Virtual Research Environments as integral to e-Infrastructure Commons

Dr Dimitris Koureas
Lead, Research data & Partnerships
Natural History Museum London

@DimitrisKoureas

Coordinator, Distributed System of Scientific Collections (DiSSCo) initiative
Executive Secretary, Biodiversity Information Standards (TDWG)
Co-chair, RDA BDI IG & ASDOC WG
Strong political commitment
Digital Agenda
European Open Science Cloud

How does success look like?

1. Improvement of technological readiness levels
2. Number of engaged users
3. Change in the scientific mode of operation
4. Faster and more robust scientific innovation

Over the last years investments have focused towards robust, reliable and interoperable services that generate global solutions for data sharing and preservation, high performance and cloud computing, user-authorisation and authentication

A strong investment programme
Previous investments in e-infrastructures
2007-2013 (FP7): €572m
2014-2015 (H2020): €170m
Researchers still reluctant to engage

---

Main barriers for researchers to engage with ‘open science’ practices (>80% agree)
Science Europe Survey 2015

(i) Quality assurance
(ii) Attribution
(iii) Integration between infrastructure components
(iv) Limited awareness

Users percentage per domain (2015)

- Physical Sciences: 47.2%
- Biological Sciences: 43.5%
- Earth Sciences: 3.6%
- Humanities: 1.4%
- Other: 4.3%

Not all communities are created equal

Striking discrepancies in uptake rates → Different communities require different approach
Socio-cultural (less technical) distance

- Physics, astronomy, bioinformatics
- Humanities, Social studies

Backbone service
- Bus
- EOSC

Generic interfaces

Community engaging mechanisms
The ‘last mile’ challenge for research e-infrastructures

- Borrowed from telecommunications and transportation
- Cost and complexity of ‘connecting’ end-users to the backbone (core) infrastructures is high when compared to the core infrastructure itself.

Virtual Research Environments act more as the socio-cultural and less as technical brokers.

Intermediate environments that connect communities of practice to a common e-infrastructure landscape (e-Infrastructure Commons).

Community vetted mechanisms (quality assurance) to organically develop new practices.
Virtual Research Environments?
aka
Virtual Labs
Science Gateways

- Envisaged as community supporting mechanisms to **enable digital collaboration**
- Predominantly **domain-specific solutions** that played a key role in bridging the gap between traditional and digital research practices
- Received significant investments between **2005-2015** (e.g. EC, national funds)
Facilitate the research data lifecycle

VREs enabled communities to transfer their typical scientific practice into the digital realm.

Functional and operational characteristics differ significantly according to the communities of practice addressed.
VREs enable researchers to organically change their *modus operandi*.
An example from the Biodiversity domain

Scratchpads
biodiversity online

Per month unique visitors to Scratchpads sites

70,000 unique visitors/month

In total more than

5,000,000 visitors

650 Scratchpads Communities
by 7,500 active registered users
covering 125,000 taxa
in 1,200,000 pages.
The challenges with VREs

- Built in isolation
- Developed redundant technological solutions for storage, authentication, computing
- Failed to develop good sustainability models
- (failed to attract users)
Re-inventing Virtual Research Environments as integral to the e-Infrastructure Commons

E-Infrastructure Commons considers community engaging mechanisms as integral to its operation.
The development of EOSC does not reduce the importance of VREs.
In fact, it strengthens it. As they become the needed mechanisms that enable communities of practice to engage with core services,

The EOSC reduces the operational costs of VRE up-keeping and hardwires interoperability minimising fragmentation and redundancy.
Recommendations
for integrating community-specific services
to the European e-infrastructure Commons

User communities need to be able to:

(i) articulate and communicate their community-specific needs in regards to data and services, and
(ii) translate these needs into clear functional requirements that will drive the development of VREs.
Recommendations
for integrating community-specific services
to the European e-infrastructure Commons

VRE operators need to:

(i) VREs looking beyond the ephemeral timeframes of project-based approaches,
(ii) Build public-public and public-private partnerships that ensure sustainability and,
(iii) Link VREs with existing underlying e-infrastructure, building on top of available
backbone services.
Recommendations
for integrating community-specific services
to the European e-infrastructure Commons

Funders need to:

(i) further acknowledge the pivotal role of VREs in support of user community engagement and,
(ii) develop, with a particular eye to long-term sustainability, dedicated VRE funding programmes with targeted calls to discipline-specific communities.
Thank you

@DimitrisKoureas
d.koureas@nhm.ac.uk
Despite their success, for Virtual Research Environments to continue serving their communities, we need to reinvent their role within the ever-developing research e-infrastructures landscape.

Successful examples of VREs across disciplines

Past Projects (EC-funded)

Current Projects (EC-funded)

VI-SEEM
MuG
OpenDreamKit
BlueBRIDGE
VRE4EIC
West-Life
Developing the European Open Science Cloud

A set of robust, sustainable infrastructures that provide cloud services (authorisation, storage, computing)

An ecosystem of common backbone services
Researchers are still reluctant to engage

European Grid Infrastructure (a flagship initiative that delivers integrated computing services to European researchers) announced (Dec 2015) a user base of 35,959 (European Grid Infrastructure 2015)

![Pie chart showing users percentage per domain]

- Physical Sciences: 47.2%
- Biological Sciences: 43.5%
- Earth Sciences: 3.6%
- Humanities: 1.4%
- Other: 4.3%
Not all research communities are ‘created equal’

For instance, researchers working on genomics, physics or astronomy have long appreciated the value of data sharing. By nurturing a culture of shared physical and computational infrastructure, open-source software and open data, they have embraced the principles of open science.

The need to respect diversity and continuously developing needs

Research communities are developing new practices organically and that they are best placed to explore which of these contribute to the advancement of their discipline.
A global priority of changing the *modus operandi* of science

- Investments in development of e-infrastructures
- Provision of tools and services
- Training and incentives

Digital Agenda for Europe

Digital Collaborative Data-intensive research
When Bell invented the telephone...

“The Americans have need of the telephone, but we do not. We have plenty of messenger boys.” – Sir William Preece, Chief Engineer, British Post Office, 1878

“This telephone has too many shortcomings to be considered as a means of communication. The device is of inherently no value to us.” – Western Union internal memo, 1876