

M U N I
I C S

Interplay between the e-infrastructures and thematic Research infrastructures

Luděk Matyska

Masaryk University

Institute of Computer Science

Czech Republic

ESFRI workshop, La Palma, 6th November 2019

Basic commonalities

- Building infrastructures
 - Complex systems with many components that must be combined and operated in a specific way
 - Large scale
- Often distributed systems
 - Even if the major facility is in one place, there is usually a large network that shares data
 - Even the supercomputing e-infrastructures are multi-site
 - Network (the GEANT e-infrastructure) an extreme example of a distributed infrastructure
- Supporting excellent science

And differences

- Thematic RIs are simply *thematic*
 - A specific (although wide) area of scientific competence
- The e-infrastructures are more “agnostic”
 - Naturally they do their own R&D in the “e-infrastructure” field
 - But the primary focus is to build and operate an infrastructure that is used by people outside the IT community
 - They provide “service” to practically any scientific field
- The thematic RIs usually include an e-infrastructure component
 - A challenge and opportunity for the interaction with e-infrastructures

The EOSC environment

- More a concept, not an infrastructure per se
- Based on the idea of the value of the (research) data
 - Access to the data produced with public funding/support should not be restricted to their producers only
- FAIR principle in the core
 - EOSC aims to be the complex “implementation” of that core
- However, data sharing is the core of thematic RIs, too
 - So what is new?

The EOSC environment – what's new

- Make data preservation an indispensable part of any (publicly funded) research
 - Not only through the RIs (top down) but starting with individual (not organized) researcher (bottom up)
- Don't create silos around themes/content
- “Data” in a very extensive interpretation
 - Practically any digital object, i.e. also tools, workflows, ...
- And provide an implementation that “link them all”

The EOSC environment – the implementation

- Still under development
 - Very ambitious and complex problem
- Many working groups
 - Architecture, landscape, onboarding, ..., but also governance
 - But still unclear what is to be build (on top of the concept)
- Not clear role for e-infrastructures
 - EOSC is a concept (and mind set), not an (e-)infrastructure

The EOSC environment – Current focus

- The EOSC Portal
 - Primary entry point to all the “EOSC data”
- The EOSC Marketplace
 - Collection of tools, databases (data sources), ..., sharing some common features
 - FAIR data, interoperable tools, ...
- The risks:
 - Fragmentation, quantitative measures (number of items in the Marketplace), ...
 - Unclear role of both e-infrastructures and thematic RIs

The EOSC environment – Opportunities

- Collaborate (RIs and e-infrastructures)
 - Don't succumb to the idea that e-infrastructures are equal to “vendors” or even “enemies”
 - Just individual tools/datasets provisioning is not sufficient
 - Someone must take care of the selection and integration and operation
- Actively engage in the EOSC shaping
 - Don't throw away the EOSC concept, but shape it so it actually helps you
 - Take e-infrastructures as partners with their own specific expertise
 - Esp. in the integration of selected tools and pieces and in operating their complex interplay
- Don't forget resources (true capacity building)

Conclusion

- EOSC is a challenge both for e-infrastructures and thematic RIs
- There is a high risk of misunderstanding, leading to further real fragmentation of the scene
 - Just formally fulfilling FAIR principles and registering tools and dataset in EOSC Marketplace/Portal is far from sufficient
- Come forward with needs and ideas and shape the future EOSC
 - Do it in collaboration with e-infrastructures
 - Don't forget there must be real resources behind

**Thank you for the attention
Any Questions?**

The interplay

- Depends on the kind of e-infrastructure
 - Networks and supercomputers driven by the IT concepts
 - Distributed e-infrastructures (e.g. EGI) practically conceived within the thematic RI (HEP)
- Many facets
 - From “service” provisioning to close co-development