

KE meeting on Research Tools Wednesday 27 June 2012

Date and time: Wednesday 27 June 2012 (9:30 – 14:30 CEST)

Venue: SURF office, Graadt van Roggenweg 340, Utrecht

Participants: Magchiel Bijsterbosch, SURF; Bas Cordewener, SURF/KE; Marc Dupuis, SURF/KE; Gera Pronk, SURFNET; Torsten Reimer, JISC; Louisa Dale, JISC/KE; Angela Holzer, DFG/KE; Jens Klump, GFZ Potsdam; Johanna Vompras, Bielefeld University; Mark Allen, CDS; Keith Russell, Knowledge Exchange (minutes)

Apologies: Nicole Gregoire, SURFnet; Matthew Dovey, JISC; Anne Sandfaer, DEFF

Knowledge Exchange meeting exploring research tools

Background

In recent years, Knowledge Exchange partners have made significant investment into virtual research environments (VREs), a method of collaboration by researchers that is enhanced through the systematic use of information and networking technology.

Partner understanding of the application and use of virtual research environments (VREs) matures, with recognition of the challenges developing and sustaining environments. A further challenge is to balance the interests of working on a general re-usable solution yet also taking into account the needs of specific research disciplines. The focus for partner dialogue and investment has shifted away from well defined or packaged 'solutions', to a wider consideration of the use of digital technologies in supporting the researcher in the research process.

Discussions by the Knowledge Exchange working group and at a recent Knowledge Exchange workshop on virtual research environments suggest the following interventions to further partner investments:

- **Understanding researcher behaviours:** encourage studies into research practice and ways to 'reach' young researchers.
- **Informing policy / exploring impact:** ensuring sustained and effective investment in technologies to support research.
- **Supporting community networks:** for technical developers.
- **Developing success stories:** raising awareness of the value and potential of digital technologies in research.

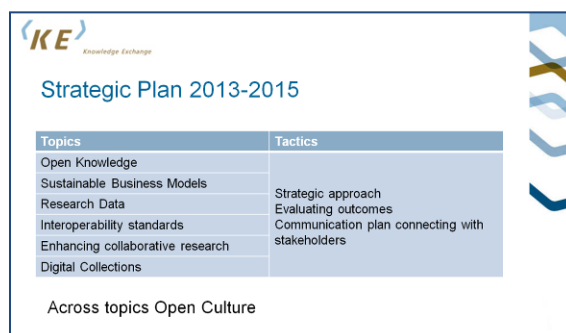
Meeting to inform KE activities

Knowledge Exchange invited colleagues from partner organisations and related organisations to a meeting to consider 'the use of digital technologies in research' or 'digital research tools' with a view to informing future Knowledge Exchange activities.

At this meeting, representatives from Knowledge Exchange partners and related organisations were invited to make short presentations on their current investments in digital infrastructures and specifically any in the use of research tools.

Report on the meeting

Bas Cordewener from SURF welcomed all participants to the meeting and Louisa Dale explained the aim of the meeting. Keith Russell provided some background information on Knowledge Exchange (KE) and Louisa Dale explained about the work on a strategic plan for



Strategic Plan 2013-2015

Topics	Tactics
Open Knowledge	Strategic approach Evaluating outcomes Communication plan connecting with stakeholders
Sustainable Business Models	
Research Data	
Interoperability standards	
Enhancing collaborative research	
Digital Collections	

Across topics Open Culture

KE. In this plan there are five topics which KE aims to address. Several of these topics are relevant to the work on research tools.

She also explained about the work undertaken on virtual research environments (VREs). [A workshop in 2010](#) investigated the concept of VREs, compared projects and discussed the need to sustain these environments. A [roundtable](#) was organised investigating challenges. Outcomes were that a better understanding of researchers behaviour is required and there is a need for policies. During a [further workshop](#) in November 2011 several findings were collected: VREs are quite varied in scale and scope and are hard to class under one monolithic concept, a flexibility of use is required. It was noted that young researchers are not getting to grips with potential of new technologies¹.

The latest work has been on preparing a [VRE knowledge base](#), this is a directory that lists projects across partners, relevant articles and background information. Based on the lesson learned the European Commission has invited KE to come to Brussels to present our findings during a lunch session. Valuable messages on research tools are also very welcome.

Several participants presented the work taking place in their country on research tools.

Mark Allen presented the [Virtual Observatory](#), this is set up by an international alliance with 19 members from various countries across the world. The alliance is an international body developing standards and sharing best practices. In the virtual observatory data is combined data from various types of telescopes from optical to long wave telescopes. It contains a registry of data with metadata descriptions and discovery tools to help identify the relevant data. Virtual Observatory standards are adopted by data centres such as the [CDS](#), providing standardized access to its services (an astronomical object identifier service, a reference service, and a data portal). The choice has been made to develop tools that concentrate on one task but can exchange and connect. Some of these are desktop tools that can connect to the registry. When setting priorities the scientific priorities define technology standards.

The alliance also works on engaging the scientific community as participants. These are typically early career researchers with a varied background.

Value has been apparent of making results visible also to a broader audience outside of science. Challenges lie in sustainability, meta data quality, take up by large data centres (connecting with others is sometimes a second priority) and providing science ready data.



Angela Holzer presented the national landscape in Germany. There are 16 governments responsible for research and education. The alliance of science organisations started a [priority initiative on digital information](#) in 2008 which is working on a national information infrastructure best suited to meeting their researchers' needs. In 2011, a master plan for the [information infrastructure](#) in Germany was published which focuses on the developments in eight domains in order to move toward a more coordinated national landscape. One of the domains sketched out in the master plan is the area of VREs: they are considered to be still in an early stage of development. Current challenges are identified: Of crucial importance is the consideration of how existing solutions can be applied in other or across disciplines, how close collaboration between information specialists and researchers can be improved, how sustainability or re-use of reference architectures can be achieved, and finally how the needs of the researchers can be addressed most adequately.

DFG has a standing programme for the funding of VRE projects and has in addition published three calls with specific and different focuses for the development of VREs. The funded projects pertain to a wide variety of disciplines and address a large number of topics. They have to be technically reusable, may offer best practices and also cover sociological studies that look at transformations of collaboration in research. Tool development is linked

¹ The report Researchers for tomorrow has just been released and is available at: <http://www.jisc.ac.uk/news/stories/2012/06/generationy.aspx>


strongly to disciplines and an open question is the possibility of a disciplinary use of tools in overarching interdisciplinary reference architectures. The focus of the current call are the questions of modularity and of sustainability. She referred to a [GRDI paper on VREs](#), the technical developing of tools does not seem to be the problem, but the societal and organizational processes are, especially to ensure take up with researcher communities and the cooperation between research and information infrastructure institutions and data centres. For funders as the DFG, the challenge is to realize that tools can be used in different context and do not necessarily relate to VREs. Therefore, tools are developed in a number of funding schemes and there is not enough knowledge in the communities about these developments.

A question arose on where to position tools, does this include tools for communication, dissemination, creation, retrieval and publication. There are tools for archiving, though this may be less relevant for this discussion. Challenges lie in the reuse of tools across disciplines, identifying user needs, evaluation of success and best practices (also the reason for lack of take up and the failure of tools), measuring the impact of tools, gap analysis on tools and investigating whether they are cost effective.

KE could take a role with regard to best practice collections, analyses of user needs and advocacy of open source requirements for tools.

3. From our point of view, what could the KE Working Group contribute ?

- ❖ Analysis of changing researchers' behaviour with regard to usage of digital tools (discipline-specific; which parts of the research process benefit from the use of tools?)
- ❖ Communities of practice and interdisciplinary use (do research tools change the way research is organised, how do they influence research topics?)
- ❖ How do we evaluate success/usefulness of digital tools?
- ❖ Financial and organisational sustainability
- ❖ Informing policy: Open Source tools, commercial tools



KE meeting on research tools
18.08.2012

DFG

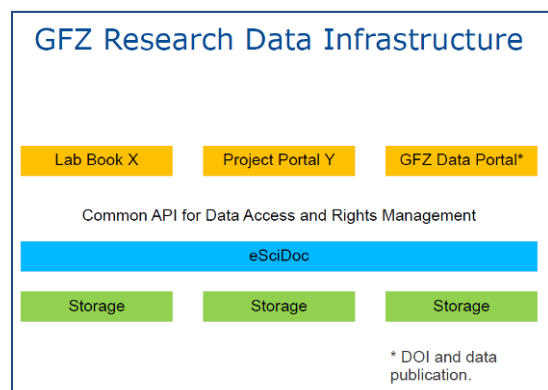
Johanna Vompras presented the DFG-Collaborative Research Centre (SFB) developed at [Bielefeld university library](#). This centre has a

information infrastructure component which is a data and information infrastructure for social science supporting and integrating information resources for all disciplines participating in the Collaborative Research Centre, focusing on research on social inequalities. Aim is in one system to have a central collection of project outputs, shorten communication lines, upgrade the research work, work on data re-use.

Aspects specific to social sciences are data privacy, documentation (throughout life cycle) using DDI as a metadata standard, research into the methodology used. Various types of data (quantitative, qualitative, simulation) which require different approaches. Working on linking publications and datasets in local systems. Actual use will be dependent on wishes of researcher, these will be investigated.



Jens Klump presented the work on data infrastructure projects at [GFZ](#). 150 concurrent projects are taking place at the centre, to support this, there was a need for scalable processes and generic tools for the whole life cycle. One of the challenges is that adding metadata is an unloved task. This requires flexible metadata schemas and where possible to make use of automated input, for example from the machine that collected the data. Data entry is only needed during active project phase, and 'Haute Couture' tools are designed to provide this. However a more 'Pret-a-Porter'



solution would suffice after the project has closed.

They have now designed a common API for data access and rights management layer on eSciDoc which collects data and passes this on to storage.

Data review is also organised in the system and the entering of metadata. In the lifecycle it supports the moving of data from private to group to persistent and access domain.

Torsten Reimer presented the work of JISC on research infrastructure and VREs. These domains are now merged in the research programme. Working on exploiting e-infrastructure, building communities across domains.

There are various reasons why VREs have been renamed to research tools. VREs are expected to do everything, which make them complex to repurpose. This has not happened as much as hoped. Research tools on the other hand are more generic, can be either large or small.

In the current call there is a demand that institutions demonstrate that tools are actually used in research. He mentioned four examples of projects: Cambridge E-

lab notebooks, TEXTUS collaborative text editing, Batmobile using mobile phones and cheap microphones to find out where bats live. Twitter Analysis Workbench Development investigated the role of twitter in the London riots.

Matthew Dovey provided a paper explaining why JISC has moved away from the VRE term. It was often regarded as one platform or portal which would solve all demands, excluding tools outside this environment. Research tools is a more generic term.

JISC
inspiring innovation

Why "Research Tools"?

1. Bring together support for research under one programme
2. Address overlap between previous programmes
3. VRE often seen as too monolithic
4. Research tools more generic, can be big or small

Marc Dupuis outlined the activities of SURF. Special Interest Groups are being set up, informal groups to discuss topics, one of these will be on research tools.

They are also working on an 'app store' collecting and presenting different tools. This was a shift from the original aim to set up a portal. Institutions indicated that they did not want such a predefined portal. The app store will be explored this year, implementation plan in next year.

SURF has an interest in sharing tools across boundaries.

Gera Pronk presented the work on SURFconext, a middleware solution linking tools. This is now being piloted at several institutions. They are also working on a project for sending very large files and supporting virtual organisations in different countries using [EduGain](#). She provided a draft e/IRG paper on data management with a request to provide comments.

International opportunities

- Still looking for **services** to connect to SURFconext
- **EduGain** use cases
- International **network** opportunities (100 G bandwidth)
- Comment on e-IRG Blue paper Data Management
- EYR4 (international e-infra contest for researchers)
- Vivo? SCLweb? Looking into it now

• Or please contact me for other ideas!

The presentations were followed by a discussion on common challenges and approaches. It is recognized that we cannot afford 'haute couture' solutions for everything and we need flexible solutions to architecture. There is a broad variety in how far disciplines are in developing such an architecture.

The challenge is how to create a flexible structure. We will have to identify which components are generic and which are not. When you have centralised services you still need local support. Data storage is an example of a centralized service that would scale really well.

An inventory of tools would be useful to get insight in which tools have already been developed and this would support re-use. The concept of an 'app store' would be relevant for more than one country. There was a discussion whether this app store should be at the end-user level or at a lower level. [Project Bamboo/DiRT](#) is working on such a store of tools for humanities, it is worth aligning with other initiatives.

It would also be valuable to have a view what has succeeded and what has failed in the past. Sometimes old code can be further developed. The discoverability of tools could be improved. There is an interest in identifiers of software versions for reproducibility. This would also allow for citations, which would be a reward for researchers for their software development.

Regarding the role of Europe: in their presentations they show VREs as a fluffy cloud. They do that have a clear demarcation of the concept.

There is also a cultural aspect to this. There is a challenge in how to include researchers and staff and general public in developing tools.

The afternoon discussion picked up on the earlier discussion from two perspectives

What is value for KE partners?

- Outcomes from specific projects could be carried to a more generic level. Share what worked.
- Alignment has to be created (Though this does perhaps not actually have to be undertaken by KE). It would be valuable to join the dots through influencing and exchanging knowledge
- Identify benefit to researchers
- Alignment between partners (though alignment of funding calls was considered less probable)
- Shared reference terms that projects should use. This would allow the transfer of outcomes of one project to the next, for example on standards to be used. This could be a recommendation to funders
- Sharing surveys planned, collecting what is already known. (For example also on early career researchers). Lessons learned regarding events and training are not only interesting for researchers but also for supporters.
- Generic is not necessarily central. A generic service can well be local (this is appreciated by researchers).

What are the key challenges we think EC might address?

- Approach researchers, this does not work by proxy. There are no incentives for researchers to share and re-use tools.
- Role for research technology support.
The example from GFZ: they are looking into required capacities (consulting role and support role, facilitators) and e-learning to train staff.
- Compare approaches already taken on serving the researcher.

A catalogue of problems, services and solutions would be helpful. This would be nice for reference for funders. This would also be of interest to researchers and research performing organisations and infrastructure providers in institutions. We could conduct a survey to inform at what level their interest would be.

Would it be worth capturing projects and success stories in a publication? This would allow the transfer of ideas from one field to the other.

Ambition (where are we in 5 years time)

- It has to run on an iPad (it has to be an app)
- Easy access and easy use
- Safe
- Free for the researchers to use
- Green
- Data access problem is solved (what about the cost problem?)
- What about certainty of maintenance of app? Some will be sustained by community, others will not.
- Commercial partners are offering services on the network. Services may be provided by a commercial provider. Quite a lot of it will be in the cloud.
- Researchers being aware of the value of technology, having an informed demand.
- VREs as they are now and will be integrated into other systems/platforms

Next steps

Louisa Dale thanked all the participants for their views and input. This is a great starting point for work in KE on research tools. It was valuable to hear that there was a clear shared perspective on the value of this work.

This is valuable input to help inform the commission on VREs and where the next steps might lie.