European Open Science Cloud: from vision to reality

e-IRG Workshop
Tuesday 15 May 2018

Athanasios Karalopoulos
European Commission, Directorate General Research & Innovation (DG RTD)
Unit A2 – Open Data Policy and Science Cloud
Why not tapping into the potential of research data?

- **Low** awareness of the value of **data sharing**.
- Need to translate recent changes in **privacy, data protection and copyright rules** to the research data domain.
- **Lack of data interoperability** (e.g. via common standards).
- **Fragmentation** and **lack of coordination** over different scientific communities and countries.
- **Not enough hardware capacity** for scientific computing, storage, connectivity.
The vision of the European Commission

"Europe's final transition must be one from fragmented data sets to an integrated European Open Science Cloud. **By 2020, we want all European researchers to be able to deposit, access and analyse European scientific data through a European Open Science Cloud.**"

Speech by Commissioner Carlos Moedas in Amsterdam, NL:
“Open science: share and succeed”, 4 April 2016
The vision of the European Commission

To provide to European science, industry and public authorities with:

- A world-class data infrastructure
- High speed connectivity
- More powerful High Performance Computational
EUROPEAN OPEN SCIENCE CLOUD
BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES

A trusted, open environment for sharing scientific data

Linking data

Connecting scientists globally

Long term and sustainable

Open and seamless services to analyse and reuse research data

Connecting across borders and scientific disciplines

Improving science
<table>
<thead>
<tr>
<th>Actions</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Commission will work with global policy and research partners to</td>
<td>As of 2016</td>
</tr>
<tr>
<td>foster cooperation and to create a <strong>level playing field</strong> in scientific</td>
<td></td>
</tr>
<tr>
<td>data sharing and data-driven science.</td>
<td></td>
</tr>
<tr>
<td>The Commission will use the <strong>Horizon 2020 Work Programmes</strong> to</td>
<td></td>
</tr>
<tr>
<td>provide funding to <strong>integrate</strong> and consolidate e-infrastructure</td>
<td></td>
</tr>
<tr>
<td>platforms, to <strong>federate</strong> existing research infrastructures and</td>
<td></td>
</tr>
<tr>
<td>scientific clouds and to support the <strong>development of cloud-based</strong></td>
<td></td>
</tr>
<tr>
<td>services for Open Science.</td>
<td></td>
</tr>
<tr>
<td>The Commission will make open research data the default option, while</td>
<td>As of 2017</td>
</tr>
<tr>
<td>ensuring opt-outs, for all new projects of the <strong>Horizon 2020</strong></td>
<td></td>
</tr>
<tr>
<td>programme.</td>
<td></td>
</tr>
<tr>
<td>The Commission will review the 2012 Commission Recommendation on access</td>
<td>As of 2017</td>
</tr>
<tr>
<td>to and preservation of scientific information to encourage scientific</td>
<td></td>
</tr>
<tr>
<td>data sharing and the creation of incentive schemes, rewards systems</td>
<td></td>
</tr>
<tr>
<td>and education and training programmes for researchers and businesses to</td>
<td></td>
</tr>
<tr>
<td>share data, in close relation with the DSM 'Free flow of data'</td>
<td></td>
</tr>
<tr>
<td>initiative.</td>
<td></td>
</tr>
<tr>
<td>The Commission will work with Member States to <strong>connect the priority</strong></td>
<td>As of 2017</td>
</tr>
<tr>
<td>European research infrastructures to the European Open Science Cloud.</td>
<td></td>
</tr>
<tr>
<td>Together with stakeholders and relevant global initiatives, the</td>
<td></td>
</tr>
<tr>
<td>Commission will work towards an <strong>Action Plan for scientific data</strong></td>
<td>By end 2017</td>
</tr>
<tr>
<td>interoperability, including 'meta-data', specifications and certification.</td>
<td></td>
</tr>
</tbody>
</table>
Not a cloud made in Brussels
EOSC Summit: 12 June 2017

- 110 key participants, all scientific fields, 23 countries, 1800 via web-stream
- 5 sessions
  - Data culture & data stewardship
  - FAIR Data principles
  - Research data infrastructures and services
  - Funding & governance
  - HPC, big data and super connectivity
- 4 Input papers
33 high level statements meant to capture our common understanding on the required & underpin the EOSC implementation

- Data culture & FAIR data,
- Research data services & architecture,
- Governance and funding

Signed by more than 70 scientific stakeholders (so called 'Coalition of the Doers').

An Action List is attached.
**Current situation: 6 lines of action of the EOSC model**

<table>
<thead>
<tr>
<th>a. Architecture</th>
<th>Architecture of the federated infrastructures as the solution to the current fragmentation in research data infrastructures which are insufficiently interoperable.</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. Data</td>
<td>FAIR data management and tools. A common data language to ensure data stewardship across borders/disciplines based on FAIR principles.</td>
</tr>
<tr>
<td>c. Services</td>
<td>Available services from a user perspective. A rich environment offering a wide range of services covering the needs of the users.</td>
</tr>
<tr>
<td>d. Access &amp; Interface</td>
<td>Mechanisms/interfaces for accessing EOSC. A simple way for dealing with open data obligations or accessing research data across different disciplines.</td>
</tr>
<tr>
<td>e. Rules</td>
<td>Rules of participation for different EOSC actors. An opportunity to comply with existing legal and technical frameworks and increase legal certainty &amp; trust.</td>
</tr>
<tr>
<td>f. Governance</td>
<td>Governance of the EOSC, aiming at ensuring EU leadership in data-driven science but requiring new governance frameworks.</td>
</tr>
</tbody>
</table>
Federation of research/data infrastructures

- A pan-European federation of research/data infrastructures, committed to providing services as part of the EOSC and by a compliance framework including notably the Rules of Participation, will provide access to a wide range of publicly funded services supplied at national, regional and institutional levels, and to complementary commercial services.

- The process of federation will be implemented:
  - Gradually;
  - On a voluntary basis;
  - According to a well-defined extent of involvement per data infrastructure, in terms of the data sets and services contributed to the EOSC; data infrastructures will continue to follow their own rules outside of their specific commitments to the EOSC;
  - Based on simple guidelines consistent with existing good practices.
Time plan

2018

FAIR data EG
- Recommendations
- FAIR data action plan
- Report on DMP

Study on FAIR
- Cost of not having FAIR research data
- Costs and benefits of FAIR research data & policy recommendations
- Measuring mechanism

FAIR data working group

FAIR data work plan

2019

FAIR data accreditation / certification schemes

EU framework for FAIR data

FAIR data maturity model

Work in progress
Work to be planned
EOSC services for European researchers

A researcher will find **five types of services** in the EOSC:

- **Identification** and **authentication**;
- A **protected & personalised work environment/space** (e.g. logbook, settings, compliance record and pending issues);
- **Access to relevant information** (e.g. status of EOSC, list of infrastructures, policy-related information, compliance framework) and **specific guidelines** (e.g. how to make data FAIR, certify a repository or service, procure joint services);
- Services to **find, access, store, re-use** and **analyse** (e.g. analytics, data merge/fusion, mining) the data generated by others, catalogued appropriately;
- **Services to make their own data FAIR**, to store them and ensure long-term preservation.
EOSC access & interface

- A user should have the **choice between different entry points for accessing EOSC services**, for practical reasons and to ensure a smooth transition from legacy systems. These different entry points to the EOSC should be similar but not equivalent, all consisting of:
  - a **web-based user interface / front-end**, tailored to the specific needs and context of particular user communities, including individual data infrastructures. These entry points would need to comply with a set of minimum requirements.
  - a **common platform (building on the "EOSC Hub" project and further developed in the INFRAEOSC-06-2020 call)** or back-end, offerings access to all EOSC shared resources (cf. data pillar) and to the full range of EOSC services via machine-to-machine interfaces.
- By default, EOSC services should be accessible via the **EOSC portal**, which should act as a **universal entry point for all potential users** via a full-fledged user interface, irrespective of geographic location or scientific affiliation.
Rules of participation for EOSC actors

- The rules of participation define the rights, obligations and accountability of the various EOSC actors (notably data producers, service providers, data/service users).

- There will be room and need for differentiating the rules applicable to different EOSC actors depending on their maturity and role and taking into consideration:

  - The **specificities** of different scientific disciplines
  
  - The **diversity and level of readiness** of infrastructures and services at discipline, MS and EU level (RIs, eInfras) and the differences in their established rules and processes

  - The **variety of service providers** and users that will be involved in the EOSC (e.g. public vs private; horizontal vs specialised)

  - **changing needs and practices** regarding the implementation of the rules, in particular concerning compliance with existing legal frameworks (e.g. GDPR) and emerging ones (e.g. free flow of data)
Functions: Strategy, implementation, monitoring, reporting

Staged approach in setting governance:

- **Phase 1 (<end 2020>):** steering and overseeing the initial EOSC development, primarily led by MS and EC, with stakeholders consulted and advising
  - Separation between advisory role, decision-making and implementation
  - Stakeholders (mainly) advise, propose and implement, while funders (MS/EC) (mainly) set orientations and endorse proposals,
  - **Low intervention cost, light mechanisms, high accountability**

- **Phase 2 (>2020):** (following a thorough evaluation) steering and overseeing initial EOSC operations and further development, largely stakeholder-driven, with MS/EC keeping a higher-level oversight role

- Any proposal for governance in 2nd phase would be included in FP9 proposal
Governance

Three layer structure as per Horizon 2020 WP 2018-2020

- EOSC Board of MS/AC and EC representatives to ensure effective supervision of EOSC implementation
- Executive Board of stakeholder representatives to help ensure proper EOSC implementation and accountability
- Stakeholder Forum to provide input from a wide range of actors

Further orientations for the set-up and composition of the Governance framework, to inform the consultation of MS without prejudging possible future decisions by the Council and the Commission
EOSC implementation milestones

**Q2 2018**
- Launch of a stakeholder consultation on rules by HLEG EOSC

**Q3 2018**
- FAIR data Action Plan

**Q4 2018**
- Initial EOSC Catalogue of services
- Prototype of EOSC Portal accessible
- EOSC Governance established (EOSC Board, Executive Board, Stakeholder Forum)
- First annual work plan adopted by the Governance structure

**Q1 2019**
- Initial EOSC Rules of Participation

**Q2 2019**
- European framework for FAIR research data agreed
- Initial EOSC catalogue of datasets accessible
EOSC implementation milestones

Q4 2019
- Final EOSC Rules of Participation
- FAIR Persistent Unique Identifier policy defined
- FAIR Certification Scheme available
- Initial registry of data infrastructures federated of the EOSC
- Initial EOSC federating core in place
- Updated EOSC catalogue of services & EOSC portal

Q2 2020
- Preliminary connection of most data infrastructures and services to the EOSC

Q3 2020
- Recommendations on strategic and financing orientations and organisational settings post 2020
### EOSC policy milestones

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 2017</td>
<td>EOSC Summit forming the coalition of the willing</td>
</tr>
<tr>
<td>Oct 2017</td>
<td>EOSC Declaration published for endorsements and to seek commitments</td>
</tr>
<tr>
<td>March 2018</td>
<td>EOSC Roadmap presented for consultation to Council RWP</td>
</tr>
<tr>
<td>May 2018</td>
<td>Council conclusions endorsing the EOSC Roadmap</td>
</tr>
<tr>
<td>Fall 2018</td>
<td>Establishment of the Governance structure; MS designate representatives to the EOSC board. Selection of members to the governance structure</td>
</tr>
<tr>
<td>Nov 2018</td>
<td>MS agreement on EOSC governance structure (Austrian Presidency event)</td>
</tr>
<tr>
<td>End 2020</td>
<td>MS+ EC agreement on the future strategic orientation and financing scheme for the EOSC</td>
</tr>
</tbody>
</table>
2nd EOSC Summit: 11 June 2018, Brussels

- Organised by the European Commission, together with the **EOSC High Level Expert Group & the FAIR Data Expert Group**.

- To take stock on progress and achievements towards the EOSC, a year on since the first Summit.

- To reflect on the steps ahead to ensure an effective implementation of the EOSC.

- **To launch a stakeholder consultation** on the draft ‘Rules of Participation of EOSC’ and on the draft 'FAIR Data Action Plan', two key inputs for the future EOSC governance.

- Registrations open this week.
Thank you

Website: http://ec.europa.eu/research/openscience/eosc
E-mail: EOSC-RTD@ec.europa.eu & Athanasios.Karalopoulos@ec.europa.eu