 22 people so far). We found that there is a huge interest in exchanging views about this topic. We are preparing the next structured version of the distributed document, which, when ready next week, will be offered on Google Drive, so that everyone can make his/her comments. A first commented version should be presented in three weeks or so and we should communicate it to everyone interested. It is a honest snapshot about how EOSC is being perceived and can motivate actors to take measures.

We are currently also preparing an analysis about the technologies used by the different RIs. We did a fist survey in 2014 from RDA Europe which was the kick-off of several RDA groups. Now we are preparing a new survey since much has changed and we do this in close collaboration with GOFAIR and ENVRI-FAIR and it is called “technology matrix”. The web-based questionnaire has already been tested with three Guiney pigs (climate, material science, languages). Now we are looking for another 10 test candidates that will fill in the questionnaire to make sure that the questions are understandable etc. At the RDA Plenary in Helsinki we will have a side meeting to report about the first results and to launch the matrix as a big action.

The other great action in GEDE is the organization of discussions on digital objects with a group of about 150 experts. There is a broad understanding that FAIR DOs can really help to solve some of the eminent interoperability problems. All information can be found on the RDA GEDE website or within short on github which we want to use as a central place for interaction.

In addition, we have

* a small group on blockchain technology and a first draft report is out.
* a small group on repositories which also produced a first draft report.
* a small group on citation harmonization

All papers are accessible via the web-site and will be made available in a more compact way soon via Github.

Now the question is whether GEDE should contribute to EOSC and it is now the moment to decide about this. And if we contribute how? If we participate, we need to adhere to some constrains:

1. We need to remain a **neutral group**, not being involved in power games.
2. We need to organize the process as a special topic/sub group in GEDE. Such work will cost time and regulations need to be followed (deliverables, KPIs, etc.), so money is needed, and responsibility needs to be defined. The suggestion is to start a small group (core group) that follows up the interactions with Bert. **So**: **do we want this?**

If we come to decisions today, we also need to include the GEDE members who are not present at today’s vico. We can do this very quickly after this report is sent around.

Ari Asmi: GEDE should do this, however, usually people who could do such work, do not have the time and funds to organize this. Providing them financial support is a good idea. We should use the GEDE platform, to discuss for example policy aspects since RIs have acquired much knowledge in this respect. A big part of EOSC is policy harmonization, some clusters can help, but there is no across dimension as GEDE can provide it.

Michael Lautenschlager: I support the policy dimension. Technical aspects are not the only one, GEDE can provide. Our expertise often comes from agreeing on common regulations. It should be a central part of what GEDE steps into.

Emiliano Degl’Innocenti: Since I am working on polic[[2]](#footnote-2)y and relations with other clusters within SSHOC, I would like to strongly support this idea. I would volunteer to be in the core group to support this type of action.

Suzanne Dumouchel: I would be happy to be part of the core group as well.

Franco Niccolucci: Some RIs are not represented in clusters, like cultural heritage. They are planning on creating a heritage cloud for example. Clusters seem to have some priority in EOSC, but there many others.

L. Milanesi: It's important to have such a group to coordinate between initiatives sincd there are also national initiatives that need to be heard.

P.W.: It should be noted that originally GEDE was started by RDA Europe, but we also have members from outside of Europe. We are an RDA Coordination group and our outputs are recognized by RDA. If a paper is finished we should put it to RDA.

Carlo Maria Zwölf: I think it’s important for GEDE to have such subgroups as mentioned in B. M. ‘s presentation on possible topics where GEDE could contribute. But what we want to reach in this subgroup? A synthetic view of existing approach or promoting new, innovative technologies?

P.W: Guess that GEDE should not exclude one of these two dimensions. In the case of DOs we already have a huge group. We need to ask this group whether they want to become active. But we could also reach out to RIs who have for example massive problems with GDPR due to sensitive data. But of course, we cannot take everything on GEDE, we need to be selective.

M. Lautenschlager: GEDE may play in role 1) in the ideas and requirements of the different RIs, like policy, rules, expectations and in role 2) where Bert for example made suggestions for the co-creation projects and implementation level topics such as PID and citation. This could be coordinated by this core group.

P.W.: GEDE indeed had a strong group working on PIDs (about 40 experts who contributed). We should summarise this paper and submit it to the WGs.

M. Lautenschlager: Yes, we should capitalise on existing work of GEDE.

Emiliano Degli Innocenti: We should create a sort of permanent collaboration between us and the people who are working on thematic clusters around issues of technical implementation, requirements etc.

Karl Presser(METROFOOD-RI): We are building the food cloud initiative within one of the EOSC initiatives – should contribute.

Carlo: what do Michel and Bert expect from GEDE?

P.W.: I think they expect to make use of existing platforms which already have shown that they can bring experts together.

Rob Hooft: I think that saying that EOSC’s content is not-well defined is too pessimistic[[3]](#footnote-3). I agree fully that bottom up experience is essential for EOSC, and in my view integrating bottom-up opinions are already is already happening. The collaboration level of science in Europe progressed from groups, to projects, via clusters, evolved into ESFRI infrastructures (who BTW each consider sustainability still as a major issue); EOSC now is there to slash down the remaining walls between data silos. EOSC is not starting a new infrastructure from an empty set, just a new higher level of organization. This can trigger, taking down the walls between clusters. For that we need a substantial set of standards and agreements. We have to make sure to reduce the number of standards, in order to increase interoperability.

EOSC organisation is trying to reach out to all the existing communities and GEDE is a very efficient way to do so. The GEDE/EOSC group should do 2 things: (1) pro-actively communicate its knowledge to the EOSC WGs and (2) follow the output of the WGs to make sure it represents the common knowledge and experience of the ESFRIs. Note: Some members of GEDE are in the EOSC WGs.

**Peter asked again whether there are participants who disagree with forming this core group. Since no one opposed, he concluded that GEDE should proceed with this idea and suggest this to the whole group. We already had some volunteers to form the core group.**

To do:

* Finalize this meeting report and send it to all for final endorsement.
* Create the new group asap to get started and ask the whole group for volunteers.

Volunteers for the core group:

* Peter Wittenburg
* Michael Lautenschlager
* Emiliano Degli Innocenti
* Suzanna Dumouchel
* Franco Nicolucci
* Zsuzsanna Szeredi (supporting the work)

1. We realise that volunteer organisations are difficult to maintain, since the RI experts are overloaded, and their visibility in tee big EOSC game is low. [↑](#footnote-ref-1)
2. There are misunderstandings of the word “coordination”. GEDE is called “coordination group” by RDA – this is a simple classification within RDA. This does not mean that GEDE is “coordinating the participating ESFRIs/ERICs. Everyone who actively participated knows that GEDE just brings together experts. This should clarify Rob’s critica l comment.about the term “coordination”. [↑](#footnote-ref-2)
3. See here the various views in th)e survey. [↑](#footnote-ref-3)