



A European Open Science Cloud

Frédéric Hemmer

CERN

Chairman of the IT Working Group



EIROforum is a partnership of Europe's eight largest intergovernmental research organizations. In EIROforum, these organizations pursue joint initiatives, combine resources, and share best practices. The eight EIROforum members are:

- CERN - European Organization for Nuclear Research
- EUROfusion - European Consortium for the Development of Fusion Energy
- EMBL - European Molecular Biology Laboratory
- ESA - European Space Agency
- ESO - European Organisation for Astronomical Research in the Southern Hemisphere
- ESRF - European Synchrotron Radiation Facility
- European XFEL - European XFEL Free-Electron Laser Facility
- ILL - Institut Laue Langevin



Challenges and Opportunities

- Research Infrastructures need more cost-effective approaches to collecting, processing, distributing and re-using the rapidly growing amounts of data they produce.
- Existing services do not cover the full lifecycle of research from proposal submissions requesting access to Research Infrastructures, through to data acquisition, sharing and publication.
- Researchers are by-passing their in-house IT departments and publicly funded e-Infrastructures to make use of commercial cloud services.
- Procuring cloud services from commercial and public providers on a pay-per-usage model offers both flexibility and scalability.
- Many research organisations that operate research infrastructures do not have the mandate to provide IT services to their users for the management, processing and long-term preservation of their experimental data and will require assistance to bridge the gap from data to knowledge acquisition.
- A European Open Science Cloud has the potential to greatly improve the provisioning of IT services for Research Infrastructures to address the big data needs of their users.



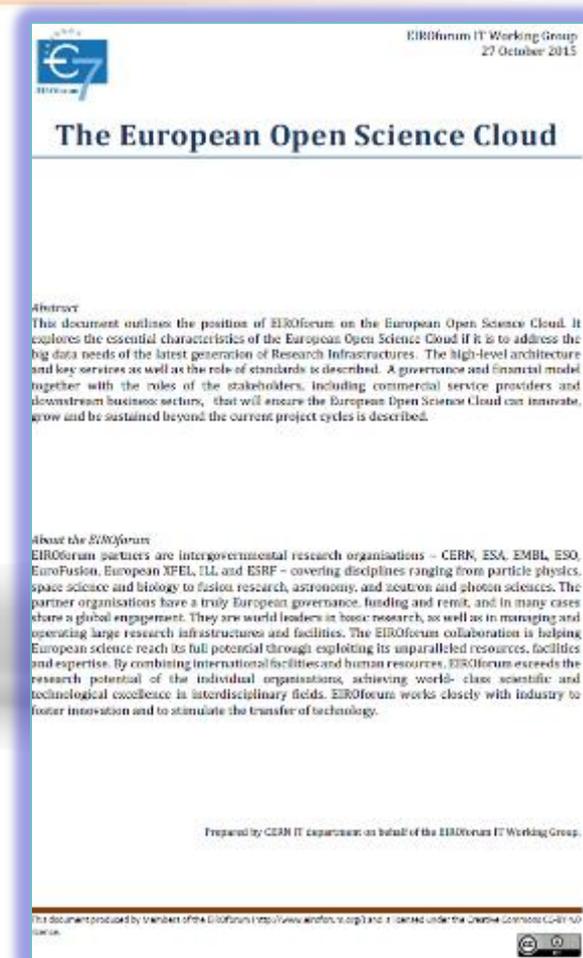
EIROforum position paper on the European Open Science Cloud

This position paper is a rallying call for adoption of a strategic approach

<http://dx.doi.org/10.5281/zenodo.34264>

November 2015, 26 pages

Endorsed by the Director Generals of all EIROforum members and accompanied by a statement of intent to enact this strategy





Stakeholders and relevant DG CNECT co-funded projects

Table 1 – major stakeholder groups

National funding agencies <ul style="list-style-type: none">• Policy makers• Third sector• Granting bodies
European Commission <ul style="list-style-type: none">• DG CONNECT• DG RTD
Research communities <ul style="list-style-type: none">• Thought leaders• Peers• Scholarly publishers
Research Infrastructures <ul style="list-style-type: none">• Policy-makers• Operational staff• Data users
Public e-infrastructures <ul style="list-style-type: none">• Service providers• Host organisations• Technology providers
Commercial cloud service providers
Independent Software Vendors
Open Source developer communities
Standards bodies

Table 2 - relevant EC co-funded projects

AARC	https://aarc-project.eu
Cloud for Europe	http://www.cloudforeurope.eu/downloads
EGI	https://wiki.egi.eu/wiki/Main_Page/
EUDAT	http://www.eudat.eu
GÉANT	http://www.geant.net/
Helix Nebula	http://www.helix-nebula.eu
Indigo Datacloud	https://www.indigo-datacloud.eu/
OpenAIRE	https://www.openaire.eu
PICSE	http://www.picse.eu/
PRACE	http://www.prace-ri.eu/
SLALOM	http://www.slalom-project.eu/



Financial considerations

- E-infrastructure costs will become an integral part of the cost of doing science and, consequently, must be cost-justified in terms of benefits and impact.
- Most publicly funded research organisations lack detailed cost models inhibiting financial comparisons between traditional and cloud-based solutions.
- Services will be provisioned from commercial suppliers when they are not available in-house or can be delivered externally on better terms (i.e. at shorter notice, lower cost or better performance etc.).
- Publicly funded data centres will continue to guarantee long-term data preservation and service supplier independence.
- **A significant difference compared to the current model is that funding agencies and research organisations will no longer provision services *exclusively* from their own in house resources**



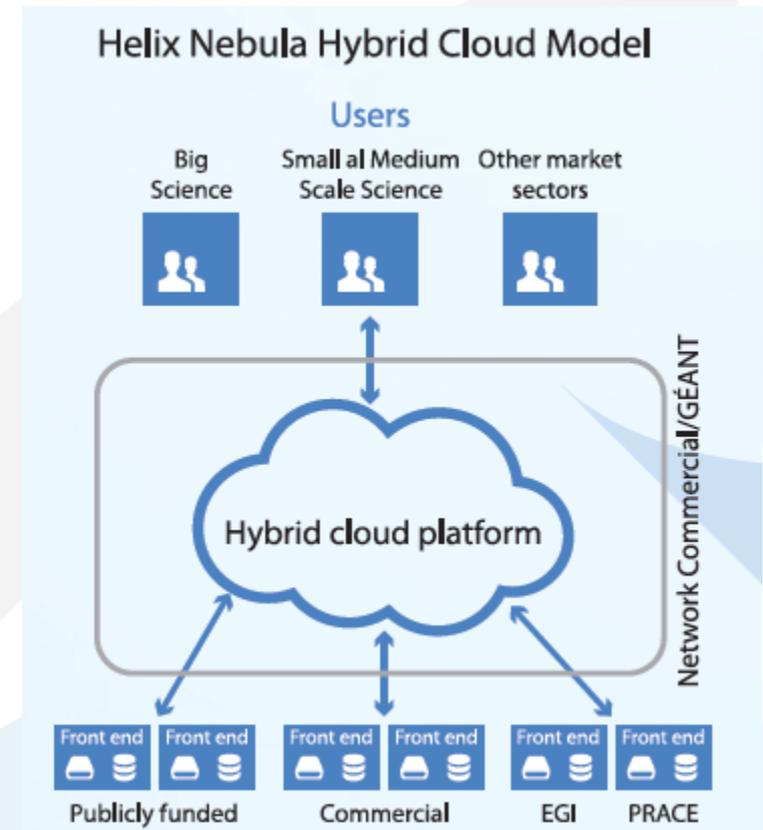
Implementation considerations

- A European Open Science Cloud should offer an initial portfolio of services based on those documented by eIRG with the technical characteristics identified by the High Level Expert Group on Scientific Data in their “Riding the Wave” report
 - Implementations for the majority of the services already exist
- The key challenges are
 - integrating frequently changing technologies
 - managing the complexity
 - identifying the optimal organisational and financial models
- Researchers must be convinced that they will not lose control of their data
 - data centres operated by public research organisations can guarantee safe copies
- Start bottom-up with IaaS
 - Integrate services via a common catalogue and a federated identity management system



The Helix Nebula Initiative

The Helix Nebula initiative has brought together research organisations, data providers, publicly funded e-infrastructures and European commercial cloud service providers to develop a hybrid cloud model with procurement and governance approaches suitable for the dynamic cloud market



The preferred model for public research organisations is a hybrid cloud that combines in-house resources with public e-infrastructures and commercial cloud services



Governance

A European Open Science Cloud will be a cornerstone of an open science commons and its governance model needs to take into account the realities of the public research sector

1. Enable integration of existing e-Infrastructures with commercial cloud computing effectively and efficiently
2. Ensure alignment with the Digital Single Market, foster coherence, equitability and inclusiveness
3. Ensure participation of all stakeholders and fair balance of their needs and interests
4. Ensure transparency, openness and responsiveness
5. Ensure value for money and fair incentives and returns
6. Continuously manage legal and ethical compliance and other risks
7. Ensure accountability and responsibility of stakeholders and decision makers
8. Manage the identity and brand of a European Open Science Cloud and ensure sustainable innovation and growth.



Summary

- Cloud computing represents a paradigm shift in the way IT resources are provisioned for research communities.
- We expect commercial cloud services to play an increasing role in the computing models of Research Infrastructures.
- A European Open Science Cloud represents a strategic vision that can be a vector for introducing change in the service provisioning and computing models for the publicly funded research sector in the medium to long term.
- EIROforum is ready and willing to take a leading role with other stakeholders and projects to prepare a European Open Science Cloud