Open Science Commons: A Participatory Model for the Open Science Cloud

Tiziana Ferrari

EGI.eu Technical Director
EGI-Engage Technical Coordinator
Outline

• Open Science Commons
• E-Infrastructure Commons – The EGI example
• EOSC vision
The needs of modern research

• In research and society, ICT has become one key enabling factor for progress
  – Changing the modus operandi of research, new possibilities for geographically distributed collaboration and sharing
  – Data-driven science and more open access to data and scientific results transforming how science is made

• All large-scale Research Infrastructures and the long tail of science depend on ICT resources
  – Need to find synergies and to develop ways to tackle the ICT challenges at a generic level
  – Effective and cost efficient services that can be of wide and general use
Researchers from all disciplines have easy, integrated and open access to the advanced digital services, scientific instruments, data, knowledge and expertise they need to collaborate and achieve excellence in science, research and innovation.

They feel engaged in governing, managing and preserving these resources for everyone’s benefit, with the support of all stakeholders.

https://www.opensciencecommons.org/
http://go.egi.eu/osc
European Council conclusions, May 2015,
Institutionalised community governance of the production and/or sharing of a particular type of resource (from natural to intellectual)
Open Science Commons: Definition

• A set of interrelated resource systems governed as commons that support the open creation and dissemination of scholarly knowledge.
Applying the Commons to Open Science

- **Shared resources**
  - Integrated, easy and fair access ➔ Compute, Storage, Software and Tools

- **Engaged communities**
  - Participating in the process ➔ hybrid federation including community infrastructures
  - Culture of sharing ➔ open access to resources and services with opportunistic use
  - Collaborating in the management and stewardship ➔ thematic community-specific services

- **Governance**
  - Rules to access and participation ➔ for users and providers
  - Rules to resolve conflicts ➔ boards for User Community, Technology and Operations Management coordination
  - Rules to balance quality vs. openness ➔ SLAs and OLAs

- **Financial support**
  - For long-term availability ➔ National funding for capacity building, EC and national in-kind contribution for operating the federation fabric
1. Open in design, participation and use
2. Publicly funded & governed with the 'commons approach'
3. Research-centric with an agile co-design with researchers and research communities
4. Comprehensive in terms of universality and inclusiveness of all disciplines
5. Diverse & distributed empowering network effects
6. Interoperable with common standards for resources and services
7. Service-oriented as well as protocol-centric
8. Social connecting diverse communities
E-Infrastructure Commons

The EGI example
• A globally distributed ICT infrastructure that federates the digital capabilities, resources and expertise of national and international research communities in Europe and worldwide.

• Mission: empower researchers from all disciplines to collaborate and to carry out data- and compute-intensive science and innovation.

http://www.egi.eu/case-studies/
Horizontal and thematic infrastructures

EGI contributes to the “e-Infrastructure Commons” with generic solutions for shared needs and requirements

@ ESFRI Roadmap 2016
• **Policy-based**
  - users are granted access based on policies defined by the EGI resource providers or by EGI.eu; such policies usually apply to resources being offered “free at point of use” to meet some national or EU level objective

• **Wide access**
  - users can freely access scientific data and digital services provided by EGI resource providers

• **Market-driven**
  - users can negotiate a fee to access services either directly with EGI resource providers or indirectly with EGI.eu

• **EGI aligned with the charter for access to RIs**
• Major national e-Infrastructures: 22 NGIs
• EIROs: CERN and EMBL-EBI
• EGI Foundation
• (ERICs)
 EG I Federation, 2016 QR1

- 650,000 CPU Cores
- 500 PB of storage
- 21 Cloud providers
- 325 resource providers
- 46,000 users
- +4,000 research papers > 790 in 2015

04/05/2016 Science 2.0 Conference
A system of open e-Infrastructures

http://www.egi.eu/infrastructure/

EGI

24 European participants

6 Integrated e-Infrastructures

1 peer Open Science Grid/US
A system of open e-Infrastructures
## High Throughput data analysis

<table>
<thead>
<tr>
<th></th>
<th>Logical cores</th>
<th>Computing power (HEP_SPEC06)</th>
<th>Online storage (PB)</th>
<th>Nearline storage (PB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>February 2016</td>
<td>652,000</td>
<td>5,842,000</td>
<td>264</td>
<td>239</td>
</tr>
<tr>
<td>Increase from January 2015</td>
<td>+ 23%</td>
<td>+ 38%</td>
<td>+ 11.8%</td>
<td>+ 42%</td>
</tr>
</tbody>
</table>

**EGI Installed HTC compute capacity (logical CPU cores), 2011-2016**
EGI Federated Cloud

https://wiki.egi.eu/wiki/EGI_Federated_Cloud

21 providers
- Publicly funded, one commercial
- Offer of virtual access to GPU
- + 2 sites ready to enter production (JINR, BITP)

28 projects over the last 12 months
- Activities involving Elixir, MoBrain, LifeWatch

Training infrastructure
- 5 sites, +10 events run on the infrastructure
Users and disciplines

46,000 users
- Natural sciences, 66%
- Medical and Health sciences, 6.5%
- Engineering and Technology, 6.3%

9 RIs using the EGI federation

SLAs
- MoBrain/INSTRUCT and Structural Biology
- BILS/Bioinformatics
- DRIHM/Hydro-meteorology

7 RIs preparatory stage
NMR and structural biology

- EGI SLA: 50 Mhours CPU for the coming 2 years
- > 100 publications/year
- 3.4 Million computational tasks in 2015

Servicing the long tail of science
https://www.wenmr.eu/

Reaching 1800 registered WeNMR VRC users

Submitted by amijbonvin on Tue, 2015-05-19 13:19

We are slowly (or actually not so slowly) reaching 1800 registered users in the WeNMR VRC!

The user growth is sustained, and this more than 1 1/2 year after the official end of the EU funding.

Users and usage are thus key to sustainability.
A vision for the EOSC federation
**EOSC Research Objects Hub**

- Thematic services (data products, pipelines, software, virtual appliances..)
- Hub-specific service management processes, business processes, policies
- Research outputs
- Storage/data management
- Cloud compute
- Cloud container compute
- HTC and HPC

**EOSC Research Objects Hub**
Federation services and processes

EOSC federation services and activities (examples)

Research objects libraries

Markeplaces

Federated IdP, Auth, Authz

Research Object Indexing and discovery services

Federation services/processes (accounting, monitoring, ...
Business processes and channels

Standards and policies

Knowledge and training
• A mechanism for achieving a **minimal level of interoperation** across distributed services
• Various levels of integration possible from **loose to tight**
  – The level of integration critically affects the user experience and determines how easily services in a federated environment can be accessed
International loose federation

• Achieved today with
  – Single Sign On
  – Research object identification (PID)
  – Harmonized security and access policies
  – Community platform ensuring interoperation between different technologies to handle the complexity of different interfaces
    • Risk of community-specific ad hoc solutions
    • Some agreement achieved on data management protocols/interfaces for access and transport
    • Plethora of standards of de facto standards for distributed access to compute resources
European tight Federation

Definition of minimum set of security policies (security for collaborating e-Infrastructures initiative)
Harmonization of access policies, accounting and support channels (e.g. policy-based, excellence-driven, membership-based)

Regional/National infrastructures
- Nat. Access policies
- Revenue streams
- Customer segments
- Service Portfolio
- Service management processes

Federated service management

European infrastructures

Science 2.0 Conference
Federated service management

- FitSM standard to simplify service management implementation in a federated environment

- Service portfolio management (SPM)
- Service level management (SLM)
- Service reporting management (SRM)
- Service availability and continuity management (SACM)
- Capacity management (CAPM)
- Information security management (ISM)
- Customer relationship management (CRM)
- Supplier relationship management (SUPPM)
- Incident and service request management (ISRM)
- Problem management (PM)
- Configuration management (CONFM)
- Change management (CHM)
- Release and deployment management (RDM)
- Continual service improvement management (CSI)

http://fitsm.itemo.org/
A worldwide OSC
Service providers (publicly funded/commercial)

B2B Relationships

EOSC

Research Community services
Composite service providers
E-Infras
Data providers
RIs

B2C Relationships
Public & private research/ Educators/ Citizen science/ E-gov/ Innovators

Science 2.0 Conference
Bridge data preservation infrastructures and computing

Via a federation of service hubs, offering storage, computing, software, thematic tools, geo-replication of research data, with sustained national and European public co-funding
Open Science platforms: sharing of open tools, applications, scientific software, research data

Community platforms can be discovered, offered, supported, shared and accessed through grants
From “services” to “solutions” involving multiple providers (e-Infras, RIs, research communities, data providers, commercial service providers...)

Co-design, harmonized access policies, federated service management processes, support, training, service discoverability
Summary

• **EGI**
  – a service component of a global system of e-Infrastructures that support Open Science

• **European Open Science Cloud**
  – EOSC: a window opportunity to address the governance, funding, sustainability of national and European e-Infrastructures
Thank you for your attention.

Questions?

Acknowledgements
This presentation used icons made by Freepik from www.flaticon.com

This work by Parties of the EGI-Engage Consortium is licensed under a Creative Commons Attribution 4.0 International License.