Virtual Research Environment

Collaborative Landscape Study

A JISC funded project
(January 2010)

Dr Annamaria Carusi
(Oxford e-Research Centre, University of Oxford)
Dr Torsten Reimer
(Centre for e-Research, King's College London)

Contents

1 Introduction

1.1 Foreword

This study was undertaken in response to an initiative of JISC, the UK's Joint Information Systems Committee. JISC has taken an active role in supporting VRE development in the UK through its VRE Programme that has just gone into its third phase[1], and had commissioned four studies to look into different aspects of VREs. One of the studies was meant to look at VREs and research collaboration in a wider international context – the VRE Collaborative Landscape Study.

The project ran from July to December 2009, and was a collaboration between the Centre for e-Research at King's College London and the University of Oxford e-Research Centre.

The major parts of the study and report were undertaken by Torsten Reimer and Annamaria Carusi; Craig Bellamy co-authored the original bid and was involved at initial stages of the project, with Valentina Asciutti joining after he had left CeRch. Other project team members included Anne Trefethen and Marina Jirotka at the OeRC.

1.2 Acknowledgements

First of all we would like to thank JISC and especially Frederique Van Till, the VRE programme manager, for initiating, funding and supporting this study. We are grateful for the chance to look into the international VRE landscape and we would also like to thank Frederique for helpful suggestions and her support.

This study could not have been written without the participation of people from all over the world who responded to our survey invitation and took part in the follow-up interviews; in particular, we want to thank all the interviewees (please see the appendix for a list of all interviewees) who took the time to speak to us in detail about their perspectives on VREs and collaboration. Their input was invaluable to this study.

Our project is one of several studies on VREs funded by the JISC at the same time. We would like to thank the other researchers taking part in these studies for the cooperation and willingness to share findings and discuss issues and ideas. In particular, we would like to thank Jim Farmer and Paul Miller.

We would also like to thank colleagues at Oxford e-Research and the Centre for e-Research at King's College London for their support and suggestions, particularly Sheila Anderson who gave valuable input into the original bid that we submitted to JISC. Particular thanks go to Ian Stephenson, Department of Education & Professional Studies at KCL, who provided invaluable input into the design of the online VRE survey. David Robey at OeRC acted as an external advisor and provided useful input on the arts and humanities side of VREs.

Last, but not least, we would also want to thank our four external reviewers for giving feedback on the draft report and also for providing very helpful input and information as our project moved along: David de Roure, John Doove, Nancy Wilkins-Diehr and especially Ann Borda, whose ongoing input over the course of the project we particularly appreciated; Ann also wrote a country perspective on Australia for this report. John Doove was kind enough to share a draft of the Collaboratory Study commissioned jointly by SURFfoundation and SURFnet; this study is looking into environments that can be used to build a VRE and makes recommendations; it should be of interest to readers of our report.

1 http://www.jisc.ac.uk/whatwedo/programmes/vre.aspx [Accessed 18/12/2009].

1.3 Executive summary

The VRE Landscape Study aimed to investigate international developments in Virtual Research Communities (VRCs) and to evaluate them in relation to the activities in the JISC's VRE programme. The study examined programmes in a number of key countries along with significant projects and communities as well as some countries where developments on this front are just beginning. There has been a great deal of activity over the past few years in terms of prototype and demonstration systems moving into the mainstream of research practice. Notable trends are emerging as researchers increasingly apply collaborative systems to everyday research tasks.

1.3.1 Key recommendations

1. Fundamentally, the most important point to have emerged from our study is that VREs need to be conceptualised as **community building projects rather than technology projects**.
2. VREs have the potential to benefit research in all disciplines at all stages of research. The access to data, tools, computational resources and collaborators that VREs facilitate leads to **faster research results** and **novel research directions**.
3. By far the most important challenge faced by VREs is **sustainability**. VREs have to be supported and used by research communities in order to be viable, but without assurances that VREs will continue to be sustained in the medium and long term, it is far more difficult to get the commitment of time and effort by researchers. VREs need to be seen as **vital elements of the general research infrastructure** requiring the same long-term commitment as other parts of the infrastructure.
4. VREs have an **international dimension**, especially because of their promise to integrate resources from different origins. There is a need to gain clarity on legal, ethical and other policies and frameworks that govern the sharing of data and other resources, and to communicate these clearly to researchers and developers. When policies and laws constitute needless barriers to research, it is necessary to attempt to modify them.
5. Both the further development and sustainability of VREs require **international co-operation**. The establishment of an international VRE forum to coordinate activities could be a way forward, and funding councils could play an extremely important role in this.
6. A multi-pronged approach to sustainability needs to be implemented, from the planning stage and throughout the life of a VRE:
    - In order to ensure **usability**, VREs need to be focused on the needs of researchers and specific research communities, putting them in the driving seat of VRE development. This implies a user-driven, bottom-up mode of development. There is no one-size-fits-all mode of developing VREs. Instead VRE users need to be put into a position where they can create their own environments with tools and other resources that are relevant to their research. In terms of technologies, the more lightweight and the more customisable the better.
    - In order to ensure broader **uptake**, VRE initiators and developers need to plan in advance for engaging the broader research community which will sustain the VRE in the medium- and long-term. It must be ensured that people in the necessary roles are involved in the VRE, including managers, librarians or archivists, champions or promoters, and the appropriate mix of established and early researchers. Awareness raising, targeted training of different types (face to face as well as web-based) and other engagement events are crucial for the uptake of VREs.

- The broader institutional and broader **context** of VREs must be taken into account at the planning stage, including the priorities of specific research institutions, the relationship with national and international institutions, the relationship with libraries and digital repositories, the relationship between research and publishing.
- A **business plan** taking into account this context and the research community supported by a VRE needs to be considered from the outset.

1.3.2 Key findings

The creation of a VRE is a social as much as a technical achievement. This is by far the most important point to emerge from the survey. It is not an entirely new point, as it has been highlighted in previous literature on VREs. However, it is a point which was strongly consolidated by the study. Any forward-looking strategy on VREs needs to approach them in a holistic fashion as community projects.

- Potentially the most important trend identified by this study is an increasing focus on providing **general VRE frameworks** that can be used to develop and host different VREs. The frameworks would provide core services (such as authentication and rights management; repositories; project planning, collaboration and communication tools) and allow the development or easy integration of modules for specific uses.
- It is clear that the most effective way of approaching the development process of VREs is **a participatory mode of development**, with researchers closely involved in generating the requirements and evaluating their implementation. Development needs to occur in an iterative fashion, with constant feedback form researchers. There is also a need to support researchers through training opportunities.
- The development of a VRE needs to be broached not as a technological project but as **a community building project**, since without community buy-in, the VRE cannot fulfil its function. Community outreach beyond the initial community is also essential for the future sustainability of VRE projects, with members feeding new applications, new content, and the social context that allows for effective use of data and other resources, back into the VRE.
- In terms of technology, the strongest trend seems to be in the **preference for a Web 2.0 style of development** and implementation. In general, more lightweight, customisable solutions are preferred.
- **Sustainability is a key issue and a major concern** for many of the projects that we studied. It is important that national and international strategies, funding councils and institutions work together to address sustainability, since by their very nature, VREs cross institutional and national borders and only a concerted strategy will be successful. At the same time, it is important to consider new business models that will enable VREs to become self-sustaining as far as possible.
- There is a need for funding councils to **move beyond the traditional funding models** of projects funded for the short-term. At the same time funding councils will increasingly play a role beyond national boundaries in the wider international development of research infrastructure.
- **Libraries have a key role to play** in several aspects of VREs. They are increasingly becoming key stakeholders in the sustainability of the data outputs of VREs, and possibly for further aspects as well. Libraries also play a very important role in the development of VREs since they are instrumental in the data and resource management. This is likely to

affect both the practices of librarians, and the research practices and processes they are supporting.

- Institutions investing in VREs are motivated by their pursuit of research excellence. However, the **model for the VRE needs to be driven by researchers** rather than being imposed by the institution.
- A major shift in research practices will occur through the formation of common vocabularies as researchers collaborate with others across disciplinary, institutional and national boundaries. This will occur through, for example, the **production of common taxonomies, data standards and metadata**. Semantic web approaches are seen as helpful in this context.
- It is extremely important that all stakeholders in the development of VREs come together to promote **a set of policies and legal frameworks** that will allow sharing of data and other resources in a transparent and comprehensible way.

1.4 Methodology

The study used a combination of desk research, case studies, surveys and interviews in order to gain an understanding of the main developments, approaches and trends in VREs. A draft report of findings was sent to four international reviewers, and their feedback was incorporated into the report.

1.4.1 Challenges

Scoping the international VRE landscape and identifying trends is not an easy task, especially if there is only a limited amount of time and resources to do it. Identifying what actually qualifies as a Virtual Research Environment and what does not is the first challenge. This is not made easier by the fact that, even in the English speaking world, different countries and disciplines use different terms: Collaborative Virtual Environment, Collaboratory, Gateway, Science Gateway, Virtual Organisation, even the term Portal is sometimes used to refer to an environment that, following the JISC definition, could be classified as a VRE.

Another difficulty with identifying VREs is that not all of them are publicly visible: VREs that are just for internal use might not be publicised or not even have an outward facing website. As Virtual Research Environments are a relatively recent development, it may also be assumed that a considerable number of VRE projects may be in an early stage of the project and hence more difficult to find, unless the project undertook outreach activities from the start. The only thing that can be said with certainty is that no one knows how many VREs currently exist.

Against this background, it is impossible to obtain an exhaustive picture of the current global position regarding VREs. Our study aimed to gain an overview of the main features of the general landscape rather than to try to fill in every detail. Thus our study was never meant to be representative in the statistical sense or to be a conclusive, definitive study of Virtual Research Environments. What we do hope, however, is to contribute to an ongoing discussion on VREs and virtual research collaboration by identifying trends and issues from an international perspective, based on the experiences of a selection of those engaging with VREs and on the lessons learned from projects and through a review of recent literature. Based on initial desk research, which was followed by a widely circulated online survey, our project then focussed on conducting qualitative interviews and developing case studies on VRE projects and programmes from four continents.

1.4.2 Desk research

For our initial research, project staff looked into recent reports and publications on VREs and virtual research collaboration (see the bibliography in the appendix); we also followed up on recommendations from interviewees and participants of the survey. It should be noted that the bibliography could potentially be much more extensive, since aspects of research collaboration are referenced over a wide range of publications. However, because of the short project duration we mostly concentrated on those publications that were more specifically focussed on the topic of our study. We also identified funding bodies that supported VREs and looked at recent calls and key projects funded through these programmes.

1.4.3 Online Survey

Based on lessons learned through the desk research, we developed an online survey to gather feedback from the international VRE community and to help us to identify people and projects to engage with in more detail. The survey was developed using the Bristol Online Survey (BOS) platform and ran from 23/08-30/11/2009. In order to spread the survey as far as possible, project

staff sent out over 200 emails to relevant institutions, projects, networks, individuals and mailing lists across the world. The starting points were JISC mailing lists and VRE projects in the UK, followed by e-research and e-science centres, also internationally. Following that, we contacted international funding bodies and VRE related programmes, always asking for both the circulation of the survey invitation and further recommendations on who to contact. We circulated invitations via relevant mailing lists, newsletters and -sites, such as InetBib[2], International Science Grid This Week (iSGTW[3]) or the eResearch Australasia Newsletter.[4] We also conducted web-searches to find individuals and projects with an interest in VREs, using a variety of related keywords in different languages. In addition to that we made use of the network of contacts of our colleagues at both Oxford e-Research and the Centre for e-Research at King's College London. As a result, we received 207 responses, 86 of which completed the full survey. Results from the survey were used in the following ways: 1) to identify individuals with expertise relevant to our study who were willing to be interviewed in more detail; 2) to identify projects for case studies; 3) to collect recommendations on literature and relevant contacts and resources; 4) to gather some qualitative material, particularly suggestions on important trends, general comments on VREs and recommendations; 5) to gather some statistical information on VREs and the user communities, with the understanding that this data would not be representative, but rather illustrative.

Participants in the online survey

It is interesting to see who responded to the survey invitation. In geographical terms, most of the participants came from the UK and continental Europe, with North America and Australia/Oceania being almost as well represented. Despite the project team making a serious effort to invite responses from other continents, for instance through sending emails to individual researchers and centres in the Middle East, Asia, Africa and South America, only a handful of participants came from those regions and especially from the less developed countries.

This is the geographical spread of the participants who decided to share their location:

Geographical location of survey participants		
Africa:	1.40%	3
Asia:	4.30%	9
Australia and Oceania:	12.10%	25
Europe (not United Kingdom):	29.00%	60
United Kingdom:	37.70%	78
Latin America and Caribbean:	1.40%	3
Middle East and North Africa:	0.50%	1
North America:	13.50%	28

However, this geographical distribution should not be mistaken as proof that VREs are necessarily used significantly more in the UK than anywhere else in the world. First of all it has to be said that, mostly because of JISC, the term 'Virtual Research Environment' is much more prominent in the UK than anywhere else. Together with the association of this study with the JISC VRE Programme and the fact that both project partners are based in the UK, this may be reason enough to explain the prominence of responses to a survey on VREs from the UK. Nevertheless, both the literature and our interviewees agree that the UK is leading the field in VRE develop-

2 http://www.inetbib.de/whatisinetbib.html [Accessed 04/01/2010], a German language list with over 5,000 subscribers with an interest and expertise in Internet and digital technology in libraries.

3 http://www.isgtw.org/ [Accessed 04/01/2010].

4 http://www.eresearch.edu.au/ [Accessed 04/01/2010].

ment; however, countries such as Germany, the Netherlands and Australia are also showing many activities in this field, as reflected in the geographical spread of participants.

In terms of participating disciplines, the arts and humanities were most strongly represented. We can offer two reasons for this. First of all, both project partners have a very good network in this domain, which in itself may explain the higher response rate. Secondly, there are some signs that the arts and humanities currently have a special interest in VREs – for a more detailed discussion see the section on disciplines below. It is not surprising that computer science is strongly represented in a field where at least the VRE builders need to have many technical skills.

Participants by subject domain		
Arts and Humanities:	34.30%	71
Biotechnology and Biological Sciences:	4.30%	9
Computer Science:	20.80%	43
Engineering and Physical Sciences:	5.80%	12
Economics and Social Sciences:	11.10%	23
Medical Sciences:	3.90%	8
Natural Sciences:	1.90%	4
Other:	17.90%	37

Participating librarians and information managers mostly did not see themselves closely associated with a particular discipline and instead chose the 'other' option. The largest group of participants in the online survey were researchers:

Participants by role		
Administration:	3.90%	8
Librarian or Information Management:	16.90%	35
Management:	10.10%	21
Professor, Reader or Associate Professor:	24.20%	50
Researcher (Research Associate or Assistant):	18.40%	38
Teaching staff (Senior Lecturer or Lecturer):	6.30%	13
Software Developer/Architect:	7.20%	15
Support Staff:	1.90%	4
Other:	11.10%	23

1.4.4 Interviews

Over the course of this project, we interviewed 26 individuals, mostly over the phone or using Skype software. Interviewees were mostly selected through the online survey, but also following recommendations from other interviewees and through existing contacts of the project partners. Interviews took about 45-60 minutes and followed a semi-structured approach that allowed the interviewee to develop ideas or explain issues in the necessary detail. Where we had permission, the interviews were recorded and then transcribed; in a few other cases the interviewers took notes during the interviews.

1.4.5 Case studies

Based on the information we gathered, the project developed twenty one case studies, seven of which focus on the experience from international VRE programmes (or VRE related programmes) and countries, the other fourteen on specific VRE projects. The programme case studies were selected based on our initial desk research and with the aim of looking at a range of countries, and not only those with mature programmes in place. As we only had a short time span in which to conduct the interviews, further criteria for both the project and programme case studies were the availability of interviewees and the responses that we obtained from our invitations to participate in the study. Project case studies were selected from the projects that we identified through the initial desk research and, mostly, from the survey. Choosing from that sample, our aim was to select a broad range of projects in order to capture different perspectives and experiences, using the following criteria: region of origin; size of the user community; different technologies used; different subject domains; single or multidisciplinary communities; VREs used by one or several institutions; VREs in different stages of development; VREs that mostly applied existing technologies or those who developed new ones; project specific VREs vs. VRE framework projects; institutional vs. project specific VREs.

1.5 Key findings from literature

Specific literature on Virtual Research Environments is still somewhat scarce and while there are several works on online collaboration, the term VRE is still not very much in use. A prominent study of collaboratories was undertaken by the Science of Collaboratories Project[5], and the findings of that project are largely relevant to VREs too.

The definition of VRE is also not yet agreed, nationally or internationally. Following the literature, a VRE can be described as a set of web applications, online tools, systems and processes interoperating to facilitate or enhance the research process within and without institutional boundaries; it enables collaborative research activities beyond geographical barriers. Industry, universities, other research institutions and government are all involved in collaborative research projects.

Several other terms are used instead of VRE, such as: CVE (Collaborative Virtual Environment), Cyberinfrastructure/e-Infrastructure, Collaborative e-Research Communities, VRC (Virtual Research Community), VO (Virtual Organisation).

The main benefits of VREs are: scholarly collaboration over a distance, exchange of information among scholars, access to skills, knowledge, research data and computational resources situated in remote locations and cooperative writing of academic material.

Information can be shared by telephone, email, instant messaging/chats, forums, wikis, blogs, meeting tools, project management tools, video conferencing, data-based conferencing, Access Grid, project calendars, task assignments.

The main challenges and problems for the development and success of VREs are: space, time, funding, isolation, procrastination, poor motivation, trust, commitment, working style, ownership, data access, difficulty of learning software and technology, lack of appropriate skills and ready access to technical support and extensive training needs.

For a successful VRE it is important to have clear ownership of the data, a mutually agreed project plan among the collaborators, clearly defined objectives and responsibilities, and adequate resources. The literature stresses the need for the software and technology to be easy to use. It is also important to promote the systems and make the scholars aware of how the systems and tools can simplify their workflows and more widely disseminate their work.

The overriding impression in the literature is that there are substantially more VREs in the natural sciences compared to social sciences and humanities. It is thought that scholars in the science disciplines are likely to already be using digital repositories and VRE systems. Historically, in the arts and humanities collaboration has tended to be based around smaller networks of scholars; however over recent years there has been a marked shift towards larger collaborations, along with a significant rise in the volume of digital data, the increasing use of ICT-based methods and the rapid advantage of technology. Among the arts and humanities disciplines, archaeology has been remarkably successful in developing e-research infrastructure.

International comparisons have revealed that the UK is well advanced in its understanding of the area and has the world's best structured programme of developments under way. In the UK most of the research and development has been funded by the JISC. UK developments have included work in both the sciences and humanities. Many other countries around the world are engaged in developing VREs including: Netherlands, North America, Australia, Germany, South Africa, Japan, India and Brazil.

5 http://www.scienceofcolaboratories.org. [Accessed 18/12/2009].

2 VRE Landscape

2.1 The term 'VRE'

Any review of the international VRE landscape has to start with the term 'Virtual Research Environment'. While it is widely used in the UK, mostly due to the impact of the JISC VRE programme, other terms are also used here to describe identical or very similar environments. Even considering only other English speaking countries, one realises that a variety of terms are used to refer to collaborative online environments for research and it is not always clear what the differences are. Apart from the term 'VRE', these are some of the related English terms used:

- Collaborative e-Research Communities
- Collaborative Virtual Environment
- Collaboratory
- (Science) Gateway
- Virtual Organisation
- Virtual Research Community

The most notable difference is that all the definitions using the word 'community' focus on the group of researchers using a virtual environment as opposed to the VRE itself. Following this line of thinking, one could argue that the VRE is the environment through which a researcher engages with, and becomes part of, a Virtual Research Community.

In order to better understand the different terms, it is helpful to look at the definitions suggested by some of the organisations that champion these concepts. The most important one, in our context, is the definition of a VRE as given by JISC:

> *The purpose of a Virtual Research Environment (VRE) is to help researchers from all disciplines to work collaboratively by managing the increasingly complex range of tasks involved in carrying out research on both small and large scales. The concept of a VRE is evolving. The term VRE is now best thought of as shorthand for the tools and technologies needed by researchers to do their research, interact with other researchers (who may come from different disciplines, institutions or even countries) and to make use of resources and technical infrastructures available both locally and nationally. The term VRE also incorporates the context in which those tools and technologies are used. The detailed design of a VRE will depend on many factors including discipline, context, and security requirements. The intention of this programme is therefore not to produce a complete VRE, but rather to define and help to develop VRE frameworks and associated standards, and to encourage the development and population of these frameworks with applications, services and resources to create VREs appropriate to particular needs.*[6]

Several key aspects emerge from this definition. The definition is fairly vague and deliberately so. Playing devil's advocate, we could say that a combination of telephone, letter and typewriter could be understood as a VRE, since they are tools that allow collaboration between researchers and are all part of a technical infrastructure. The definition reflects the fact that the concept of a VRE is indeed evolving and all representatives of funding bodies we spoke to agree that being too prescriptive in the definition might stifle innovation. Yet, what clearly emerges from the JISC definition is that a VRE facilitates collaboration between researchers and provides access to data, tools and services through a technical framework that accesses a wider research infrastructure. This is also reflected in the definition of VRE given by the DFG, Germany's research

6 http://www.jisc.ac.uk/whatwedo/programmes/vre.aspx [Accessed 18/12/2009].

council. The DFG defines 'Virtuelle Forschungsumgebung' (a literal translation of 'Virtual Research Environment') as:

> *a platform for internet-based collaborative working that enables new ways of collaboration and a new way of dealing with research data and information.*[7]

It has to be noted that the DFG also does not put too much emphasis on the details of this definition in order to give researchers and developers a certain amount of flexibility when developing their concepts. Again, the focus is on collaboration and on an environment, here clearly located online, that contributes to a research process.

Looking at other definitions from this perspective, the different terms are not actually that different in terms of their meaning. 'Collaboratory', for instance, is one of those terms. The term is a hybrid of 'collaborate' and 'laboratory', and even though that may suggest a focus on the sciences it is not actually limited to a particular domain. The Dutch SURFfoundation, for instance, defines 'collaboratory' as follows:

> *A collaboratory is a web-based collaborative electronic environment that enables researchers based in different locations to work together and share their knowledge and facilities, thus enriching and speeding up both national and international research.*[8]

This definition adds different locations and a focus on speeding up research to the definition of the DFG, but these two aspects are actually as important for the DFG as they are for any other funding body (see the DFG VRE programme case study for more details).

It is interesting to look at one more definition, this time the definition of the term '(Science) Gateway' as given by TeraGrid Gateways:

> *A Science Gateway is a community-developed set of tools, applications, and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is customized to meet the needs of the targeted community.*[9]

Compared to all other cited definitions, this one focuses much more on the technical aspect of the actual virtual environment, but again the importance of access to data and tools/services through a framework to support a specific community clearly emerges.

It has to be noted that there also seems to be a consensus about what a VRE is in relation to research infrastructure in a wider sense. While 'e-Infrastructure' (or 'cyberinfrastructure' in the US) refers to all aspects of the digital side of research infrastructure, VREs are understood to be an interface to that infrastructure, allowing easy access to data and services in an environment that is focussed on a particular research activity. One of the interviewees elaborated on this in relation to the term e-research:

> *The VRE, to me, is in the first instance about collaboration across boundaries. A VRE both sits on top of the electronic architecture but is also embedded in it. A VRE is almost like an intranet designed exclusively for researchers irrespective of their location; it pulls together in one place absolutely everything that could assist research staff to become very competent and efficient collaborative researchers (it gives them access to all the electronic tools that they need to 'do' research successfully, all the systems they need to manipulate, all the documentation that needs to be accessed or completed, all the bits and pieces that go missing at crucial points of the*

7 http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/projektfoerderung/foerderziele/virtuelle_forschungsumgebungen.html [Accessed 18/12/2009].

8 http://www.surffoundation.nl/en/themas/openonderzoek/collaboratories/Pages/Default.aspx [Accessed 18/12/2009].

9 http://www.teragrid.org/gateways/#what [Accessed 18/12/2009].

> *research process)! It is about making use of the VRE to take care of the 'organise and control' aspects of project management.*

An interesting aspect of this view is that it indicates an integration of physical and virtual research spaces, where researchers, through the VRE, move from one to the other. A participant in our survey also related to that aspect, arguing that speaking of *Virtual* Research Environments was maybe not the right term as it obscured the physical aspect of research:

> *Also, a VRE must fully integrate with physical research environment – not all research is done online, and not all online activity is done on the desktop. I prefer 'Hybrid Research Environment' to VRE.*

Even though we did not find any use of this particular term, it might be argued that as VREs become more integrated into every day research activities, it will become more useful to just see them as part of a general research environment. An interviewee commented:

> *The names we are using now are from a specific focus point. For instance, these environments are not really virtual any more. They are happening in everyday research life now, so calling them 'virtual', or 'cyber', for that matter – sounds a bit like something futuristic, whereas we are actually trying to make it an everyday part of the researcher's day-to-day practice.*

VREs may still have a long way to go in this respect, but it could be that this transformation is happening in the same way it happened to the small letter 'e' being attached to all sorts of concepts. A few years ago, for example, it may have made sense to refer to 'e-content', whereas now most content is digital and the 'e'-suffix has become somewhat meaningless. It seems to us that on a practical level this is already happening with VREs. It can be illustrated by thinking about Virtual Research Environments, Portals and Digital Libraries. For instance, when a portal environment goes beyond allowing access to distributed data, and also allows for the sharing and annotating of data, is it of much relevance to the researchers who use it whether it is called a portal or VRE? As an example, the NCeSS VRE (see case study) is called a portal and this does not seem to be of any relevance to its users. The Australian MemRE project is building a digital library of materials on membrane technology – but in a collaborative way that also uses a wiki in order to create a dictionary.

Libraries have for many centuries been places in which actual research takes place, so it is not surprising to see that the same happens within digital libraries. Partly for this reason, libraries are developing a keen interest in VREs, as our study shows. Because of these reasons we suggest that it may be of more use to focus on what actually happens with and within VREs instead of arguing the differences between different ways of referring to them. This view was widely shared among our interviewees too, for instance by John Doove of the SURFfoundation:

> *We had the same sort of discussion here. 'Collaboratories' just emphasises the collaborating part, but it has more to offer than just collaborating. 'Virtual Research Environment', on the other hand, sometimes has the connotation of just being a virtual lab in which researchers can work. So it is difficult to find a name, although we have figured out that researchers do not really care about what name you give it, as long as it helps them in the research process. So we stuck to the collaboratories name, because it is not really worth discussing over and over again.*

To summarise, we found that the term used was not important, though the understandings associated with the terms 'VRE', 'Collaboratory' and 'Gateway' are converging on a set of characteristic features: an electronic web-based environment for a) access to data, tools, resources; b) co-operation or collaboration with other researchers at the same or different institutions; c) co-operation at the intra- and inter-institutional levels; or d) preserving or taking care of data and other outputs. Not all of these environments serve all of these functions, but they generally serve two or more.

2.2 Motivations for funding, building and using VREs

2.2.1 Funding strategies in an international context

Before looking at different VRE programmes in more detail, it is useful to consider why funding bodies support the development of VREs in the first place. While some of them, such as JISC, have a special remit to support the use of ICT in education and research, this is not true for all funding bodies that fund VRE development. Even so, funders are prepared to invest money in VREs because they feel that these environments potentially offer many benefits including:

- support for geographically dispersed research teams
- facilitation of international collaboration
- support for general networking
- support for interdisciplinary research
- increasing the productivity of researchers
- facilitation of access to (expensive) research infrastructure
- increasing speed of general communication
- faster dissemination of research results, including preliminary findings
- better preservation of research outputs and also the information on the process that led to them
- and, perhaps most importantly, a new quality of research outputs.

VREs are not the only way of addressing at least some of these issues, so it should not come as a surprise that funding programmes can have a different focus. Generally, we found that funding strategies regarding VRE or similar programmes fell into three main categories. In the first category are dedicated VRE or similar programmes; in the second category are programmes which do not explicitly see themselves as promoting VREs as such, but where there is an overlap with explicitly VRE or VRE-like programmes; in the third category are programmes which do not target anything like VREs.

Examples of funding strategies in the first category are those pursued by JISC's VRE programme in the UK, the Collaboratory component in the SURFshare programme in the Netherlands (see case study), the Science Gateways in the USA (see case study), and the DFG in Germany (see case study). All of these programmes share a relatively similar vision of key elements of VREs, as shown above, and they are specifically aimed at facilitating the shared use of digital infrastructure by researchers through the provision of shared environments.

Examples of funding strategies in the second category are those pursued by Adonis in France (case study) and to a lesser extent Norgrid in Norway (case study). These strategies do not explicitly see themselves as forming communities which co-operate or collaborate in VRE-like structures. However, many of the projects funded under these programmes do finally operate in ways substantially similar to VREs.

Examples of funding strategies in the third category are those pursued in Spain and in Italy. In these countries there are substantial high performance computing and other e-science resources that are mostly targeted at physical sciences, but we could not discover initiatives similar to VRE programmes.

Most countries have a mixed approach for different disciplines. For example, in the Netherlands, VREs are targeted more at the humanities and social sciences, and in France, of all the TGE (Large Scale Research domains) only Adonis seemed to have VRE like features.

An important motivating factor for particular funding strategies is the international context. Thus the First European Roadmap for new Large Scale Research e-Infrastructures is an important in-

centive for European strategy.[10] This was underscored in the 2009 FP7 Research Infrastructures Call for Proposals, for the creation and support of cross-European infrastructure for research collaborations, with a strong emphasis on ICT-based e-infrastructures.[11]

and the need to collaborate internationally is a further incentive. This was seen in several of the case studies: for example, Adonis, NorGrid, HubLab. In countries with no VRE initiative of their own, such as Italy, European funding is often the only route to international collaboration. In developing countries, such as Ghana and South Africa (see case studies), international collaboration via an e-infrastructure is often the only way in which to conduct certain kinds of research. However, these countries often face challenges regarding the technical infrastructure, such as bandwidth.

2.2.2 Academic institutions

The main motivation driving VRE development in institutions is the need for optimisation of information and knowledge management, in a wider sense. From the institution's perspective, potential synergies are often not realised because information is lost or hard to find. This is making grant development more difficult and time consuming; it increases project management workload and may also reduce the use of resources in which the institution has invested money. This is of particular relevance to libraries that want to make sure that researchers are aware of databases and other resources that the library provides, often at great cost. With regards to research data management, there is also a danger that data will be lost if they are not properly managed and curated, while the promotion of research outputs, especially publications, is made more difficult when information is stored (or not stored) in a variety of systems and websites, with or without proper security in place. Furthermore, VREs are seen as having great benefits for the support of geographically dispersed research teams and remote working of staff, two trends that seem to be gaining in importance.

Large institutions in particular have an interest in developing VREs as part of a more general research infrastructure, without which individual centres or research groups might waste resources by developing solutions for problems that other units have already addressed. While the main focus from the institutions' perspective may be to support existing research, there is also an increasing awareness that VREs, especially interdisciplinary ones, have the potential to enable new research in a way that is much more than 'just' a supporting role.

Facilitate grant development

Academic institutions have a strong interest in increasing the number of (successful) grant applications to external funding bodies, as they provide additional income, increase research outputs and lead to more academic prestige that in turn can translate into more students. They are also aware that academic staff can spend considerable time on the administrative side of developing research grants, a task that rarely interests researchers and may also not represent the best use of their time. This makes providing easy access to information on developing grant applications, together with up-to-date information on recent funding calls, a priority from an institutional perspective. While this can be provided as a stand-alone module, it makes sense to integrate it into VREs so that project development, management and the actual research collaboration can happen in the same environment. Researchers also have an interest in this where a VRE has demonstrable benefits for their work, for instance through saving them time. A good example is the VRE of the Noguchi Memorial Institute for Medical Research (case study), where grant management support is the first module to be released.

10 ESFRI European Roadmap for Research Infrastructures, 2008.

11 http://cordis.europa.eu/fp7/dc/index.cfm?fuseaction=UserSite.CapacitiesDetailsCallPage&call_id=263 [Accessed 18/12/2009].

Encourage collaboration within the institution

Researchers are often very active in developing networks in collaborations with partners from their subject domain, whereas they are not necessarily aware of relevant research in other departments within their institution. VREs that associate people with projects, tasks and research interests can thus be seen as a way of making researchers more aware of potential collaborators within their institution, especially when the VRE has a social networking component. However, staff are not necessarily always interested in these tools, especially in smaller organisations, such as the Virtual Knowledge Lab, where it was decided not to implement dedicated social networking support because researchers felt their personal profile on the website was sufficient.

Support geographically dispersed teams and remote working

From an institutional perspective, supporting geographically dispersed teams is becoming increasingly important. The most obvious reason for this is the internationalisation of research, driven by specialisation and international funding streams and also encouraged by the increasing speed of communication. While the same is also true for collaboration on a national level, geographically dispersed teams do also operate within the same institution – even comparatively small ones such as the VKS in the Netherlands. Furthermore, digital communication is also encouraging remote working for staff, be it because of regular travel or increasing use of the home office in a more flexible working environment.

To support work across locations, facilities are needed for sharing of data and for communication, with the addition of other tools and services as used by a specific project/group. Opening VREs to collaborators from outside the institutions' networks requires an infrastructure for secure access and authentication, especially where sensitive research data are concerned. The e-Resource Centre VRE (see case study) is a good example for such a project: researchers from Australian government institutions and universities needed a secure environment to share, discuss and analyse data across institutions.

Provide general project management support

Project management tasks such as organising meetings can take up a considerable amount of time, especially in collaborations across institutions, disciplines and countries. Freeing up as much time as possible to enable staff to focus on their research is in the interest of institutions; also, this allows administrators and project managers to better understand specific projects and what support and resources they might need. VREs can provide tools for scheduling and planing meetings; tracking of tasks; sharing project-related information; costing bids; etc.

Information and knowledge management

Looking at information and knowledge management in a wider sense, VREs have the potential to provide answers to a variety of issues that concern academic institutions, especially libraries. First of all, there is the issue of preserving research data. While institutional repositories do exist in some universities, they are mostly used for publications and not for actual research data. Furthermore, repositories can often seem somewhat alien to researchers as they reflect the institutional perspective and not necessarily how researchers work. Integrating an architecture for data management into a VRE can address both of these issues, especially as our research has shown that researchers do care about secure storage of their data, provided it is (relatively) easy to use and they can keep control of their work. Encouraging researchers to use a VRE for managing their data can also help with preservation, especially if the environment comes with a well thought out data management plan and the tools provided to use and create data in documented formats. Sustainability of VREs is a critical issue with regards to preservation of research data. Because VREs are (ideally) well integrated with the research process, they also make a good

access point for institutional publication repositories. Academic institutions, especially libraries, are not only concerned with managing data created by but also for researchers. Libraries invest an increasing amount of money in subscriptions to databases and online resources and want to make sure that these resources are used. Integrating these resources into domain specific VREs is thus of interest to libraries, as shown by the example of the British Library which uses the RIC VRE (see case study) to give biomedical researchers access to a range of specific resources such as PubMed Central.

2.2.3 Researchers

Perhaps somewhat unexpectedly, we found that the motivations of individual researchers and research teams overlap substantially with those of their institutions. Despite this apparent overlap in motivations, the same motivation can look very different from the perspective of an institution and that of an individual researcher. A digital publications repository designed to contribute to the reputation of an academic institution, for instance, can lead to a very different implementation of the same repository management system than one that suits the needs of a researcher.

As a generalisation, researchers tend to think in terms of their own subject domain, their career and their specific way of working, which can lead to certain tensions with a more institutional perspective. Another difference seems to be the time-scales in which researchers and institutions seem to think: it was noted by several interviewees that while researchers do seem to show an increasing awareness of the benefits of shared access to data and information, that does not necessarily translate into an understanding of the importance and issues of long-term preservation and re-use of data.

A point on which the views of institutions and researchers often differ is access to the institution's network infrastructure. Restrictions that may seem reasonable from the perspective of securely administering a university-wide network can cause problems for researchers who need to share certain resources with partners outside their institution. In the same way it is possible to look at tools designed to facilitate grant development either as a way of supporting researchers or of forcing them to follow an institutional model that may contradict their way of working. It is important not to forget or downplay these tensions in the light of many similar motivations.

Our survey gives an indication of the activities of researchers as supported by various virtual environments:

Which of the following functions is the VRE used for?	
analyse and process data:	32
collaboratively annotate data:	41
provide access to tools, services or an infrastructure:	55
share data with others:	72
support communication in a team:	64
support project management:	44

Storing and sharing data

> *One of the first things researchers want to do when they are in a collaboratory is to share data amongst their peers. Not in an open-access data kind of way to the public, that is a whole different story, but really among their peers, to work on the same data set. (quote from an interview)*

Our research has shown that researchers are highly motivated to access and share data. In many instances, it is not possible to do research in a particular domain without being able to ac-

cess resources that are geographically distributed. This is true for researchers in the physical sciences, such as those in the Nordic Data Grid Facility who are accessing data from CERN through their Virtual Research Organisations, as well as for researchers in the humanities and social sciences, attested by the several projects under Adonis which aim to make specific data sets accessible. Sharing data is also a prime motivation for projects such as the VKS collaboratory (see case study) and HubLab (see case study). It has to be noted though that the interest in sharing data can differ across disciplines and contexts: biologists working on drug design, for instance, might not be significantly less likely to share their data then climatologists or other disciplines.

A further motivation relating to data and presentation is that there may be no other way for it to be presented than in a virtual environment. For example, the Orlando Project (see case study) was originally motivated by the existence of a great deal of material which could not be included in a book because of space constraints.

This increasing interest in sharing data in collaborative projects is also noted by funders; the DFG, for instance, has reacted by offering special funding to the research groups it supports in order to enable them to build and operate facilities to store and share data.

Workflows

Beyond data, accessing workflows for experiments is a further important motivation for participating in VREs, especially in the sciences. MyExperiment (see case study) is a leader with regards to infrastructure for accessing and sharing workflows. Adopting a Web 2.0 social networking type of functionality, this VRE allows the embedding of workflows within rich contexts, including not only the data and metadata, but also the social contexts of the workflows. User-generated content such as tags, number of downloads, reviews and discussions around the workflows, as well as social networking aspects such as the creation of groups among which workflows are shared and friend lists, all contribute to a rich context for workflows that makes them properly usable by others. Like data, workflows are not stand-alone digital objects, but work best when interlinked with other objects and activities around them.

Linked data and information resources

A further motivation for VREs is to provide access to the added value created by linking data and information resources.[12] This is, for example, the approach of Alzforum and related sites (see case study). Being able to link abstracts of articles with data, as well as with comments and discussions, enables researchers to follow the development of hypotheses as they are formulated and tested, and the evidence for or against them.[13] This is a semantic web approach, sometimes described as Web 3.0, as it tries to bring together the power of social networking and the semantic web.

Collaborative activities

Researchers are also motivated to build and join VREs by the collaborative activities that they can participate in through the VREs. These include, among others:

Collaborative analysis and interpretation

In several research domains, including humanities and social science research, researchers share data and collaborate in the analysis and interpretation of the data. For example, VREs cre-

12 On the concept of linked data see http://linkeddata.org/ [Accessed 04/01/2010].

13 See for example the 'Enhanced publications' concept of the SURFfoundation: http://www.surffoundation.nl/en/themas/openonderzoek/verrijktepublicaties/Pages/Default.aspx [Accessed 04/01/2010].

ated for archaeology and classics enable collaborators to access geographically dispersed data, and also to collaborate in the analysis thereof, reading a text together.

Writing and publication

A very important activity for researchers is collaborative writing, which is often supported by external sites such as Google Apps. Collaborative writing is an important aspect of writing grant proposals as well as presenting the outcomes of research in journal articles or other publications. In the case of the Orlando Project, collaborative writing was achieved within the site through the use of an XML editor. Another example is Revues.org (part of the Cléo project) which supports the collaborative publication of journals by providing access to the Lodel software developed by the team.

Communication, discussing and debating

A high priority is placed on being able to access other researchers through the VRE. Communicating with other researchers is a further strong motivation for individual researchers. Sometimes communication is simply a matter of discussion and debate. This is highly valued by researchers ranging from the biomedical sciences (Alzforum, case study), and Hypothèse.org, under Cléo (case study). At other times, VREs facilitate communication among colleagues who are geographically dispersed, yet needing to collaborate on the analysis of data or on writing research papers or grant proposals. Having an integrated environment where data and other resources can be shared and collaborative writing can be carried out is a strong motivating factor for humanities scholars in particular (see for example the case studies on CWRC and VKS).

Participating in the VRE also stimulates further discussions outside and beyond the VRE. For example, one interviewee remarked:

> *There is more openness and more discussion – especially during the workshops. So much discussion is not that common and is new. So in that sense, this is not business as usual.*

Accessing software, tools and supercomputing resources

Access to supercomputing resources can also be the focal point of particular VREs, such as many of the Science Gateways. Apart from the use of these resources for large databases, researchers also require access in order to run simulations and visualisations or other computing intensive processes.

Software and tools are also accessed and used in VREs. For example, apart from data facilities TextGrid gives access to tools such as an XML editor, Lemmatizer, Tokenizer, Metadata Annotator, and a Dictionary Search Tool, whereas one of the Adonis projects, Archeogrid[14], gives access to 3D visualising tools for archaeological objects.

Project management

Staff in academic institutions can spend a considerable amount of their time on aspects of project management, including the development of grant applications. Reducing that time and providing better access to information about new funding calls and the application process is in the interest of both researchers and the institution. Researchers do not necessarily always take a great interest in project management. However, including tools for this purpose into the same environment that researchers use for their research practice may increase their interest as they do not feel they are forced to use an application that is 'just' about administration. Archiving information relating to project management in an accessible way may also allow other researchers to profit from work that colleagues did previously.

14 http://www.tge-adonis.fr/?Le-projet-ARCHEOGRID-Conservatoire [Accessed 18/12/2009].

Apart from actual bid development, VREs support scheduling (even though some interviewees expressed a preference for the use of free online calendars such as the one provided by Google) and often offer a collaborative space, usually through a wiki, to plan project meetings and archive meeting notes. Tasks can also be managed and allocated in that way.

It could be argued that supporting project management through a VRE was not research in the strict sense and that email and a file space could provide the same functionality. However, one interviewee made a particularly strong case for integrating such functionality in a VRE: he argued that information related to project management information provides the context for the actual research, and if it is scattered across email accounts and external systems, that context can be lost – and with it the chance to understand how certain decisions were made.

News and dissemination of information

VREs are also used to disseminate all sorts of information relating to a project and also news from the wider community in which a project operates. All collaboratories supported by the SURF Foundation, for instance, have a news section. This is also found in Alzforum and Schizophrenia Research Forum, which include news about the general community including very personal things, like births and deaths.

Contribution to research

The most important motivating factor for researchers is ultimately the contribution that VREs make to research. Many VREs are still in early stages, and therefore it is difficult to assess their contribution to research as yet. This may also be reflected in a fairly high percentage of participants of our survey who were 'not sure yet' whether or how far the VRE had already contributed to their research. Again, these figures are only indicative and not conclusive, especially as it appears that some researchers would not count the support that a VRE can provide to project and data management and to communication as a contribution to the actual research. The seventeen participants who chose 'not applicable' mostly engaged with VREs as support or development staff:

Has using the Virtual Research Environment made a contribution to your research?	
yes:	52
no:	1
not sure yet:	19
not applicable:	17

The transformative role of VREs

Well functioning VREs demonstrate clearly that 'the whole is more than the sum of its parts'. This is related to the most important motivating factor for individual researchers or individual centres of research, namely the potential of VREs to take research to the next level. For example, the global sharing of data facilitated by the HubLab project (see case study) will enable the writing of a world economic history, which would not otherwise be possible.

> *We can only do this when we have people working in every country or every continent, helping to create one large collection of data. For that of course you have to have tools or means to get these people together and to let them work more intensively than it would be possible if you can only bring them together once a year for a workshop.*

This has also been the main motivation for existing VRE projects which are now beginning to deliver on this promise, such as the eSAD project which has enabled the re-interpretation of an an-

cient document[15]. For Alzforum and related projects (see case study), the semantic web enabled integration of articles, journals, comments and data has proven to be a tremendously important resource for the progress of research in this field.

Another transformative role that was ascribed to VREs was to enable interdisciplinarity as a way for new research, especially in relation to access to large datasets. Bringing together data and approaches from different disciplines was seen as key for creating new research findings. One of the examples given was the potential that linking and analysing migration and climate data over a long period could have. VREs, partly as a means to access data storage in a wider research infrastructure, have an important role in such research.

15 Bowman et al (2009).

2.3 Building VREs: community development projects

In terms of development of VREs, three main trends emerged from our research: 1) VREs are developed as a collaboration between researchers (users) and infrastructure developers, with the researchers ideally taking the lead on producing the requirements. 2) Development is mostly or ideally taken forward in an iterative process and driven by feedback from the users. 3) At least the research-oriented libraries increasingly see the development of VREs as part of their remit as they are aiming to support researchers throughout the whole research lifecycle.

Because of these trends, we are discussing the development of VREs in the context of a process that, ideally, integrates the user community right from the start of a project – if they are not actually the ones who initiate it. The importance of not building a VRE *for*, but *with* the researchers cannot be emphasised enough. One interviewee summarised it from the perspective of a funding body:

> *You really need to answer the needs of the researchers instead of creating something for them and hoping that they will take it up. That was a direction that we actually started from a few years ago, when we thought, well this is going to be very cool for researchers, so let's develop it, and let's throw it at them and see what they do with it. If you take that approach they will not do anything with it at all. Some enthusiasts might, but you will not reach the whole community. So we are actually listening better to our researchers. Maybe we should have done that from the start! We are now actually almost sitting next to the researchers and seeing what they are doing in the research environment, and how we can make things more efficient in a whole lot of ways, from collaboration to research data sharing to communication.*

2.3.1 Researchers and requirements

In all the projects we looked at, researchers were, at least to a certain extent, involved in the definition of requirements and the creation of the original concept; all projects involved the researchers in their development, a process that was often taken forward as a partnership. Several interviewees noted that researchers are overall not yet technically aware enough to formulate their requirements clearly without some guidance, even though there are signs of an increasing awareness, at least for the benefits of repositories for sharing and preserving data. This is especially true for researchers collaborating across countries, institutions and disciplines as they do clearly see the benefit of a good file storage system.

We have also come across projects that were not only initiated by the research community, but where researchers approached a development partner, usually a library, with a very clear idea of functionalities and systems. The Membranes Research Environment (MemRE) is one example of this. However, even in such cases researchers still need partners for development, especially if the environment is meant to be sustainable, something which libraries can more often guarantee than projects or even research clusters. Sometimes that role is taken by e-research centres, as in the example of VeRSI as developers and hosts of the e-Resource Centre VRE. However, at least in some countries such centres are still rare and with the future of many of them not yet clear libraries will probably continue to play a key role in long-term preservation.

Our case study projects used a variety of methods for gathering requirements. In several cases, researchers already had a set of requirements before the project started. For instance, this was the case in the TextGrid project where several of the participating researchers had worked with digital tools for text editing and encoding before and thus had some ideas about what they wanted from the VRE. Surveys, focus groups, informal conversations, examples from existing VREs and more detailed analyses of the ways in which researchers work where also used. For the RIC VRE, for instance, Microsoft and the British Library consulted existing studies and com-

missioned their own research, especially looking to identify 'pain points' in the work of researchers where the right use of technology could make certain tasks easier. Especially where larger, institutional research infrastructure is concerned, requirements also come from the librarians. This was the case for the eSciDoc project where information specialists looked into building an underlying architecture on top of which specific services for researchers were then built. No matter how the requirements were gathered, though, all interviewees agreed that the needs and interests of researchers must be the driving factor for VRE development; otherwise, there is a clear danger that the project might fail because it would not be taken up. Mutually agreed project plans, clearly defined responsibilities and objectives are the basis of successful VRE projects. Funding programmes also stress the importance of this kind of collaboration, and the DFG even requires projects to bring together researchers and infrastructure developing organisations right from the start.

2.3.2 Development methodology and communities

The more researchers are involved both in the definition of requirements and in the actual development of a VRE, the more important iterative development methodologies become. All the projects we have considered used an iterative approach to a certain extent, as this method allowed them to let the researchers guide the development. Often prototypes or demonstrators were built to enable the researchers to give concrete feedback. In the RIC project, several closed beta groups were set up to provide feedback from different perspectives. For the e-Resource Centre VRE, researchers who attended the launch of a demonstrator were surveyed to provide further ideas and see what they were most interested in. In other cases, researchers took an active part in even the earliest stages of the development, although this is clearly limited to those with higher technical expertise and/or interest. It also should be noted that it may not always be easy to define when a development process has finished – in many cases development is ongoing and only limited by availability of resources. In these cases researchers will regularly provide feedback and make requests for further development. It was noted by interviewees that once an environment is used more often, researchers develop new ideas on what else they could do – which again suggests iterative development as the most suitable approach for building VREs.

As the construction of VREs appears to be a highly collaborative venture, it is unsurprising that development through a community is seen as increasingly important, especially in relation to larger projects. Funding bodies often require projects to encourage re-use, while few institutions have the knowledge and expertise to be able to work on all aspects of sometimes complicated development. For this reason, projects are actively working on developing communities around them; this is one of the stated goals of the RIC project which aims to develop a development community for VREs. It is important to remember that not only the research in VREs but also the development of VREs has an increasingly collaborative aspect and that development of new solutions can in itself be a research project. For examples on how to approach this, many projects turn to the open source development world and models that were developed there.

2.3.3 External applications

The open source world and freely available web tools and services also have another impact on VRE development. As both our survey and consultations indicate, many researchers use software created by open source community projects and freely available online tools. For example, Google Apps[16], including calendar, email and Google docs, is very popular because of its ease of use, free availability and the collaboration it enables without the need to set up or maintain any software. Zotero, a bibliographical tool developed by an academic open source project,[17] is another example of a popular application. Some projects, such as the VKS Collaboratory, made

16 http://www.google.com/apps/ [Accessed 18/12/2009].

the decision not to develop similar functionalities but rather to either accept or integrate the use of these widely used tools. Partly because of the interest researchers have in using tools such as Zotero, the SURFfoundation is also encouraging the development of VREs as open platforms that allow for easy integration of such tools.

2.3.4 Important roles in the development process

Apart from the software engineers, several key roles emerge in the development process and the uptake of VREs. The first is the (subject) librarian who in several projects stood out as a key interface between developers and users. As information specialists, librarians not only understand information and data management, they are also aware of the needs and interests of their respective subject communities. This gives them unique skills to act as interpreters between researchers and developers as they can, to a certain extent at least, speak the languages of both. They also act as ambassadors for digital technologies within their communities. Their role was for instance noted in relation to the eSciDoc project, where librarians in the respective Max Planck institute act as liaison to the central digital library unit. They also helped to identify researchers and groups that were interested in testing new technologies.

Of equally great importance are subject or institutional champions. These are usually senior academics who are well respected within their subject domains and who have a keen interest in the potential of new technologies. As this allows them to make the case for the use of ICT in research in a very convincing way and as insiders of the research community, their recommendations are of importance to other researchers. The TextGrid project, for instance, has profited from the engagement of senior academics who give the project academic credibility. Unfortunately, this kind of engagement may not always be seen as immediately beneficial to researchers as these projects can be considered as a detour for an academic career in fields that do not recognise engagement with research infrastructure in the same way as traditional research. However, at last some of our interviewees reported signs that this may slowly be changing.

A third key role is that of people like moderators or project managers who facilitate discussions within the VREs, or who negotiate between different stakeholders. These roles are often filled by people who are themselves members of the disciplinary domain. For example, Alzforum and related VREs all have moderators and editors who have backgrounds in the biosciences; similarly the collaboratories within HubLab have project managers who are themselves economic historians, and whose careers in economic history will be enhanced by being involved in the collaboratory.

It has also been noted that on-location support is very helpful, especially in the early phase of a project when researchers might still have difficulties not only in using the technology but also in articulating their experience. This can even apply to institutions whose staff has a fairly high level of computer literacy such as the VKS.

2.3.5 VREs as community development and research projects

Virtual Research Environments are designed to support collaborative research and working. In practice, the process of developing them is usually also a collaborative research process, that might cross institutions and perhaps even countries, but that certainly happens across subject domains. Librarians, computer scientists, support staff and researchers from at least one subject domain were or are involved in most of the projects we looked at. This means that just for delivering the VRE, projects would already have succeeded in creating a virtual research (and development) community.

17 http://www.zotero.org/ [Accessed 18/12/2009].

As some of the technologies used and more often the ways they are combined are new, this collaborative work often also has a research aspect to it. This turns VREs themselves into research outputs, especially when one considers the research undertaken in how researchers work and how they can best be supported. This makes VRE development more challenging, as if the pure technological aspects of the work were, at least in the context of some of the more ambitious projects, not big enough. Several interviewees remarked that this can lead to a tension between conducting research and delivering a stable platform. While in research 'failure' may be an acceptable outcome if it leads to new lessons learned, this would cause serious issues for a project that also needs to deliver a working VRE. It is important to keep this in mind when looking at VRE projects as this field is still at an early stage. Projects could potentially overcome this issue by providing both a test and a production environment, but this may require additional resources and a very structured approach to development that may not be suitable to all projects.

It is also important to keep in mind that some VREs have turned into large scale development projects, sometimes spanning more than a dozen institutions with many different work packages. One interviewee pointed out that universities and libraries are not necessarily always well equipped to lead such projects, and expressed hope that trustworthy commercial partners could be found that would supply the necessary experience. The interviewee felt that while some IT companies would have the necessary skills, they needed not only a certain understanding of academia, but also an interest to be involved in such projects not only for the sake of financial reward. Even where institutions such as libraries had the capacity and interest to set up development units, this in turn would have an impact on the institutions themselves. For instance, it could be argued that libraries will start to compete with publishing houses by developing VREs and publication platforms. Furthermore, to be able to participate in community development projects, especially in an open source context, the institutions might have to become more flexible and open.

2.4 Differences across the disciplines

When looking at the different ways different disciplines engage with VREs, it is important to consider that our study is not based on research that is representative of VREs with statistical significance. If, for instance, our survey were representative of the VRE domain, we would have to conclude that VREs offer services to disciplines in a fairly even way, with the exception of the arts and humanities that according to our survey are by far the best supported domain with regards to VREs:

What disciplines does the VRE support?	
Arts and Humanities:	45
Biotechnology and Biological Sciences:	29
Computer Science:	29
Engineering and Physical Sciences:	30
Economics and Social Sciences:	25
Medical Sciences:	28
Natural Sciences:	24
Other:	12

We cannot say to what extent this is actually representative of the disciplines supported by VREs. The large number of VREs in the arts and humanities can also be explained by the fact that our invitation to participate in the online survey drew the greatest number of respondents from the arts and humanities. However, at least some interviews indicate that there may be a particular interest in VREs from the arts and humanities. The 2008 and 2009 VRE calls of the DFG, for instance, had a response rate of more than 50% from the humanities as opposed to the sciences and other domains.

While no clear picture emerged about this, representatives from several funding bodies remarked that the sciences overall might be less likely to apply for such calls for two reasons: 1) They had better access to development support and digital infrastructure than the humanities, which means that they were in a better position to develop technical solutions through their own resources. 2) Scientists in general appear to be more computer literate than humanities scholars. Also, because of the way they have been working with IT infrastructure for decades, scientists are more used to interface with research infrastructure, for instance, by directly submitting jobs to the grid. As 'latecomers' who also do not (yet?) regularly work with very large datasets, the humanities may now be much more drawn to the more intuitive Web 2.0 interfaces that have emerged in the last few years and tend to be associated with VREs. In addition to that, at least one interviewee remarked that arts and humanities scholars and practitioners put more emphasis on face-to-face collaboration, which is why they might be more interested in collaboration within VREs that can be better personalised than the much more abstract research infrastructure.

2.5 International and cross-organisational research environments

Results from our survey confirm the general impression of VREs as places for international collaboration. It is striking that about two thirds of all the VREs supported international research, while only a third of them had a mostly national focus:

Does the Virtual Research Environment support mainly national or international collaboration?	
national:	30
international:	18
both equally:	41

This was also reflected in some of the interviews and case studies as even VRE projects that aim only to support a single institution often provide access to research partners in other countries. Not only may this indicate that VREs are indeed supporting the more international research of the future, the collaborative nature of many ventures does also become quite clear when looking at the high percentage of projects that involve more than one institution:

How many institutions are involved in the Virtual Research Environment?	
1 Institution:	15
2-5 Institutions:	32
6 or more institutions:	34

It must also be noted that VRE projects do not only support and involve academics, but partners from industry as well:

What type of institutions/project partners are involved?	
academic:	84
commercial:	29
government/public sector:	33

2.6 Technology

Judging from the survey, literature and interviews, the technologies subsumed under the 'Web 2.0' label currently generate the most interest in the VRE community. However, this does not mean to say that a clear preference for Web 2.0 approaches emerged or that they were used exclusively. Our survey shows that in current VREs a range of technologies are being used; this use mostly happens in a complementary way with different technologies supporting each other:

What technologies are being used for the Virtual Research Environment?	
cloud technology:	9
grid technology:	19
portal technology (for instance JSR 168/286):	32
repository management system:	27
publication technology (to publish materials directly):	19
services:	33
other:	19

The two most important technologies mentioned under 'other' were webhosting solutions and content management systems as well as 'semantic web' technologies that also attracted a lot of interest from our interviewees. The survey also showed that the use of open source software is much more prominent than that of proprietary solutions (49 vs. 13 responses) and that while a fairly large number of projects built their work on existing software (32), almost as many claimed to have developed the software they used from scratch (26). While this may only refer to key components, it illustrates how much pioneering work is being done not only in terms of combining existing solutions, but also relating to developing new software and services.

In terms of technological trends, we came across a variety of technologies and concepts that seemed of interest for the future. Web 2.0 technologies have already been mentioned, and they are seen as particularly interesting because of their ease of use and more lightweight nature. An obvious example for a more heavyweight architecture would be the Grid, an architecture that seems to be of particular importance in the sciences, especially physics, but is now also been used by the humanities (see the TextGrid case study), even though that use is very rare. The Grid offers authentication and, through it, secure access to storage and computing resources, something that will continue to be of importance. While there clearly is interest in the Grid, it seemed that generally speaking it had not been taken up as much as it was expected a few years ago, and other solutions are being looked at. For instance, there seemed to be a certain feeling that the Cloud may be able to provide similar functionality, perhaps in a somewhat reduced yet more easy to use way – it can be seen as the commercial, simplified version of the Grid. Generally speaking, there was no clear preference for any of these or other technologies, but an interest in integrating or combining technologies, such as adding Web 2.0 features to Grid architecture. Interestingly, the Dutch SURFnet has recently announced the 'Collaboration Infrastructure and Federated Collaboratories (CIFC) project that will investigate, among other issues, 'the ability to create an infrastructure platform that loosely couples collaboration services'.[18]

At least among our interviewees there was a consensus that as long as a technology worked in a certain context, it did not matter too much which particular technology it was: as long as it was interoperable, as far as needed. The Science Gateways, for instance, are not proponents of any one technology and different gateways exist; recommendations are only made when the choice of a software or technology creates problems for interoperability or other technical issues. Per-

18 SURFnet (2009), p. 8.

sonal preferences or the history of a project aside, interviewees were much more interested in what technology might offer and what would be needed than which specific solution was used:

> *The use of GRID does not seem to be getting up as fast as we thought a few years ago, and I think there might be other tools for sharing data, and for distributed computing. It could be cloud computing, it could be something completely different, but I think that it is difficult now to say that one of these ways of sharing and collaborating will be more important than others. But what we see is that data services that can handle large amounts of data, update this data and share this data will be in really great demand. Because that overflow of data that we see now is only just the start. So tools for accessing, sharing and curating data, I think, will be much more important than tools for sharing computational power.*

As these 'data services' become more important, a need also emerges for improving the searching and linking of resources. Especially in this context, semantic web applications are likely to be in increasing demand. As more data and other resources are digitised and deposited, and as more research activities take place in digital environments around these contexts, it will be necessary to be able to conduct intelligent searches in order to make best use of the resources available.

3 Challenges

3.1 Sustainability

A major challenge faced by all the VREs is that of sustainability. It was agreed by all that the enormous investment in VREs, both in terms of money and effort on the part of project teams and users would in most cases only be worthwhile if the VREs are seen to have long term support. Increasingly the data produced during research is seen in itself as an important outcome, and funding bodies want to obtain value for money from funded research by requiring that data be annotated and archived. Digital repositories need to become an integral part of national research strategies, in a similar way to other parts of e-infrastructures such as communication networks, high-performance computing, and distributed grids. Adonis is an example of a nationally directed strategy, including a common platform (ISIDORE) with which other funded projects need to be interoperable. This underscores the need to strike a balance between to tackle a sensitivity to the needs of specific communities, and the need for convergence, possibly through a common technology, such as a platform. In the case of Europe, this strategy crossed the boundaries of individual countries and extends to the whole European union with the European Roadmap for Research Infrastructures[19]. Thus the broader context of national and international strategy will strongly influence the discussion on sustainability.

It is difficult to determine how the development of research infrastructure can best be supported. Clearly, projects of two or three years duration (as is normal with research councils) are not seen as the right way to develop and then sustain a project that often only starts to become operational after at least a year or two. Funding councils are aware of this, but can be restricted by their institutional context, for instance through annual budgets that are not really suited for supporting projects in the longer term. Also, not all funding councils have the remit or the flexibility to support infrastructure in the longer term.

The key strategies for sustainability[20] adopted by projects and programmes are the following, either singly or, more often, in combinations:

3.1.1 Further funding

Several projects aimed to seek further funding beyond the formal end of the project from the research funding body that had initially funded them. In some cases, such as with funding from the Canadian Foundation for Innovation, a further four years of funding beyond the formal end of the project is in-built. The Cléo project has gone to the extent of declaring itself a 'precarious insider' in order to draw attention that action is needed in order to keep it going: that is, while appearing to be established inside academic institutions, it is in fact in a precarious position.

Funding of VREs is not entirely dependent on (national) research funding, but can also occur through institutions' budgets or in partnerships with industry. An example of the former are the projects initially jointly funded by the SURFfoundation and institutions, later to be funded entirely by institutions or the wider SURFshare (that is, Universities of Applied Sciences); and an example of the latter partnership is the RIC which involves a partnership between the British Library and Microsoft. As the VRE becomes increasingly integrated into the work of the institution, it will increasingly be a matter for institutions to sustain (see for example eSciDoc).

Just as VREs are often inter-institutional, so is their sustainability often an inter-institutional challenge, requiring a strategy that is bought into by clusters of institutions that are stakeholders in a VRE (see for example MemRE). A key role can be played by libraries in institutional strategies

19 ESFRI European Roadmap for Research Infrastructures, 2008.

20 For further models on sustainability see: Maron, Nancy L., Smith, K. Kirby, et Loy, Matthew (2009).

for sustaining VREs, since the digital outputs that they produce for future as well as current communities often require the expertise and infrastructure offered by libraries.

Results from the survey suggest that VRE projects are originally funded mostly by funding agencies or institutions' budgets; in the longer term, however, they are mostly sustained by institutional contribution and/or further grants.

How is/was the project development funded?	
research council or public sector funding agency:	60
industry or commercial:	5
institution's budget:	27
other:	12

(Under 'other' philanthropic contributions and/or the Mellon Foundation were most prominently listed as ways of funding a project.

How is the Virtual Research Environment sustained or how are you planning to fund it in the longer term (i.e. after the initial development)?	
institutional contribution:	43
volunteer effort:	18
further grants:	42
membership/usage fees:	7
other:	17

It should be added that a large percentage of those choosing the 'other' option had no clear idea yet about how the VRE may be sustained.

3.1.2 Business models

Several projects had developed business models to make VREs self-sustaining. In one case that we studied, the Orlando Project, the outcome of the VRE was an electronic publication which is now in the catalogue of Cambridge University Press, paid for by libraries and individual users. However, there are also models where the VRE is partially self-sustaining, with most functionalities being free, but some charged. Again, the largest remuneration would still be from libraries which would pay for particular functionalities. This is often a way of shifting the costs from individual VREs to institutions.

3.1.3 Community support

Ultimately the sustainability of VREs is dependent on their acceptance and use by the communities that they are intended for. The more VREs prove themselves to be indispensable for research activities, the more likely they will be seen as priorities for continued funding and support at the institutional level and ideally, also beyond. This reinforces the need to develop VREs in close collaboration with research communities, since they need in some ways to be self-built by research communities in order to ensure that different kinds of research can actually be supported by them.

Clusters of communities are as important as individual communities. If there are a number of different communities using a similar kinds of infrastructure, the cost of building and maintaining the infrastructure decreases over the long term. This is the approach adopted in the cluster of

projects related to the Science Collaboration Framework (Alzforum and related VREs), and can also be seen to be operating in myExperiment as an ultimate goal.

3.2 Barriers to the use of VREs

Our study has shown that there are quite a few very active virtual research communities out there. However, it is also important to consider some of the barriers that keep researchers from making more use of VREs – or, for that matter, keep them from engaging with VREs in the first place. In our survey we asked the question: 'What, if any, factors have kept the Virtual Research Environment [you are engaging with] from making a greater contribution to your research?' These were the answers that participants gave:

technology not reliable:	8
too difficult to use:	14
does not suit our research practice:	11
not enough technical support:	22
not enough institutional support for training etc.:	20
user community too small:	15
security/trust issues:	11
other:	36

'Other' lists a range of different issues, some being closely related to the options we gave participants. It was, for instance, remarked that it is not the size of the user community as such that was a problem, but the size of the active core group of users, those who would actively contribute rather than those who mostly followed activities. Several participants also mentioned that their VREs were hampered by a lack of funding, which restricted necessary development work, or that the environment was just too slow to be used effectively. Participants also noted that they had to use so many different systems that they had neither time nor interest in engaging with yet another platform.

Leaving aside the subject of sustainability, the following main barriers emerge from both the online survey and the interviews:

3.2.1 Lack of support

From both the interviews and the survey, a lack of support emerges as one of the most critical barriers to the uptake of VREs. There is, on the one hand, the ongoing technical support for fixing bugs and further development. On the other hand, direct support of researchers engaging with VREs is also a crucial issue. This applies especially during the early stages of a project when the researchers have to learn how to engage with a VRE. Dedicated on-location support was seen as critical in this regard, especially as researchers may lack the language to describe their issues in a way that a remote, purely technical support can understand. It is therefore important that support and training are provided by those who understand both the technology and the way in which researchers work. Local subject librarians would be one group with the necessary skills and some projects, for instance eSciDoc, make use of them as bridges between the development team and the users. Interactive online training or webinar type events can also be useful, especially with a larger user community for which the provision of on-location support would be too difficult.

3.2.2 Unsuitable for research practice

A smaller number (11) of the respondents to the survey said that the VRE they are engaging with was not ideally suited for their research practice. The impact that this can have on research-

ers varies, ranging from being a mild nuisance, to being a barrier so severe that it would make the VRE mostly useless. We did not come across cases of the second type in our study, although it must be assumed that researchers who decided not to engage with VREs because of this issue would also not be very likely to participate in a survey on the use of VREs. It would therefore be pure speculation to say how important this barrier actually is, apart from stating that it can be very severe. All interviewees were very aware of this barrier and all projects we looked at took care to engage the researchers not only in the gathering of requirements (if these did not actually come from the researchers themselves), but also in the development of the VRE. While such an approach cannot guarantee the uptake of a VRE, it seems to us that without it there is a high risk of failure.

3.2.3 Reliability of technology

Many VRE projects use cutting edge technologies in new and innovative ways. While this makes such projects interesting challenges from a developer's point of view, this can also lead to frustration among the users. We would assume that this is actually a bigger problem than the numbers in our survey indicate, as many of the researchers who responded are early adopters or were even involved in the development of a VRE from the start. Because of that it has to be assumed that they are, to a certain extent, used to working with technology that is not always reliable. The TextGrid project, for instance, is very aware of this problem and it was decided not to put too much priority on increasing the user community before the project has matured (both in terms of stability and functionality) to a stage where it is ready for use by a larger, less computer literate group of researchers.

3.2.4 Critical mass of active users

A key issue for any virtual community, not just those engaging in research, is to encourage enough contribution to make it seem interesting enough for others to join and/or engage with that community. While this is not so much of a problem for VREs that are focussed on providing access to services or tools, it is important for those that rely on either a contribution of content/data or an engagement in discussion and collaborative activities. Without enough participation, such VREs will soon be seen as inactive and face the risk that researchers turn away. In order to provide better environments, it is important for VRE developers and managers to gain an understanding of both those researchers who do engage with the VRE and those who do not, through for example talking to them and analysing literature citations. Getting the critical mass of users together is also important for the sustainability of a VRE, as further funding will be very difficult to secure without visible uptake in the research community.

3.2.5 Legal and ethical issues

Privacy concerns and lack of understanding of copyright and copyright compliance are two of the major challenges listed in the literature[21] for the success of a VRE. However, this seemed to depend on the type of VRE and the type of data accessed. For example, any medical data must meet ethical and legal standards (see for example Biogrid) and hosting medical data is often avoided (see for example Science Gateways). In the humanities and social sciences, authorship and copyright issues are more common than ethical and legal issues. However, on the whole, it did not seem that legal issues relating to IPR were experienced as overly problematic. The experience of the DFG VRE programme, for instance, suggests that intellectual property rights (IPR) would be more of an issue with long term projects, for instance when projects get passed on to other institutions etc. Ownership of content/data that are collaboratively shared and ac-

21 For example see Silipigni Connaway, L. et Dickey, T.J. (October 6, 2009).

cessed can also be an issue, needing to be managed by auditing and archiving policies – a discussion that the members of the e-Resource Centre VRE are currently engaged in. Similar to the DFG findings, ownership becomes problematic if a VRE is transitioned to another hosting/development organisation or is decommissioned. Digitisation of content remains problematic, although this is not a VRE-specific problem. In a VRE with international collaborators it can be problematic to share resources for which one institution has a subscription and the other does not.

In at least one country, that is South Africa, there is a law preventing data that is produced by research funded by the state from being shared beyond the borders of the country without prior permission. However, researchers are not always sure to which geographical location they are actually moving data. The question of geographical location is also important because of the legal jurisdiction under which data falls. There can also be questions regarding what can be done with data that originated in one country, but are stored on the database of another. For example, the HubLab project had the experience of not being permitted to upgrade data that had originated elsewhere. The NMIMR project similarly has experienced difficulties regarding storing of data generated in different countries. Geographical location is also important when sharing data relating to people. For example, the European Directive on Data Protection prohibits sharing data with countries which do not have the similar measures for data protection as Europe, which in some case would make it illegal to share data with the USA.[22] Anthropological data can often raise difficult ethical and legal questions.

Thus the major legal issues relate to international collaboration. It is often difficult for project teams to even know whether what they are doing is strictly legal. Further clarification is needed; however legal barriers to sharing and storing data in different countries could be restrictive for VREs wishing to conduct international collaborations, and would need to be addressed in advance.

3.2.6 Interdisciplinarity and different ways of working

Working across disciplines can create several issues, mostly relating to different languages spoken, domain specific ontologies and different ways of working. Sometimes apparently simple tasks can appear in a very different light if seen from the perspective of a different discipline. The VKS Collaboratory, for instance, has a module that is meant to support the writing and sharing of draft documents. It became apparent that 'working on a draft' could relate to very different activities: in some disciplines it is a way of letting your colleagues know that a document is mostly ready for publication, while in others writing and discussing a draft document is seen as part of a highly reflective process during which some of the research findings will be constituted. In an interdisciplinary environment, it is important to look at both the ways researchers work and communicate in different disciplines.

22 Carusi & Jirotka, 2009.

3.3 Desirables

3.3.1 Awareness raising

While we have found signs that researchers are becoming more aware of some of the benefits of using VREs or digital research infrastructure, much remains to be done in order to enable researchers not only engage to with existing infrastructure, but also to formulate their requirements. For instance, while researchers increasingly seem to realise the benefits of using repositories in larger or geographically dispersed research teams, they do not yet appear to be fully aware of the difficulties of digital preservation and curation. Requirements to do this from a funding agency are often successful, though training must be offered in how to do it correctly. Although awareness varies across subject domains (with the arts and humanities overall appearing to be somewhat less computer literate than the sciences), the benefits of VREs need to be communicated better and more training needs to be provided. Given the importance of community acceptance in order to ensure the sustainability of projects, there is a clear need for advocacy, publicity, marketing and promotion of the systems and for making scholars aware of how the systems and tools can simplify their workflows and more widely disseminate their work. This was clear from several of the case studies (for example, HubLab) and from the literature[23].

3.3.2 An international VRE forum

Those engaging in building and researching VREs are very interested in establishing a dialogue that goes beyond platform specific communities. The Science Gateways, for instance, have established a forum for VRE developers which is considered a valuable aspect of the programme in its own right. However, as more research collaboration happens in an international context, it is becoming more important to discuss VREs in an international forum that also includes funding bodies.

The main aim of such an (inter)national dialogue could be to develop joint strategies to guide investment and research in VREs and identify areas in which funding bodies can develop joint programmes and initiatives. While it may not be realistic to assume that this process would lead to the establishment of a few generic VRE platforms, ongoing dialogue could help to prevent the constant reinvention of the wheel in some fields and ensure a more seamless interoperability of emerging national infrastructures. Researchers involved in VRE projects pointed out the importance of building on the findings of previous VRE and digital repository projects. This would involve linking to, expanding and improving already-developed systems instead of reinventing new systems from scratch. Such a forum could make an important contribution to further development and uptake of VREs, and it can only be hoped that the establishment of the Knowledge Exchange, a group of four European funding bodies with an interest in VREs, will be a first step in that direction.[24]

3.3.3 National and international integration of resources

With an increase in (inter)national research collaboration – a trend actively encouraged by funding bodies and policy makers – there is also an increasing interest in, and need for access to, resources across institutions and state borders. Research teams from different countries collaborating in a virtual environment will currently have to accept that they cannot all access the same resources because some members of the team may have a license for a database through their host institution that other partners do not have. Some countries have national licenses for ac-

23 For example, Silipigni Connaway, L. et Dickey, T.J. (October 6, 2009).

24 http://www.knowledge-exchange.info/ [Accessed 04/01/2010].

cess to a wide range of digital resources, but the problem can even affect researchers in a single institution. It also has to be noted that laws regarding IPR and various regulations about access to data and licensing can be different across countries, which adds to the complexity. Even within just one country it can be problematic. The VKS, for instance, is a distributed group with members in different cities in the Netherlands, attached to different host institutions, and despite using a VRE these researchers cannot all access the same data resources.

Precisely because VREs have the technical capacity to bring together various resources in one environment they also draw attention to this issue. Researchers, VRE developers and even representatives of funding bodies see this as a problem that needs to be addressed on an international level in order to fully enable the potential of research collaboration. One interviewee commented from a European perspective:

> *A critical question for Europe is to be able to share resources across national boundaries. And it seems to be very difficult to have, say, one country invest in this type of infrastructure in another country, and then share the use of it. For the user it simply should not matter where the resources really are located. But for the national government and the national centres, it seems to be difficult to invest in other countries. And that means that there's a scale problem in Europe, because basically, there are lots of small countries. We will have lots of smaller infrastructures that need to be tied in together in an accessible way. The tools are one thing, but the policies for allowing this are a very critical point.*

Integrating different resources, ideally across disciplines and countries, is a big challenge. However, it is also seen as having enormous potential for enabling new research and finding new answers, as has been discussed before. To unleash that potential access to data resources is crucial.

3.3.4 Authentication and single sign-on

Controlling access to resources can be in the interest of both institutions and individual researchers and is sometimes a legal necessity. Trust can be increased by knowing who has access to data and also by knowing that the person you communicate with is actually who you believe they are. Providing secure and reliable authentication is of the utmost importance in this context. While single sign-on technology, such as Shibboleth, exists that can facilitate that, the practice of implementing shared access across institutions is still a big issue on both practical and policy levels. The e-Resource Centre VRE, for instance, needed to implement a Shibboleth IdP for its university partners who are part of the Australian Access Federation, and an OpenIdP for the government researchers (DPI) who are not part of the Federation but who also require access to e-RC). More granular requirements that may need to be achieved through authorisation mechanisms (for example in limiting access to certain types of data by role) is yet another level of complexity for VREs.

3.3.5 Usability

There was a clear consensus among the interviewees that ease of use is one of the most important aspects of developing a VRE, an observation that is also confirmed by the literature. Some scholars are reluctant to use new technologies not because they are not interested in them, but because it appears to be difficult to learn new systems and processes – especially with a feeling that there is not even enough time for core research tasks. A simple interface and user-friendly tools are high on the list of researchers' desirables. Support through workshops and demonstrations, more technical support, and greater institutional support for training are also needed. In order to support collaborative and cooperative activities, it is important that virtual en-

vironments offer the means to access appropriate information as well as communication. The literature raises the argument that current technologies and implementations do not adequately support the key concepts of communication and community[25]. An interviewee stresses the point that

> *the goal of VREs really ought to be, how do we make the technology work for the researchers in such a way that they are spending less time fussing with the technology and more just doing their research?*

A key aspect of usability in VRE design is the art of reducing the many options that powerful software frameworks potentially can offer to those features that the user community actually needs. Experiences from projects clearly showed that, maybe somewhat counter-intuitively, fewer features mean more usage. The VKS, for instance, saw deciding which SharePoint features not to use as a key part of the work on their internal collaboratory, and another interviewee remarked:

> *People tend to get confused very quickly with elaborate software like SharePoint. And once they have been confused once or twice they just drop using the tool.*

In terms of designing a VRE, the same person also had a very clear guideline: 'What we tried to design is a tool with as few buttons as possible.'

3.3.6 Integration of Web 2.0 technologies with infrastructure

The participants in the survey as well as our interviewees suggested that the development of collaborative tools and platforms should occur in the light of informed understanding of Web 2.0 technologies as well as social applications (e.g. SlideShare, Flickr, Connotea, Facebook). It seems important to move away from monolithic portals towards more lightweight and adaptable solutions, such as Web 2.0 solutions. While there were several projects that built on repository (Fedora; Dspace) and/or Grid technology, there are at least as many, if not more, that used more light weight Web 2.0 technology such as wikis and blogs, which are much easier to set up and customise. However this did depend on the disciplinary domain and type of research activity. For example, it is difficult to use a cloud technology approach for extremely large data sets (at the peta-scale) and impossible for supercomputing; in addition, cloud computing brings its own legal and regulatory risks for data which need to be secured[26].

There are also a few criticisms of aspects of Web 2.0 technologies. An interviewee who works on developing research infrastructure for a larger organisation raised the matter that Web 2.0 is not considered trustworthy enough by their researchers who also consider several of its applications to be 'just a toy'. However, they do use Wikis and blogs. Another interviewee reported similar reactions as

> *e-researchers cringe a little bit when we start talking about Facebook for researchers. That's not the way they think about finding their colleagues. [...] From a technology perspective, it's very simple to build a Facebook app that feeds some of the data out and into Facebook. From a user-acceptance point of view, however, there's not a real crisp scenario that we have heard yet from the researchers that they really want this.*

While this could be seen as a certain scepticism against combining social networking and VREs, Web 2.0 appears to be the most interesting technology in relation to VREs. With a growing number of students and young researchers using applications such as Facebook and other Web 2.0

25 Redfern, S. et Naughton, N. (2002).

26 See for example, Christopher Millard (2008) *What's all the fuss about cloud computing?* http://www.bristows.com/?pid=46&nid=1203&level=2. [Accessed 18/12/2009].

platforms and techniques for sharing information and for communication, it may be assumed that there will be increasing expectations among the next generation of researchers that their VRE tools have a similar ease of use.

3.3.7 Common standards

There is a clear need to create standards for the context associated with data, including the data format, the semantic meaning of descriptive metadata, and the procedures for parsing the data. The standards represent the consensus of the community that wants to promote interoperability and broader use of the data. On the level that is closest to conducting research with the data, without formatting and metadata standards, deposited data are very difficult to interpret by other researchers. Even though the view was often expressed that this is particularly difficult in the humanities where there is no tradition of creating standards and metadata, in fact this is an issue throughout the disciplines considered. There is a very wide range of approaches to going about implementing standards and metadata. At one end of the spectrum, researchers gather in face-to-face workshops to work on a set of standards together and create a final version. An example is the Dublin Core descriptive metadata used in the digital library community. At the other end of the spectrum, an evolutionary approach is used in which the consensus is represented by the set of policies and procedures that are actively used to manage the collection. The policies are turned into computer actionable rules, and the procedures are composed from computer executable functions.

Both the policies and procedures can be modified over time to improve the ability of the data management system to enforce the desired properties of the data collection. An example is the integrated Rule Oriented Data System (iRODS), developed by the DICE group led by Reagan Moore. The policies in question can relate to any aspect of data that is important to researchers (access controls, required descriptive metadata, retention, disposition, distribution, replication). The policies are enforced on data deposition into the collection, on data access, and on every operation performed upon the data. Reagan Moore's experience with data sharing, first with the Storage Resource Broker data grid, and now with IRODS, showed the importance of a consensus regarding data standards that evolves over time. The policies used to share data within a project can rely upon knowledge shared by project members. When the data are published in a digital library, a richer context is required that enables non-project members to understand the data. When the data are organized into a reference collection for preservation, the context needs to be understood by future researchers. At this point the standards become those of the discipline rather than of the project. For humanities research, for example, to adopt a similar policy-led data management system would enable development of policy sets that simplify education, or that promote collaborative research, or that promote publication of digital collection analyses. In education, students learn standards for data organization, interpretation, and analysis. In collaborations, the project members develop a consensus on a unifying context that transcends personal collections. In publication, standards for a discipline are applied for classifying data.

The creation of data standards also needs to meet the challenges posed by a variety of different modalities of data. The survey showed that even though the majority of data are still textual, there is also an increasing amount of data in different modalities:

If you access/share data through the Virtual Research Environment, in what format are they?	
audio:	21
textual:	75
images:	56
other:[27]	19

Because libraries are important partners in many VRE projects, particularly with respect to sustainability, metadata issues will gain even more prominence. Librarians are typically concerned with future users of data, whereas researchers are often more concerned with the immediate project of research. Thus it is to be expected that librarians will continue to play an important role to play in the standardisation of metadata.

3.3.8 Trust and social design

Both the literature and the results from the survey agree that crucial elements for the success of a VRE project are mutual trust and respect among the researchers, personal compatibility and good personal relationships. Effective communication, transparency and clarity are at the basis of a positive collaboration. Trust is also at the basis of data sharing, since data can only be interpreted correctly in a climate of trust regarding the history of the data: that is, its provenance, the methods involved in producing it, and a lot of other information. It is also often important for researchers to be able to interact in an interpersonal way in order for the data to be shared in the most fruitful ways. To work optimally VREs need to provide a rich context in which data are trustable.

27 Most examples given here do actually fall under the other categories, such as 'PDF' or 'bibliographic data', although in a few cases 'binary data' and software were also shared/accessed.

4 Conclusion: key findings

At the outset of this study, we were expecting to conclude with a set of findings distinguishable into ‘people’, ‘technologies’ and ‘institutions’. Instead what we have found is that these three elements are very closely inter-related, to the point where it is very difficult to try to take them apart. Moreover, it is not helpful to try to take them apart if we are to consider strategies for taking forward VREs so that they can fulfil their potential as research enablers. The strategy for VREs needs to be able to recognise and work with the integration of these three elements.

That said, there are perhaps particular recommendations that can be made to the different stakeholders in VREs so that they can better operate in these highly integrated systems.

- While many **different definitions, terms and concepts** relating to VREs exist, the differences do not appear to be of interest to researchers as long as the offered solution is seen as useful as part of a research process.
- The line between **portals, digital libraries and VREs** is becoming increasingly blurred as the former add means to share data and collaborate.
- Potentially the most important trend identified by this study is an increasing focus on providing **general VRE frameworks** that can be used to develop and host different VREs. The frameworks would provide core services (such as authentication and rights management; repositories; project planning, collaboration and communication tools) and allow the development or easy integration of modules for specific uses.
- Even though there clearly is an interest in VRE frameworks, there is also a consensus that a **one-size-fits-all approach to VREs will not work**; researchers' needs, even within the same discipline, are too different. General environments could be used, but they needed to be flexible to adjust them to specific needs.
- VREs need to be developed with the **researchers' needs** in mind and have to be seen as contributing to the research process; otherwise there is a danger that they will not be taken up by the research community.
- VRE development should be taken forward based on researchers' requirements and as a constant **dialogue between researchers and developers**.
- As integrating feedback from an early stage is of such high importance for developing a VRE that meets researchers' needs, **iterative design methodologies** seem particularly suitable.
- A key for developing VREs is a strong **focus on usability**, which can often mean not implementing all possible functionality, but focusing on delivering the most important parts so that they can be used in the easiest way possible.
- Despite signs that researchers may become more interested in virtual research collaboration, **awareness raising** emerged as an important desirable in order to better enable researchers to formulate their requirements and understand how they can benefit from VREs.
- **Accessing and sharing data** emerged as the aspect of VREs that researchers were most interested in, especially if this was set up in a way so that they did not have to give up control of their data.
- The most problematic aspect of VREs appears to be the **unresolved question of sustainability**. While funding for developing VREs is often available, their survival often depends on the success of projects to raise further funding or the ability of institutions to

support them. This may be a further argument for general VRE frameworks that could be supported by larger consortia, thus reducing the cost for each participating institution. However, this does not address the issue financing a national and institutional research infrastructure.

- There is a need to **develop and coordinate policies and structures relating to research infrastructure**, both on a national and international level. In a world that increasingly encourages and demands international research collaboration, the promise of VREs can only be fulfilled by allowing researchers to access data and services across institutions and borders.
- Other countries look to the **UK as an international leader in VRE development**; this is at least partly seen as an effect of the JISC VRE Programme.
- VREs are seen as having a key role in **facilitating a new type of research** that is highly international, interdisciplinary and that relies on distributed data.
- Even those researchers heavily involved in a particular VRE still rely on **external (web) applications** and are not likely to give up both general and specific tools such as Google Apps and Zotero that they perceive as very practical solutions; VREs should aim to integrate such tools and only develop competing solutions if there is a very good reason.
- **Google Apps** (including email, calendar, docs etc.) appears as a kind of lightweight VRE that supports basic requirements such as communication, shared editing, project planning and sharing of documents. The software is especially popular as it requires no set-up, is easy to use and allows collaboration across institutions, without the need to deal with regulations or applying for web space or central support. It will be interesting to see what impact Google Wave will have on online collaboration, but it is currently too early to tell.
- Research-oriented **libraries get more involved with VREs**, for a variety of reasons: 1) VREs are seen as a good way to feature library resources, especially domain specific ones that researchers are not always aware of; 2) The more libraries are concerned with curating digital research output, the more they realise that the best way to make that task easier is to be involved in the creation of these materials from the start; 3) Libraries who see themselves as supporting the whole research lifecycle want to better understand the impact of VREs on that process.As libraries have always been places of communication and collaboration it may be seen as a logical consequence that **digital libraries become more like VREs**.
- **Web 2.0 technologies** generate a lot of interest in the VRE community, especially as they are seen as suitable for rapid development and deployment, easy to use and geared towards collaboration.
- Generally **not one particular technology** emerged as more important than any other; the consensus seemed to be to select the best tool for a particular task.
- There is, however, a certain sense that **not all promises of the GRID** may have been realised and that Grids can be too complicated for certain tasks or too difficult to implement. In some cases, combining the GRID and Web 2.0 features was seen as promising.
- There is also a noticeable **interest in Cloud Computing**, despite some concerns about losing control over how and where data is stored; this can be a critical issue for projects looking after medical and other sensitive data. It should be noted that there are not yet many examples of Cloud Computing being used in VREs.
- The **integration of semantic web technologies** into VREs can be helpful to make it easier to find and interlink information.

- **Social networking** approaches generated mixed responses; while there clearly are very useful implementations, some researchers appear to reduce them to 'Facebook for researchers', which may give the impression that it is just a toy. However, this may be more an issue of how certain approaches are presented.
- Some VRE projects turn into **very large development projects**; libraries and universities may find it necessary to change their structures to be able to deal with those and/or find commercial partners.
- A key factor for the uptake of VREs is on-location support, ideally provided by people who understand both the technology and the researchers well enough to be able to act as interpreters.
- Another important factor for the uptake of VREs is the involvement of **academic or institutional champions**, who promote VREs or e-research in general and who are well respected in their field or institutions.
- The **provision of training** was seen as another key factor for the uptake of VREs.
- VREs are supporting both synchronous and asynchronous communication; there was, however, a clear demand for the integration of easy to use, robust and scalable **video conferencing in VREs**.
- Several interviewees expressed their hope that funding bodies would concentrate more on supporting and expanding **existing solutions** with a wider user base instead of constantly chasing new ideas that could not be supported. While this can be seen as an argument for general VRE frameworks, there was also a request for focussing on more limited projects that maybe only addressed one issue, but did that really well.

5 Case Studies

5.1 Programmes and countries

5.1.1 DFG VRE Programme

Region	Europe (Germany)
Subject domain(s)	All subject disciplines
Project started	2008
Website	http://www.dfg.de

Introduction

The Deutsche Forschungsgemeinschaft (DFG – German Research Foundation) is the central, self-governing research funding organisation that promotes research at universities and other publicly financed research institutions in Germany. The DFG serves all branches of science and the humanities by funding research projects and facilitating cooperation among researchers. The DFG has a long-standing interest in virtual research collaboration and started its first related programme in 2000 ('Themenorientierte Informationsnetze', issue-focussed information networks). In 2008 and 2009 each, the DFG issued specific VRE calls ('Virtuelle Forschungsumgebungen').

Origin and motivations

The idea behind the VRE calls is to support collaborative working across disciplines and over the whole research lifecycle, from collecting and sharing of primary data to analysis, publication and preservation. VREs are seen as essential for the support of the growing number of geographically and nationally dispersed, interdisciplinary research groups. In addition to support for this ongoing transformation of the research landscape and the networking that is part of it, the DFG also sees VREs as beneficial as they can increase the speed of publication of preliminary research findings and facilitate preservation of project data through the use of shared repositories. Most importantly, however, it is the DFG's view that Virtual Research Environments can enable new research and thus increase the quality of research outputs.

Features and technology

The DFG does not feature any particular technology through its VRE calls as it sees the technology as a means to support a good concept. As different technologies can be suitable for different research questions, the calls are non-specific with regards to what approach will be used. However, the DFG wants new software development to follow open source principles and to demonstrate an awareness of the state of the art and relevant standards. The VRE calls particularly support the development of infrastructure and testbeds, with a view that research conducted as part of a VRE project would mostly be for testing and using the infrastructure. Two types of projects can be funded, development projects ('Entwicklungsprojekte') that develop something new and transfer projects ('Transferprojekte') that apply existing solutions. Because of the diversity of the funded projects, it is difficult to identify clear trends in relation to technology. The eSciDoc platform was quite prominent in the latest call and quite a few projects used repositories and grid architecture. However, at least as many projects seem to focus on more lightweight Web 2.0 technologies – if any then this is the trend that emerges from looking at the applications and projects.

User community and development

A requirement for bids is that projects have to be a collaboration between researchers and infrastructure developing institutions (libraries, computer centres, e-research centres). An ongoing dialogue between both is seen as a precondition to build VREs that are both technically sound and fit for the intended research purpose. With collaboration at the heart of the call, the DFG welcomes interdisciplinary and international projects. It was noted that both in 2008 and 2009 slightly more applications came from the humanities than from the sciences; this may have been because the humanities may have less technical resources available and would thus have to rely more on external money for VRE development projects. In the 2008 round, 15 bids were received, six of which were subsequently funded. 2009 the number of applications had increased significantly to 48; the bids are currently being reviewed. The increase in applications may to a certain extent be due to an increased focus on outreach activities: the DFG had promoted the 2009 call mostly amongst the research community, while the 2008 call was directed more towards institutions. There are still signs for a growing interest in VREs amongst researchers, especially compared to the programme launched in 2000. The DFG also sees an increasing demand from researchers to provide shared data storage through a virtual environment, which is why special funding for DFG supported research groups (SFBs) is available for developing such environments to be used right from the project start. While researchers may be more aware of the use of technologies for sharing data, only a comparatively small group seems to think beyond this, especially in the humanities.

Ethical and legal issues

The DFG has so far not been made aware of any major legal or ethical barriers impacting on its VRE programme. IPR can be an issue, but this mostly relates to digitisation of content and is not VRE specific.

Future plans

Part of the future plans of the DFG is to consider whether it should move away from the very generic and open call to a more focussed approach to make the development more coherent and increase interoperability. Different approaches are currently being considered, including a focus on research methods or preservation. However, these are just preliminary considerations and the DFG also realises that VRE development is still in early stages and may benefit from a broader approach. Another intention of the DFG is to increase VRE-related collaboration between funders, the various German states and the large research organisations such as the Leibniz or Helmholtz societies, for instance through the Priority Initiative 'Digital Information' of the Alliance of German Science Organisations.[28]

Sustainability

Project host institutions are required to ensure that projects outputs are sustained, but the DFG usually does not outline a minimum time period. Libraries and computing centres at or related to the host institutions have a crucial role for preserving research outputs, especially as there is not yet a national preservation strategy.

Lessons learned:

- VRE projects need to be driven by researchers' needs and research questions
- a dialogue between infrastructure specialists and researchers is absolutely necessary for a successful project and should start with the development of a project bid
- using VREs can contribute to the preservation of research data and outputs

28 http://www.allianz-initiative.de [Accessed 18/12/2009].

- VREs have the potential to facilitate new research findings, not only to support existing research
- researchers are more aware of the importance of preservation and the benefits of shared access to repositories (no matter whether they contain research data, publications or other materials)
- awareness raising is still very important
- there seems to be an increasing interest in lightweight, easy to use solutions
- usability is key for the uptake of a VRE
- the technology behind a VRE it is not of particular interest to most researchers, as long as it does the job

5.1.2 ReInfra

Region	Europe (Norway)
Main project partners	Norwegian Research Council, University of Oslo, University of Bergen, Norwegian University of Science and Technology, University of Tromsø, Meteorological Institute and Uninett Sigma
Subject domain(s)	All subject disciplines
Technologies	Grid, HPC
Project started	2008
Funded by	Norwegian Research Council
Website	http://www.forskningsradet.no/english (website of the research council)

Introduction

ReInfra, the Committee for Investments in eInfrastructure, is an advisory committee set up by senior Norwegian scientists, to establish an e-Infrastructure strategy for Norway. The Norwegian Research Council is the overall funder, although the Committee has worked across various funding streams in the research council, notably with NOTUR which is a project dedicated to computational infrastructure[29], UNINETT Sigma, which coordinates the operational activity and long-term development of the national infrastructure for computational science[30], and NORGRID, which looks after Norwegian Grid Initiative[31].

Origin and motivation of the project

Infrastructure strategy is seen as an extremely important part of Norway's ability to maintain its place in research, and for it to continue to be a player in European initiatives, as well as other regional initiatives, such as those involving Nordic countries. So far, the high energy physics research community has made the most use of the infrastructure. A main driver for the infrastructure strategy is to enable this community to access and analyse data generated by the Large Hadron Collider at CERN. In recent years, this has included a Virtual Research Organisation the Nordic Data Grid Facility[32], which is a collaboration of the four Nordic countries (Norway, Sweden, Denmark and Finland). However, there is a commitment to opening up the infrastructure to other disciplines.

A further important driver is the European Strategy for Research Infrastructures, exclusion from which would be disadvantageous to Norwegian research, even if Norway is not a part of the European Union.

Features and technology

The infrastructure strategy applies to:

- Hardware including operations: high end computational resources, storage facilities, high-speed network
- Software, including: system software and basic tools, application software, grid middleware
- Support: basic help-desk support, advanced user support

29 http://www.notur.no/ [Accessed 18/12/2009].

30 http://sigma.uninett.no/index.en.html [Accessed 18/12/2009].

31 http://www.norgrid.no / [Accessed 18/12/2009].

32 http://www.ndgf.org/ndgfweb/home.html [Accessed 18/12/2009].

- Services for: end-user functionality, performance guarantees, quality assurance.[33]

It did not seem that a specific VRE strategy had been included in the infrastructure strategy. However, this does not mean that there are not VRE-like developments – the Nordic Data Grid Facility, which sees itself as building Virtual Organisations, has already put in place initiatives to create VOs in the bioinformatics and climate research communities. These have features that are closely related to VREs.

User groups

Apart from the high energy physics community, a user group in computational chemistry is in the early stages. Biosciences, medical sciences and other groups, such as linguistics, also have projects starting up.

Ethical, legal and institutional issues

While no major ethical and legal issues had been encountered so far, it was felt that there are likely to be issues emerging when the resources are extended to medical data. For Europe, a major issue will be making data available across national boundaries.

Future plans and sustainability

The strategy is to extend and develop further the Norwegian e-infrastructure, both for current user communities, and for new user communities across other disciplines.

The National Computing Infrastructure project was initially funded for ten years, but a more long term strategy is recognized as necessary, and the ReInfra strategy group has tried to take this project forward. Like other research councils, the demands on the Norwegian Research Council change from year to year, and it is difficult to predict how much it will be possible to set aside for sustaining infrastructure. Because of the annual budgets allocated to national research funding, it is difficult to plan ahead. The project has identified two main strategy lines, depending on funding which is made available: a low budget scenario and a high budget scenario. In fact, only in the high budget scenario will the development of the e-infrastructure be at reasonable levels.

33 ReInfra Committee, October 2008. 'Investment in e-infrastructure for computational science: An investment plan for the period 2007-2017'. http://www.forskningsradet.no/publikasjoner [Accessed 18/12/2009].

5.1.3 Science Gateways

Region	North America (USA)
Main project partners	Teragrid, incorporating eleven partners
Subject domain(s)	Cross disciplinary, but mostly natural and physical sciences
Technologies	Grid; Supercomputing
Project started	2004
Funded by	NSF
Website	http://www.teragrid.org/gateways/

Introduction

Science Gateways aims to facilitate the use of Teragrid resources by scientists. Funded by the National Science Foundation, the anticipated outcome of the Science Gateways is a greater take-up of TeraGrid and of High Performance Computing. There are eleven partners in the TeraGrid: Indiana University Research Technologies Division, (LONI) Louisiana Optical Network Initiative, NCAR (National Center for Atmospheric Research), NCSA (National Center for Supercomputing Applications), NICS (National Institute for Computational Sciences), ORNL (Oak Ridge National Laboratory), PSC (Pittsburgh Supercomputing Center), Purdue University, SNDC (San Diego Supercomputer Centre), TACC (Texas Advanced Computing Centre), Argonne National Laboratory.

Origin and motivation of the project

The idea behind Science Gateways initially came at the inception of the TeraGrid project from then-director Rick Stevens of the Argonne National Laboratory. Rick observed that science was increasingly going digital and scientists were designing their own interfaces to digital data and analysis capabilities. Supercomputing could be made more relevant to these researchers by 'clip[ping] it on the backend of interfaces that researchers were designing for themselves'.[34]

Features and technology

New projects applying for a Science Gateway are going through an open peer-reviewed process of evaluation. Currently there are 35 Gateways, which use TeraGrid as well as other resources, including some Cloud-based ones. The Gateways are mostly in the natural and physical sciences, with only one in the social sciences. Most of the projects use computing resources rather than applications geared towards data management, although they are also starting to see more interest in making data available (for example, in expensive petascale simulation).

The Science Gateways programme is non-prescriptive about technology or software; they use a bottom-up and user-driven approach, with the different Gateway projects designing their environments according to the needs of the community they are serving. The kinds of resources that can be accessed and used via a Science Gateway include: workflows; visualisation software and hardware; access to data collections; data analysis and movement tools; resource discovery; and job execution services.

The website lists the following three instantiations of the Gateway as being the most common:

- Web portal: The user interface is a Web browser-based application with users in front and TeraGrid services behind

34 Interview, Nancy Wilkins-Diehr.

- Desktop application: The interface is an application or suite of applications that run directly on users' machines and that accesses TeraGrid services.
- Grid-bridging gateway: Some communities run their own grids that are devoted to their areas of science. In these cases, the gateway is a mechanism to extend the reach of their community grid so its users can also use the resources of the TeraGrid.[35]

The interface to the TeraGrid is controlled by the different Gateway projects.

Ethical and legal issues

The different Gateways all institute their own security measures, and the TeraGrid is not responsible for security. Gateways must comply with the institutional requirements of TeraGrid's resource providers. For example at most sites, HIPPA (Health Insurance Portability and Accountability Act of 1996) data cannot hosted by the TeraGrid since this requires special certifications. Only anonymised data can be hosted. Gateways develop their own access restrictions, but many impose few restrictions. 'Open scientific environments are recognized as being beneficial for all.'[36]

Sustainability

The impact of Science Gateways will only be felt in the long term, and therefore they need long-term commitment. Sustainability is thus seen as being the most important and urgent challenge faced by the Science Gateways. There is a project starting which will look into sustainability options, including considering which metrics to use in order to understand which projects deserve long-term funding.

Lessons learned

- There is an increasing interest among researchers to make more data available.
- VREs can be a useful interface to supercomputing resources.
- It should be left to the specific communities to decide which technologies suit them best, as long as they are interoperable.
- Sustainability is a key consideration for developing any part of research infrastructure.

35 http://www.teragrid.org/gateways/ [Accessed 18/12/2009].

36 Interview, Nancy Wilkins-Diehr.

5.1.4 SURFshare

Region	Europe (Netherlands)
Main project partners	Royal Netherlands Academy for Arts and Sciences, Netherlands Organisation for Scientific Research, and most major higher education institutions in the Netherlands (including research and applied science universities)
Subject domain(s)	Cross-disciplinary but mostly arts and humanities and social sciences
Project started	2007
Funded by	50% Government funding, 50% institutions
Website	http://www.surffoundation.nl

1) Introduction

SURF is a major organisation in the Netherlands promoting collaboration among the higher education institutions on matters relating to ICT for education and research. SURFfoundation is the main user facing division of SURF, whereas SURFnet is the main technical and development division. SURFfoundation covers a broad range of ICT-related areas, including Scholarly Communications. Under this rubric, SURFshare is a programme dedicated 'to creat[ing] a common infrastructure that will facilitate access to research information and make it possible for researchers to share scientific and scholarly information.'[37] It has six major themes, of which Collaboratories is one. The 2007 and 2008 calls for proposals under the programme concentrated on three themes: Enhanced Publications, Collaboratories and Knowledge Dissemination at Universities of Applied Science. In this round, the Collaboratories theme funded three collaboratories, all in the humanities and social sciences.

Origin and motivation of the project

The promotion of collaboratories are an integral part of the SURFfoundation's overall ICT for research and education strategy. The emphasis on collaboratories in the present round of calls for proposals occurs against the background of the Digital Academic Repositories programme (DARE) which ran from 2003 to 2006[38]. This programme focused on digitising, and making accessible the outcomes of research such as the journal or book publication (a portal that aggregates these outcomes can be found at www.narcis.info). While this is an ongoing concern of the SURFshare programme (which has another theme dedicated to 'Enhanced Publications'), there was also a felt need to support the research process as well. This is where the Collaboratories theme comes in. This theme is designed to support specific research processes that lead to publication: for example, archiving, using and re-using research data; the tools for working on those data and for collaborating. The main goal, therefore was 'to focus on the things that happen in the research process instead of just on the output',[39] such as, for example, supporting the authors of documents to share their documents, to collaborate on one document, and so on.

Features and technology

SURFfoundation does not itself develop technology; this is undertaken by SURFnet. There is no one technology supported by the group; instead they have tried to be as flexible as possible in meeting researchers' needs. Funded projects were given freedom to test any environment that they thought might suit their needs. Data sharing emerged as an important focus in the collaboratories, which was unexpected as humanities and social science research does not have a tradi-

37 http://www.surffoundation.nl/en/themas/openonderzoek/Pages/Default.aspx [Accessed 18/12/2009]

38 http://www.surffoundation.nl/wiki/display/dare/_English [Accessed 18/12/2009]

39 Interview, John Doove.

tion of data sharing and re-use in the same way as physics and other natural sciences (see HubLab case study). However, this has become a central feature of the VREs supported.

User community and development

The three projects that are currently running are user driven. Developers work closely with researchers, observing their research practices, and working with them to see what processes could be carried out more efficiently and how. This close interaction does not stop with the development of the tool, but is ongoing with hands-on workshops so that researchers can get acquainted with the tools. This includes producing visual instructions (such as videos) and step-by-step instructions. It has been found that without this kind of support there is poor uptake of the tools – even if the tools were requested by the researchers themselves.

Ethical, legal and institutional issues

Sharing data internationally is one of the main motivations for VREs, but it is difficult to know what exactly are the legal conditions to be met. A further issue has been with sharing research resources that require institutional subscriptions with researchers at institutions which do not have the same subscriptions.

Future plans

SURFfoundation has identified the need for an open platform into which researchers can drag tools that they are already using, or that they find meets their needs. This open platform is being developed by SURFnet based on open standards, with some basic functionality, which it will then be up to researchers to configure for their own needs and purposes. This includes tools such as Zotero[40] and other open tools that use open standards. Thus, the idea is to enable researchers to create their own 'customised, personalized, research environment'.[41] A pilot is scheduled for late 2010, with an initial group of users to test the open platforms. Alongside this, there are several planned activities to raise awareness about the usefulness of collaboratories, including roadshows and workshops.

Sustainability

The funding model adopted (50% national funding and 50% institution funding) may serve to distribute the costs of sustaining the VREs; however, sustainability is recognised as a major concern across the whole of the SURF organization. Since it is a broad issue (which includes buy in from institutions also at the top level), the sustainability of project outcomes such as VREs will be addressed along with that of other innovative technologies.

Lessons learned

The second round of calls for proposals was able to implement some of the lessons learned from the first call and other projects: that is, not to take the route of developing technologies and expecting researchers to use them, but rather to adopt a user-driven approach. Using a user driven approach has proven fruitful. A new challenge arises now that VREs are becoming more populair. Namely the fact that there is not one single VRE that can answer to all the needs of researchers throughout all the disciplines. Researchers have the need to fine-tune their own VRE specific to their personal research needs. This is being rolled out in the next phase of the collaboratories programme, with the open platforms approach.

40 http://www.zotero.org/ [Accessed 18/12/2009].

41 Interview, John Doove

5.1.5 TGE Adonis

Region	Europe (France)
Subject domain(s)	Arts and humanities; social sciences
Project started	2007
Funded by	CNRS
Website	http://www.tge-adonis.fr

Introduction

Adonis (Accès Unifié Aux Données et Document Numériques des Sciences Humaines et Sociales; Integrated Access to Digital Data and Documents in the Humanities and Social Sciences) is part of the TGE (Très Grands Équipements) strategy in France. TGE is an overarching strategy regarding large scale research in all disciplines. Adonis is the part of the strategy dedicated to social sciences and humanities.

Origin and motivation of the project

Adonis is part of a French National Research Strategy to archive data and promote integrated access to digital data and documents in the humanities and social sciences, a strategy which has been given a high degree of importance at national level. The ministry responsible for funding has the role of creating major infrastructures for research. This is politically and economically important, in view of European strategy on infrastructures, but also in view of the huge investment in research which is partly lost when research outputs are not archived and made available to other researchers. The Strategy is based on an understanding that individual researchers and institutions do not have the technical understanding or capacity for large scale archival storage of research data, nor are they (or should they have to be) in a position to undertake a task that needs to be addressed on a national level. Hence the formation of Adonis, which can provide – or facilitate the provision of – archival services, encourage understanding among researchers, help them to create their own resources, and raise awareness. Importantly, Adonis itself is a facilitator and not a funder of projects.

Features and technology

The four main activities of Adonis are: 1) Ensuring that various data and resources are archived, through partnerships with five national digital resource centres; 2) Establishing and maintaining Adonis Grid: Service infrastructure for storage and access; 3) Developing the Isidore platform which will provide searching and integrated access to data and documents; 4) Providing interoperable standards.

It did not seem that there was specifically a VRE programme or vision as such associated with Adonis. Still there is much overlap between the Adonis programme and other VRE programmes in terms of the kinds of projects funded[42]. While the emphasis is very much on making it easier for groups to archive, access, share and manage data, there is also a focus on developing tools, such as visualisation tools for archaeology.

A look through the different calls for proposals gives an idea of the priorities of the Adonis programme. The 2007 call aimed at attracting projects which would support and promote innovative tools and systems for the valorisation and dissemination of research data, allowing for resource

42 This was not specific to Adonis. For example, a raft of projects funded by ITEM (Institut des Textes et Manuscrits Moderne), which were geared specifically to smaller teams sharing data in secure environments[1] similarly were not classified under a VRE (or similar) rubric.

sharing and accessibility by the whole humanities and social sciences community. There is a stress on communities creating their own resources according to their own needs, reflecting a diversity of approaches but sufficient coherence to allow for transversal views, multidisciplinary work and international access. Collaboration is emphasised as is the adoption of Web 2.0 approaches.[43] The more than 20 successful projects of the 2007 call include projects around discrete collections of data or works, to projects to develop ontologies. The 2009-2010 call is divided into data processing and functionalities or tools, with an emphasis on those which will help to develop the Meta-Portal being produced by Adonis[44]. This includes archiving, search engines, metadata, software platforms, social networks for the facilitation of collaboration, and editing tools. The projects funded under these calls are seen as primary testbeds for ISIDORE.

With the emphasis on data archiving and accessibility, it is clear that the platform being developed, called *Isidore*, is of central importance. An external consultant was used for producing the requirements and specification for Isidore. Two members of this consultancy have a research background themselves, one in medieval studies and one in literary studies. They also have a lot of experience of working with libraries.

Ethical, legal and institutional issues

Ethical and legal issues had arisen around anthropological data, and the IPR of data and documents.

Sustainability

In terms of sustainability, Adonis is dependent on CNRS, so long as CNRS is committed to TGE, there will also be a commitment to sustainability. This must be seen in the broader context of the ESFRI European Roadmap that requires each government to identify a body that supports infrastructure. Adonis has been identified as that body for humanities and social sciences in France.

Lessons learned

- The importance of long-term sustainability being seen in terms of national research strategy.
- The possibility of combining a central platform with a broad range of interoperable but independent projects should be considered.
- What's in a name? 'VRE' is not a universal term, and sometimes there are no cognate terms either, yet there is a great deal of VRE-like activity.

43 Appel à projets: Outils innovants de traitement numérique pour la valorisation et la diffusion des données, Août 2007.

44 http://maquettewicri.loria.fr/fr.ticri/index.php5?title=TGE-Adonis%2C_appel_%C3%A0_projets_pour_2009-2010 [Accessed 18/12/2009].

5.1.6 Country perspective: Australia

Australia has been steadily progressing a comprehensive e-research strategy for over a half decade. Much of the national agenda for e-research was influenced by the UK e-Science Programme and the NSF Cyberinfrastructure. In 2004, the Australian government invited John Taylor (then Director of RCUK) to help inform this agenda, among other e-science experts on the international scene, and these consultations have led to a number of high-level reports and recommendations.

The Australian Commonwealth Government is supporting the enhancement of Australian e-research infrastructure through two sets of investments:

- The development of the **Platforms for Collaboration** components of the National Collaborative Research Infrastructure Strategy (NCRIS), that provide a total of $82 million between 2007 and 2011.
- The development of e-research components of the **Super Science Initiatives** that provide a further $312 million between 2009 and 2013.

Both investments are an outcome of the NCRIS Roadmap 2008[45] and its subsequent update in 2010 which underlined the critical importance of e-research infrastructure to future research competitiveness.

Platforms for Collaboration (PfC) capability[46]

The PfC investment is intended to support technological platforms that enhance researchers' ability to generate, collect, share, analyse, store and retrieve information, allowing them to access knowledge, data and information and work together seamlessly from desk to desk between organisations. In view of these perceived needs, there are VRE-like outputs that are being encouraged for development and use by research communities. The PfC capability consists of the following components:

National Computational Infrastructure (NCI)

The NCI[47] is expected to deliver an internationally significant high-performance computing (HPC) capability needed to support priority research. It will also provide a national strategy for computation infrastructure.

Australian Research Collaboration Service (ARCS)[48]

Established in 2007, major activities include engaging with its members in providing a number of e-research type services, and working in a coordinated manner with other service provider. Current service offerings include:

- Authorisation development services in collaboration with the Australian Access Federation and Council of Australian Universities Directors of Information Technology (CAUDIT);
- Video and Web Collaboration Tools;
- Data storage and collaboration, replication, and transfer services (the "ARCS Data Fabric");

45 https://www.pfc.org.au/pub/Main/WebHome/Strategic_Roadmap_Aug_2008.pdf [Accessed 04/01/2010].

46 http://ncris.innovation.gov.au/Capabilities/Pages/PfC.aspx [Accessed 04/01/2010].

47 http://nci.edu.au/ [Accessed 04/01/2010].

48 http://www.arcs.org.au/ [Accessed 04/01/2010].

- Remote and Grid Computing;
- Training, workshops and helpdesk services

The Australian National Data Service (ANDS)[49]

The ANDS component of the PfC capability will ensure researchers are able to identify, locate, access and analyse any available research data. In particular, ANDS will provide a systemic approach to research data to transform the disparate collections of research data around Australia into a cohesive corpus of research resources. A report, Towards the Australian Data Commons (TADC)[50], published in October 2007, formed the basis of deliverables required of ANDS and enabled the characteristics of its operation to be defined and broadly agreed.

Four programs of activity are being pursued: Developing Frameworks, Providing Utilities, Seeding the Commons, and Building Capabilities.

ANDS will focus strongly on engagement with the whole research sector, sourcing expertise for projects, activities and services from wherever that expertise resides. The Super Science Initiatives (see below) will also have a significant impact on the ANDS role to support the access and sharing of data through virtualised research environments and digital repositories.

The National eResearch Architecture Taskforce (NeAT)[51]

NEAT has been established to provide guidance on the evolution of the national e-research infrastructure and to identify and scope activities that broaden the appeal of e-research services. NeAT is also tasked with overseeing the development and implementation of e-Research tools and embedding these into research communities.

NeAT is responsible for recommending projects for investment within the ANDS and ARCS components. Two rounds of NeAT projects have commenced and are assisting research communities from the humanities through to biodiversity specialists – a number of virtual research collaborative environments are emerging from the NeAT calls, as well as next generation e-research tools and services.

Super Science[52]

Additional to the NCRIS funding, a second wave of significant investment was recently made available through the Super Science initiatives, announced as part of the national Government's budget in May 2009.

The e-research components of the Super Science initiative will support fundamental research, high quality research and international research.

They are intended to:

- enhance collaboration
- improve networks
- manage massive data assets
- support more sharing of more data, with faster analysis and better modelling
- provide super-computing enabled 21st Century analysis tools

The specific measures that comprise this e-research infrastructure investment are very much aligned with other international efforts, and are comparable to JISC programming in the Information Environment (e.g. Digital Repositories) and VRE areas:

---

49 http://www.ands.org.au/ [Accessed 04/01/2010].

50 http://www.pfc.org.au/pub/Main/Data/TowardstheAustralianDataCommons.pdf [Accessed 04/01/2010].

51 http://www.pfc.org.au/bin/view/Main/NeAT [Accessed 04/01/2010].

52 https://www.pfc.org.au/bin/view/Main/SuperScience [Accessed 04/01/2010].

The specific measures under Super Science relating to e-research related activities, include:

- Data Storage and Collaboration Tools - $97 million
  The project will build a national network of research data storage and collaboration infrastructure consisting of: regional data infrastructure for data retention and data access management; physical, IT and middleware infrastructure to enable data to be retained in regional data centres operated by participant institutions; physical, IT and middleware infrastructure to allow genuine and authorised researchers to access the regional data
- Data Commons - $48 million
  An Australian research data commons to support the discovery of, and access to, research data held in Australian universities, publicly funded research agencies and government organisations for the use of research.
  This investment will enable the construction of a range of ICT utilities to capitalise on and ensure greater use and re-use of existing data resources, as well as better management of new data generated in Australian research. This is likely to entail a number of domain specific and general Virtual research type environments to support access to and sharing of data.
- The National Research Network - $37 million
  The Australian Research and Education Network (AREN) will be extended and upgraded to connect regional research data centres with each other, with new and existing HPC centres, and all other high volume sources of primary research data.

5.1.7 Country perspective: South Africa

South Africa is at early stages of an e-research strategy. The technical infrastructure for e-research is progressing fast, with optical cables almost in place along the coasts, and spreading to the centre of the country[53]. With the necessary bandwidth and supercomputing resources for e-research becoming available in the very near future, several stakeholders in the country are preparing for an e-research and VRE strategy. Members of the Council for Scientific and Industrial Research (CSIR) in South Africa have been in close contact with the JISC VRE programme, as well as with e-research initiatives in Australia, and are likely to try to implement similar kinds of plans and processes. In South Africa we could see the results of lessons learned in the early developments of VREs being directly applied, and thus a leap-frogging over early developmental stages.

It is unclear how much e-research activity is taking place in South Africa at the moment. Even though there is not a specific e-research strategy in place as yet, our interviewee believed that there is in fact quite a lot of 'hidden' activity. This activity may be occurring through South African researchers working through international networks, such as EU projects which have a South African partner.

There are some areas of research which could benefit greatly through a VRE. One area which has been identified is that of the South African Malaria Initiative (SAMI).[54] SAMI is a consortium with 14 members, including the major research universities in the country. Collectively researchers associated with it have significant quantities of data, which it would be extremely beneficial for them to be able to share. Using requirement principles developed in the JISC VRE Programme – specifically BVREH and IBVRE – this initiative was used as a case study for a potential VRE for malaria research which would facilitate data sharing.[55]

A major issue for this community are the ethical, legal and institutional issues around data sharing. Apart from the usual constraints around data sharing relating to researchers' professional and institutional interests, the South African IP Act, promulgated in 2008, prohibits data generated through state funded research to be shared without permission from state structures. However, at times it is not clear to researchers that they are in fact delivering data to another country when they become involved in an e-research initiative. Therefore, this will need addressing in any VRE strategy.

Other aspects which will affect the development of a VRE strategy is the collaboration with the UK Data Archive being conducted by the HSRC (Human Sciences Resources Council) which funds research in humanities and social sciences. Here there is a programme to develop open access publishing and digital curation.

The development of the VRE programme will depend on many other factors, including the overall e-research strategy in the country. It is interesting however, that there is a perceived need for a VRE strategy, along the lines of that adopted by JISC.

---

53 Page-Shipp, RJ, Hammes, MMP, Pienaar, H, Reagon, F, Thomas, G, Van Deventer, MJ, Veldsman, S 2005. 'e-Research support services: responding to a challenge facing the South African research and information communities', South African Journal of Information Management, vol. 7, no. 4.

54 http://www.sami.org.za/ [Accessed 18/12/2009].

55 Heila Pienaar and Martie van Deventer (2009) To VRE or Not to VRE?: Do South African Malaria Researchers Need a Virtual Research Environment?, Ariadne, Issue 59, http://www.ariadne.ac.uk/issue59/pienaar-vandeventer/. [Accessed 18/12/2009].

5.2 Projects

5.2.1 Alzforum, Schizophrenia Research Forum, PD Online, StemBook

Region	North America (USA)
Main project partners	Each project as an independent organisation; Massachusetts General Hospital, Harvard Initiative in Innovative Computing; Harvard Stem Cell Institute
Subject domain(s)	Medical sciences
Technologies	Cloud computing; publication technology; web services; web communication technology; semantic web
Project started	1996
Funded by	Private philanthropic foundation, Michael J. Fox Foundation, NARSAD (A charity promoting Mental Health Research), National Institute of Mental Health
Website	http://www.alzforum.org/, http://www.schizophreniaforum.org/, http://www.stembook.org/, http:/www.pdonlineresearch.org/

Introduction

Alzforum, Schizophrenia Research Forum, PD (Parkinson's Disease) online and StemBook is a cluster of projects in biomedical research, focusing in the first instance on neurodegenerative disorders. Funded mostly by a philanthropic foundation, the projects are the result of a collaboration between the foundation, a team of dedicated staff and consultants, and an interdisciplinary team of bioinformaticians and other biomedical researchers at the Mass General Institute for Neurodegenerative Disease. The first of these projects was Alzforum, which started in 1996; following on its success, three further projects have been initiated, the most recent being PD Online, launched in 2009. It was decided to establish Alzforum as an independent organization, since it is felt that this helps to maintain the neutrality of the organisations, something which is important in areas of research where there can be a great deal of contestation between rival hypotheses.

Origin and motivation of the project

Alzforum was launched in Osaka, Japan in 1996, and is thus the longest standing VRE included in this study. The project began with the impetus of funding from a private philanthropic foundation, which was convinced of the benefits that could be obtained by using the Internet for collaboration in biomedical research. Research into Alzheimer's Disease was focused upon as the specific area in which there was the possibility of achieving great advances through Internet mediated collaboration. Specifically, this was as it was difficult for researchers to access the information needed to judge which of the proliferation of competing hypotheses were more worth pursuing, and to enter into discussion with each other. Alzforum started off modestly as a kind of community newspaper, publishing paraphrased abstracts of papers on Alzheimer's disease published globally, slides and audio of relevant presentations at scientific conferences, a collection of Milestones Papers in AD research, running a 'Paper of the Week' feature, and facilitating rapid, informal communication between researchers – for example, through live chats, and enabling comments and discussion about papers. The numbers of registered users quickly grew, with as many as 100 new users a month in the first years of its existence. From 1997, the website also began to develop as a community repository, enabling researchers to deposit data sets. Currently Alzforum maintains several databases relating to gene mutations, gene association studies, antibodies, and drug trials. Alzforum also acts as an integrator of these diverse sources, linking primary research articles to related news, papers, databases, discussions and so on.

Schizophrenia Research Forum has similar functionality to Alzforum; StemBook instead is a collection of open access chapters on stem cell research linked to related databases, allowing readers to post comments and discuss entries; PD Online focuses on enabling discussion around discovering which are the most promising research directions and emerging hypotheses. It is aimed at forging a co-operation between funders and researchers to discover which are the most promising research directions and hypotheses that merit funding.

Features and technology

Alzforum is an intensely socio-technological organisation, with editors playing a pivotal role in encouraging and moderating discussions and commentary, and discovering and integrating information. From 2000 onwards a data-driven dynamic system has been used to automatically search and download PubMed citations into a database every night, and provide tools to let editors post news and comments and crosslink them to related material. The development of semantic web tools is a current and ongoing development. These tools assist in the identification of hypotheses and related evidence in papers and discussions. SWAN (Semantic Web Applications in Neuromedicine) has resulted from a collaboration between Alzforum and Massachusetts General Hospital. Work on this continues with the Scientific Collaboration Framework.[56]

User community and development

In 2007 Alzforum had 4600 registered users[57], and currently has more than 5000 users. It is estimated that 30-50% of researchers studying Alzheimer's internationally consult the site, are registered on it, or are active on it. Researchers using Alzforum also come from an extremely wide range of institutions and countries. The other webs sites also have significant membership numbers: Schizophrenia Research Forum has 3045 registered members, PD Online (the most recent of the sites) has more than 1650.

Future plans

Further integration of the websites with Web 3.0 functionality is continuing. This includes further research being carried out on systems which will enable semantic web applications for representing hypotheses and evidence in scientific discourse, in order to assist in the sharing and discovery of knowledge.

Sustainability

Increasing the number of communities using the same model is beneficial since this helps to lessen the costs for all. In line with this, there are ongoing discussions with foundations and other funders to bring on new communities. For example, there are plans to build a community of researchers on autism, and to extend the core annotation framework to the Journal of Neuroscience as well.

Lessons learned

- Tim Clark points out 'the importance of understanding the ecosystems you are working in, and consciously seeking to place what you are doing in a producer and consumer ecosystem'. Socially collaborate with others doing the same kinds of things, helping others doing the same kind of things.
- Consciously build user communities; not building something independently of user buy-in.

56 http://www.sciencecollaboration.org/ [Accessed 18/12/2009].

57 Clark T, Kinoshita J. Alzforum and SWAN: the present and future of scientific web communities. Brief Bioinform 2007; 8(3):163–71.

- Use of the model will depend on how well you can get it accepted in the scientific community.
- Re-using other people's work, and not trying to build a monolithic independent structure yourself. Finding ways of interlocking with others' work.
- Projects being 'owned' by legal entities or consortia instead of one institution can increase the credibility of and engagement with a VRE, as it demonstrates neutrality and may prevent the impression that content is being taken over by a rival organisation.
- The lines between publication and communication platforms, repositories and VREs are becoming increasingly blurry.
- Semantic web technologies can be of great use in analysing content in a VRE.

5.2.2 Biogrid

Region	Australia and Oceania (Australia)
Main project partners	A consortium of hospitals, medical research institutes and universities, development is led by the Melbourne Health
Subject domain(s)	Medical Sciences
Technologies	Federated repository architecture
Project started	2004
Funded by	Australian and Victorian Governments
Website	http://www.biogrid.org.au/

Introduction

BioGrid Australia is a not for profit platform for life science research teams to access and share genetic and clinical research data across multiple organisations in an ethically approved and secure way. As a partnership of hospitals, research institutes and universities, it integrates multiple data sets (clinical, genomic, tissue bank, biomarkers) for ten tumour streams, diabetes and epilepsy and has extended into other diseases. Physically located within independent organisations, the data sets can be analysed remotely by querying via a federated data integrator. BioGrid Australia has developed a number of applications to assist with data collection, provides the integration, the analytical tools as well as assistance with data analysis and generating reports. It is funded by several programmes from both the Australian and Victorian Governments.

Origin and motivations

The use of medical data in a collaborative environment is crucial for research and treatment of patients. However, it is also a challenge for e-research as security and adherence to ethical issues are critical in handling very sensitive data. Allowing clinical researchers to study treatment outcomes and influences covering a range of diseases across a whole country has enormous potential, but is practically impossible with a secure distributed architecture and an accompanying set of regulations. In 2004, Bio21 Australia, a biomedical, biotechnology research cluster supporting collaborative projects, initiated a project to facilitate cross-institutional access to patient level data for research. While the pilot project, which was funded by the Victorian State Government, had a more limited geographical scope, the project was intended to and has now grown to a national platform.

Features and technology

The BioGrid Australia platform provides researchers with the capability to access, integrate and link data across many environments regardless of their existing linkage and research platforms. BioGrid Australia is a ‘federation’ of the researchers’ repositories and can connect and integrate data from the participating hospitals and research centres in Australia.

- Firstly Institutional Ethics approval is obtained for the database to be securely connected to BioGrid and to BioGrid processes. Other agreements are also signed
- A secure (virtualised) Local Repository is established at the institution
- The source data from databases at various institutions is extracted, transformed and loaded (ETL) on a nightly basis to their respective research repositories located at the institution
- The data is record-linked for the individuals using probabilistic matching or hashing and a record linkage key is assigned and stored in encrypted format at the institution

- Authorised researchers are then able to query and analyse the data via the Federator using statistical/business analysis and query tools
- Researchers must obtain authorisation to access data from the data custodians, the Scientific Advisory Committee and the BioGrid Australia Management Committee

The federator is an integrator for accessing data across physical boundaries. The data is sent to the user via a Virtual Private Network (VPN) and secure socket layer web services in de-identified form with a record linkage key. The federator does not store health data.

User community and development

The BioGrid project was envisaged and developed in close consultation with researchers. The initial requirements were established through a consultation with clinical, population and laboratory based researchers who then also helped evaluate the pilot that was built based on the requirements. These researchers have continued to be involved with advisory groups to ensure that the platform meets the needs of the research community. BioGrid now supports researchers from its group of 21 member organisations and 5 collaborating institutions. Currently, about 250 researchers engage with the project to study which factors influence disease outcomes or why particular drugs work for some patients and not others. The main focus of the collaboration is access to data sets, followed by an analysis and a secure way of sharing the results. For further documentation and discussion of the results, for instance to then develop publications, the researchers are using alternative tools.

Ethical and legal issues

Medical information is among the most sensitive data researchers deal with; storing, accessing and sharing that data is (rightly) highly regulated legally and can also be problematic from an ethical point of view. Because of this, the BioGrid project had to pay rigorous attention to ethics and privacy requirements. All participating sites must obtain ethics approval to join. BioGrid Australia complies with all privacy legislation and regularly seeks independent external legal advice to ensure the project continues to comply with all relevant privacy legislation, as well as conducting regular security audits. The research data is used in a de-identified (codified form) but the system allows the patient to be ethically re-identified, if required.

Research outputs

BioGrid Australia has enabled the implementation of many successful collaborative research projects. Publications are listed on the BioGrid website.

Future plans

BioGrid is working with a number of disease groups to facilitate national collaborative research as well as help with quality health care reporting to assess healthcare performance targets. In addition surveillance of some conditions is being developed. The project also expects that the number of participating researchers will grow significantly as data relating to additional diseases will be included.

Sustainability

BioGrid Australia is in the process of exploring options for the sustainability of the project beyond 2010; the project plan for the next year includes several key milestones relating to this.

Lessons learned

- Offering a reliable solution that addressed the issues of intellectual property and privacy has made BioGrid attractive to its stakeholders and ensured its growth

- Even when there is strong interest in a platform that offers unique features, it can still take the research community some time to understand the platform and the nature of virtual collaboration
- Local champions that support the uptake of a VRE and encourage collaborative working in a virtual environment are crucial for the success
- Even in the same subject domain, researchers do not always have the same requirements or ways of working; this may result in an interest in very different data sets
- In essence, BioGrid supports two distinct types of researchers: a) Those who require in-depth data on smaller numbers – the biomedical translational researchers; b) those who require the breadth of data – population health researchers. The data sets that give value for both groups are different but data such as death date and cause and national data gives value to both groups.

5.2.3 Cleo

Region	Europe (France)
Main project partners	CNRS, École des Hautes Études en Sciences Sociales, University of Provence, University of Avignon
Subject domain(s)	Arts and humanities; social sciences
Technologies	Editing and publishing technology
Project started	1999
Funded by	CNRS
Website	http://cleo.cnrs.fr/index.html

Introduction

Cléo (Centre pour l'édition éléctronique ouverte/Centre for open electronic publishing) currently falls under the umbrella of the Adonis TGE, although it started a long time before the inception of this programme. It is a website dedicated to facilitating the editing and publication of resources in arts, humanities and social sciences by individual researchers as well as teams. Cléo has three sub-projects: 1) Revue.org[58], a platform for the publication of journals; 2) Hypothèses.org[59], which allows individual researchers or groups to create notebooks on particular ideas or themes, and to have discussions and debates; and 3) Calenda[60], a calendar listing all events, calls for papers, calls for proposals, conferences, courses, jobs, etc. in the social sciences and humanities. These sub-projects are all interlinked.

Origin and motivation of the project

Marin Dacos is the director and originator of this project. He began it as an undergraduate student in 1999, as a project dedicated to the spread of knowledge through open publishing, in three senses of 'open': open access (as far as possible), open to communicative exchange and commentary, and open in the sense of direct, without editing intermediaries, or with as few as possible. The project soon gained community support. In 2002 it began to gain institutional support from the École des Hautes Études en Sciences Sociales, and in 2007 from the CNRS under the Adonis TGE programme. Thus it is in the first instance a project which developed through community demand rather than in response to an institutional call for projects. The project never did respond to a call for proposals; however the project was already very well known and accepted by the community, and it would have been counter-productive to create a new, institutional, publishing platform. There are now 20 members of staff.

Features and technology

The emphasis is on publishing and not on data; the project sees its task as one of enhancing publishing, giving more visibility to publications, for example, and not curating data. The project both develops its own software and uses open source software systems, such as Wordpress, which is the editing system used for Hypothèses.org. Lodel[61] is the software which has been developed by Cléo for journal publication. It is open software, and is also used beyond Cléo.

58 http://www.revues.org/ [Accessed 18/12/2009].

59 http://hypotheses.org/ [Accessed 18/12/2009].

60 http://calenda.revues.org/ [Accessed 18/12/2009].

61 http://www.lodel.org/ [Accessed 18/12/2009].

The project also supports users' use of an open access system to support each stage of the workflow of publication, from submission to final publication, through 'Les manuscrits de Revues.org[62]. Training and support for this are not free.

Collaborative writing is not currently supported on Cléo. However, in the future a plug in ('Command Press') from 'The Future of the Book' in Hypothèses.org will allow for collaborative writing.

Users, virtual community, research outputs

With the emphasis on publishing without external intermediaries, it was important that the software should be easily usable. Cléo offers frequent free two day training courses in Paris on the use of the software to anyone wishing to start a journal on Revues.org. These training courses are also an opportunity for them to get feedback on the software, and feedback continues as people begin to use the software. Thus there is continuous development according to user needs.

Cléo is well known among humanities and social science students and researchers. There are 35 000 registered readers of the monthly newsletter, has 1.2 million monthly visits, supports the online publication of 200 journals, has supported the publication of over 40 000 documents, and over 70 notebooks. The project is also beginning to be more internationalised.

Sustainability

The project has declared itself 'Précaire insiders'[63]: in response to the state increasing support to research which is carried out on a contract basis rather than on a permanent basis. This leaves projects like Cléo in a precarious position, notwithstanding its being within the state research structure. For example, notwithstanding institutional support, the project also struggles to pay all of its staff salaries. The project has a threefold strategy for sustainability: 1) continue to convince the institutions to give them stable support; 2) charge for some functionalities, at the journal level; 3) offer some bespoke functionalities to libraries which they would pay for. There is also a charge for bespoke training courses.

Lessons learned

- Training of users is important and also a good way of generating feedback that can lead to improvements of a system.
- High visibility, much content and large numbers of visitors/readers do not necessarily guarantee sustainability of a project if no funding structures for long-term projects exist.
- Projects growing out of the research community and fulfilling their needs will attract continued use by the community.

62 http://cleo.cnrs.fr/index.html [Accessed 18/12/2009].

63 http://cleo.cnrs.fr/index868.html [Accessed 18/12/2009].

5.2.4 e-Resource Centre (e-RC)

Region	Australia and Oceania (Australia)
Main project partners	Victorian eResearch Strategic Initiative (VeRSI) and DPI (Department of Primary Industries) with Monash University, La Trobe University, University of Melbourne
Subject domain(s)	Natural sciences
Technologies	Apache Tomcat, Confluence, OpenIdP, Shibboleth, Web 2.0 applications
Project started	2007
Funded by	Mixed funding; mostly VeRSI resources, contributions from DPI and in-kind contributions by partners
Website	https://www.versi.edu.au/kb/erc

Introduction

The e-Resource Centre (e-RC) is a VRE for managing ecosystem and environmental data and information products. The e-RC enables cross government-university collaboration among DPI (Department of Primary Industries), Monash University, La Trobe University, the University of Melbourne and partners in the Australian State of Victoria engaged in the VeRSI Ecoinformatics Climate Change Demonstrator project. The e-RC provides access to regional climate change scenarios, maps, models and model outputs, decision support tools, research reports and communication material. It provides a virtual collaborative space for managing state-wide ecosystem and environmental data, information and knowledge. At present, the e-RC is being piloted with the DPI VCCAP (Victorian Climate Change Adaptation Program) team and partners. It is currently being used by about 130 researchers.

Origin and motivations

e-RC originated from the need of a group of researchers spread across different Australian universities and government bodies in the State of Victoria to share environmental data – and to have access to an open 'ecosystem' data & knowledge infrastructure to support complex policy and decision making (EcoInformatics Steering Committee, October 2007). As the group was both geographically dispersed and restricted by institutional firewalls and security rules for network resources, a secure virtual environment for sharing of data and collaboration was needed. VeRSI, the Victorian eResearch Strategic Initiative, who have a remit to apply e-science methodologies to support researchers, undertook the development of a demonstrator that evolved into the e-RC.

Features and technology

Because of the needs of the researchers, developing a secure environment for access and authentication emerged as one of the key components of the e-RC. As a 'Virtual Organisation', the e-RC consists of a Confluence wiki running on an Apache Tomcat server. It is sitting behind the VeRSI VO IdP (VeRSI Virtual Organisation Identity Provider) that allows users to log in to the environment and identify themselves. The VeRSI VO IdP is composed of an OpenIdP (specifically to allow government researchers to have authenticated access to the e-RC) and the Shibboleth federated identity system (in use by the University partners and compliant with the Australian Access Federation). Accessible through a web interface, e-RC provides access to a range of shared tools and services including: virtual library, discussion board, project calendar, interactive maps, data repositories, wikis, search tools, auto-notification of new uploads and RSS feeds. Visualisation and geospatial technologies are important components of the e-RC

too. Currently, the environment provides access to more than 1000 digital artefacts including documents, maps, models and 3d objects.

User community and development

Research in climate change as supported through the e-RC is interdisciplinary and looks not only at the climate itself but also at effects on, for instance, hydrology and topography (such as wind and water erosion). Based on feedback from the VCCAP group, VeRSI developed a demonstrator that was launched in 2008 and has since been improved with feedback from the user community, using an iterative approach to development, for instance in improving and simplifying the user interface – a key request from researchers. The team also used workshops and surveys to get feedback from the users and interested researchers. Following that a steering committee has been set up to represent the stakeholders and to provide guidance on further technical development. Usage of the environment has grown steadily and the e-RC is now facilitating collaboration within a group of researchers that previously was restricted by the available technology. There are also first signs that the VRE is changing culture within the groups that are using it and it encourages a range of research collaborations across institutions.

Research outputs

The research of the VCCAP directly informs government policy, but individual members from across the partnership are also involved in producing research publications based around data objects, such as climate models and workflows, that are accessible through the e-RC. Despite being originally thought of as just a demonstrator, the e-RC environment has already contributed to a research mission.

Future plans

Planned technical development includes: improvement of 3D visualisation features (including an AutoCAD based 3D object viewer); interfaces and access to other data sources (such as spatial queries); improved workflow support; usability. The project is also planning further outreach and training activities. The platform itself has also gained interest by others and VeRSI are currently working with a group of major life sciences research institutes, hospitals and medical department to build an environment for cancer and genome research. It may be possible to develop the environment developed for e-RC into a more general platform that can be used by others (most of the software components are open source and/or modular).

Sustainability

The project in its current form (demonstrator phase) will come to an end in June 2010. The project has just submitted a transition to service document to the steering committee; further discussions are taking place to determine if and how the project can be continued. Thus far, there is a positive response from stakeholders recognising the need for the e-RC in some form to be taken forward (as of December 2009).

Lessons learned

- An iterative approach with researchers and stakeholders in the development of the VRE is essential for acceptance and uptake.
- Usability, especially with regards to simplifying the user interface, is also crucial for acceptance.
- Visualising research results through Google maps and similar Web 2.0 tools is of interest to researchers.
- Project development still needs to go hand in hand with awareness raising as researchers might not necessarily be aware of what functionality they could use for their work.

- A VRE moderator or content manger is essential to support the VRE, as well as making available access to technical support for critical issues.
- Sharing of data is an important driver for acceptance of/interest in VREs among researchers.
- Easy to use and secure authentication is a key feature for cross-organisational collaboration.
- Sharing data and infrastructure across organisations can result in tangible benefits to collaborative research and 'synergies', as well as longer-term cost benefits (e.g. shared service type provision).
- Ownership of content and IPR in a distributed and collaborative environment is one particular issue that requires an agreed policy as soon as possible.
- Keeping to an open 'system' and generic platform with some targeted research tools in the first instance has aided in bringing together much more quickly collaborating researchers across organisations and has supported transferability of the platform to other domains.

5.2.5 eSciDoc and the Max Planck Society

Region	Europe (Germany)
Main project partners	Max Planck Society and FIZ Karlsruhe
Subject domain(s)	All domains; VRE component currently arts and humanities as well as engineering and physical sciences
Technologies	Fedora Commons; PostgreSQL; Lucene; Java, JSF; SOA based (SOAP, REST interfaces)
Start date	2004
Funded by	German Federal Ministry of Education and Research (BMBF); contributions from project partners
Website	https://www.escidoc.org/

Introduction

eSciDoc is a joint project of the Max Planck Society and FIZ Karlsruhe, aimed at building an e-research platform for multi-disciplinary research organisations. It is not a single VRE, but rather a framework that allows several virtual environments (called 'Solutions', in the eSciDoc terminology) to share the same infrastructure. eSciDoc allows its users to publish, visualise, manage and work with data artefacts, including publication data and research data across disciplines. The platform is being tested by the Max Planck Society, an association of 80 German research institutes with approximately 13,000 employees working across the domains of natural sciences, life sciences, social sciences and the arts and humanities. eSciDoc is also an open source community project joined by several partner institutions.

Origin and motivations

The original motivation behind eSciDoc was to build an e-science research infrastructure for the Max Planck institutes, with a particular focus on stable, trusted and secure storage facility for sharing and preserving of data. Because of the diversity of the institutes, it was decided early on to develop a generic infrastructure into which services could be built to support specific research communities. This vision was informed by the Berlin Declaration on Open Access of 2003,[64] leading to the idea to provide an 'Open Access Platform' (which was the original project name) for publication. During the requirements gathering, it became obvious that researchers were initially much more concerned with issues such as control over their data, which made authorisation and identification key issues for the project.

Features and technology

The eSciDoc Core Services form a middleware for e-research applications. The Core Services encapsulate a repository (Fedora Commons) and implement a broad range of commonly used functionalities, including access control and authentication, using service oriented architecture (SOA). Initially, four basic scenarios were envisaged, but because of the diverging requirements it was decided to focus on only two: plans for an electronic laboratory journal and a hosting infrastructure for library resources were shelved, while work on a publication management component and a scholarly work bench to deal with research data went ahead. The repository currently holds over 4,600 publications, over 20,000 pages of manuscripts and more than 6,000 images. A collection of scanned journal articles about 6,000 PDFs is expected for next year. Due to the migration of the current publication repository of the MPG to the eSciDoc, about 100,000 publications will be added when the existing MPG repository will be migrated to the eSciDoc platform.

64 http://oa.mpg.de/openaccess-berlin/berlindeclaration.html [Accessed 18/12/2009].

On top of the eSciDoc infrastructure sit the 'Solutions', VREs for specific communities. Currently, eSciDoc offers solutions for researchers working with digital images (FACES) and digital text (ViRR).[65] Both are being developed with pilot groups of researchers as part of an iterative approach to develop a more general Scholarly Workbench. While eSciDoc relies on SOA-based infrastructure, a few Web 2.0 style applications such as Mediawiki and blogs (Wordpress MU) have been integrated too, as researchers – despite a certain scepticism towards Web 2.0 as a toy – liked their ease of use; a Wordpress plug-in[66] also allows to use Wordpress as an interface for access to some digital collections.

Users and virtual community

The eSciDoc users, currently about 300 for the publication management and about 20 for the solution FACES, are a development community consisting of librarians, local IT staff/developers and researchers. Because of their sometimes diverging interests and requirements (metadata vs. services vs. data), the project team decided to use an iterative design approach starting with basic functionality and adding more features based on an analysis of workflows, requirements and workloads and the feedback of the users. Currently, the publication management component of eSciDoc is used about equally by librarians and researchers across all disciplines, with the percentage of researchers increasing as it is being rolled out to the MPG. The image environment is used by an international group of humanities scholars from and associated with the Max Planck Institute for Human Development, and by a group of material scientists based in Japan. Users of the text environment are mostly research librarians, associated with institutes focussing on arts and humanities (specifically art and legal history) as the development was initially focussed on an editor for structural metadata. Subject librarians at the institutes were actually key factors in getting researchers to engage with the architecture as they helped to bring user groups together and could also act as translators between researchers and developers. This was even more important as the institutes of the MPG have a lot of autonomy and need to be convinced of the benefits of using a central platform such as eSciDoc. The project team works with institutes to bring them aboard and to set up collaborative ventures. eSciDoc is also used outside of the MPG, through partners in (currently) Germany, Japan and Denmark, who share in the further development of the platform. Some are also using the MPG infrastructure for storing data.

Ethical and legal issues

So far the project has not encountered any major ethical or legal issues in terms of VRE development. IPR is a problem only insofar as it affects any publication repository – librarians have to check that open access publications do not violate any third party rights before publication.

Research outputs

The focus so far has been on technical development, which in itself is seen as a research outcome. This applies specifically to the complete specification and discussions regarding functionalities, technologies and architecture. Most of it is available publicly.[67]

Future plans

The project is committed to expand the eSciDoc community and make it easier to deploy the platform in other institutions, for instance through better documentation and improved installers. The Scholarly Workbench will be further developed, as will be the specific solutions; one of the next steps for the textual environment will be to focus on collaboration aspects of textual transcriptions, a development process that is monitored by a group of researchers. Additionally, in-

65 https://www.escidoc.org/JSPWiki/en/Solutions [Accessed 18/12/2009].

66 http://colab.mpdl.mpg.de/mediawiki/Category:Wordpress_Plugins [Accessed 18/12/2009].

67 http://colab.mpdl.mpg.de/mediawiki/Main_Page [Accessed 18/12/2009].

creasing the number, size and diversity of digital resources and collections is continuous work that will be done in parallel as well in future.

Sustainability

Both FIZ and the MPG have committed to support the eSciDoc platform after the initial government funding runs out. Other members of the community are also contributing to the work. Even so, eSciDoc will have to demonstrate that it can contribute to the research mission, institutional strategy and changing requirements of the participating institutions as well as the individual researchers.

Lessons learned

- Researchers are increasingly interested in getting their data online quickly so that they can share findings more easily and collaborate better; sustainability is of interest too, but less so.
- The consequences and practicality of crossing discipline boundaries is something that the researchers themselves often do not yet fully understand a priori, which makes it even more difficult to build appropriate services and solutions to facilitate this collaboration.
- Developing an infrastructure to support such research is even more difficult as the technologies and standards often need to be further developed in the process, turning development work into research in its own right. Striking the balance between providing working environments, developing innovative prototypes and documenting complex systems is not easy.
- In the same way, developing new features while ensuring the integrity of existing data is a challenge.
- In terms of technical development, ensuring data and service interoperability, as well as conformance to standards are key factors.
- Trust is a key factor in uptake of a VRE – in both the technical infrastructure and the developers; users have to be certain that they can control and use their data in the way they want.
- Ambassadors are needed and subject librarians may have a key role to play as interpreters between their research community and the developers and general information specialists.
- It is easy to underestimate the difficulty of providing research infrastructure for a larger organisation, especially if it is comprised of autonomous institutes and research communities with very specific interests. In this context social and political aspects can be as or more important than technological ones. Organisations themselves may have to adapt and change their traditional ways of working if they want to make use of the potential of e-research infrastructure. 'This is one of the outcomes for me that VRE development is not only just about building IT solutions. The organisation has to be prepared to be a learning organisation, and that is not always easy.'[68]

68 Ulla Tschida, interview.

5.2.6 HubLab

Region	Europe (Netherlands)
Main project partners	Institute of Social History
Subject domain(s)	Arts and humanities and economics and social science
Technologies	Data repositories
Project started	2007
Funded by	KNAW
Website	https://collab.iisg.nl/web/guest http://www.iisg.nl/research/labourcollab/

Introduction

HubLab is a project based at the International Institute of Social History in the Netherlands. This Institute has created a hub comprising of five collaboratories, all in the area of social and economic history. HubLab aims to study the working practices within these collaboratories and, more specifically, the user experiences with the VRE built in Liferay. HubLab is funded under the SURFfoundation's Collaboratory Programme, and is thus co-funded by the Virtual Knowledge Studio for the Humanities and Social Sciences and the hosting Institute.

Origin and motivation

The main motivation for the Hub project derives from the long-term objectives of the Institute firstly to be a world leader in new forms of research, in social and economic history, and secondly to enable the writing of global history. This objective requires global collaboration between researchers. This still occurs largely through face-to-face events; however it was hoped that the impetus of these events would be continued through the collaboratories, which would also allow for diverse and globally distributed data sets to be brought together, and for analysis and comparison across these data sets.

Features and technology

The focus of the collaboratories is on allowing researchers across the globe to share data, and the whole infrastructure for the collaboration allows sharing to occur meaningfully. The HubLab project has designed the VRE for the Hub-project. The approach taken was to use an existing resource rather than to build something from scratch. The project team chose Liferay, which is an Open Source platform for collaboration.[69] Liferay was chosen because it is a lightweight platform which is easily navigable and easy to use across different bandwidths and conditions. This was imperative given the globally distributed nature of the research teams. However, there were several problems with the software not providing the functionality that was expected, leading to frustration. An advantage of using a product like Liferay is that there is a community site for working on bugs and problems, so the project developers were themselves involved in a joint design project. At this point, the project has carried out substantial modifications of the original Liferay software.

Data sharing is the main objective of the project, but also its main challenge. Since the humanities do not have a long history of sharing data, there is not a culture of producing standardised data. When the data are newly created, data standards, taxonomies and overarching schemes for data management are co-created by the different groups of researchers, through face to face workshops as well as through collaboratory mediated interaction. However, the vast majority of

69 http://www.liferay.com/web/guest/home [Accessed 18/12/2009].

the data sets pre-exist the collaboratories, and in these cases there are intermediate structures which usually only allow for a few variables for comparison across data sets.

User community and development

Users of the collaboratories are demographers, sociologists, economists, and social and economic historians. While the IISH Hub is based in the Netherlands the communities are global. For example, one of the collaboratories has 60 active members, and only five or six of these are in the Netherlands.

The motivations of individual researchers to participate in the collaboratories include the desire to consolidate an existing group, or the advantage of having a clearly defined network already in place when seeking funding. The need to share data is an over-riding motivation since it will enable researchers to get to the next level of research. For example, the Global Inequalities project brings together local data sets, for example on prices and wages, which when exchanged, allow for new research findings and understandings on a world level.

The role of managers is vital to the success of a collaboratory. Managers facilitate online and offline collaborations in a number of different ways. Usually, managers are themselves members of the community, with good reputations in their communities.

Ethical, legal and institutional issues

It is not always clear to researchers what the legal requirements pertaining to data actually are. For example, when wanting to upgrade a data collection which originally was physically located in the US, the IISH Hub team were not permitted to upgrade the data, even though it was stored on their site.

Future plans

Further funding is being sought to build a 'collaboratory of collaboratories', developing further the approach already used with the development of Liferay. In a future development of the project, it is expected that the central hub will harvest data from the different collaboratories. Data protocols for formatting data will be required. Thus members of the collaboratories will need to agree on data formats from the inception of the projects.

Sustainability

Since the driving force behind the Hub collaboratories is the Institute for Social History, it is likely that the Institute will take care of the sustainable storage of data. The project is also developing a sustainable storage scheme and a scheme to license data, essentially an overall global hubs strategy. However, the individual projects are very much linked to the main researchers involved.

Lessons learned

- More attention to be paid to developer/user communications, for, among other things, ensuring that user expectations remain realistic.
- Liferay turned out not to be as user-friendly as expected at the outset, leading to many frustrations.
- Live demonstrations are very important but not always successful, and not always possible for globally distributed communities. Video demonstrations are being developed.[70]

70 See Jan Kok (2008): HubLab: Towards online collaboratories for global data gathering in social and economic history. Final report.

5.2.7 The Membrane Research Environment (MemRE)

Region	Australia and Oceania (Australia)
Main project partners	A research cluster of nine Australian universities, led by Victoria University; technical development through the UNESCO Centre for Membrane Science and Technology and the University Library at the University of New South Wales
Subject domain(s)	Engineering and physical sciences; biotechnology and biological sciences
Technologies	Fedora Commons, Wiki, single sign-on
Project started	2007
Funded by	Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Website	http://membranes.edu.au/

Introduction

The Membrane Research Environment (MemRE) is a research infrastructure project within the Advanced Membrane Technologies for Water Treatment Research Cluster, a research project funded by the Australian CSIRO. MemRE is a collaborative digital library of information on membrane separation in engineered and biological systems. It aims to: 1. Accelerate the dissemination of information on the mechanics of desalination by membrane processes. 2. Identify, refine and catalogue information on techniques related to membrane systems in a way that is readily accessible by cluster members during the flagship project and the broader community at the end of the project. The technical development is led by the University of New South Wales (UNSW) through the University Library and the UNESCO Centre for Membrane Science and Technology, who also see the environment as a test-bed for developing research infrastructure and extending curatorial library services to the area of research data. The VRE combines a robust repository system with a wiki.

Origin and motivations

The sustainable management of water resources is a question of high importance in Australia and the government invests heavily in research in this field, for instance through the CSIRO Water for a Healthy Country Flagship. Following the establishment of a new research cluster on membrane technologies for water treatment, researchers at UNSW contacted the library with the request to build a digital library that would combine a repository for publications and research data with a wiki for building a collaborative 'Membrane Dictionary'. From the researchers' perspective, using this would address several issues. First of all it would accelerate the dissemination of information within the project. Secondly, it would enable cluster members to jointly build a digital library of techniques relating to membrane systems for the use of the researchers. And thirdly, the research environment could easily be transformed into a publicly accessible resource after the project duration, providing an outreach and dissemination platform that could still be expanded. The library, on the other hand, is very interested in working with research communities to establish services for data management throughout the whole lifecycle of research. It sees participation in this project as part of a strategic vision to rethink recent changes in information management and research practice on a broader level and consider how they need to be brought together. Furthermore, the library also thinks of the project as a test-bed for development of components of a research infrastructure that can be shared across projects. While the library is developing a limited number of e-research systems and VREs in different fields of research, it does not have the resources nor the interest to develop and support many different systems for separate projects.

Features and technology

The Fedora Commons repository system forms the base of the digital library within the MemRE environment. It allows the researchers to upload both publications and research data to the library, where others can search and browse the uploaded materials. The wiki component of MemRE allows collaborative editing of the Membrane Dictionary, to which media files can be uploaded too. Both components are connected through a single sign-on to form an environment. Currently, there are 35 articles in the wiki, while the repository holds in excess of 1500 materials, including articles, conference presentations, research data and digitised materials relating to membrane technology. The collection will eventually bring together literature from 30 years of membrane research, including grey literature such as conference proceedings and workshop materials.

User community and development

The research cluster that initiated the project already had a clearly defined set of requirements, especially as the researchers at UNSW had a keen interest in e-research. The repository team at the library and one of the researchers then jointly developed the concept for MemRE and presented it to the research cluster, who approved it. After that, an iterative development process started that also included some of the early career researchers in the cluster who were expected to do a lot of the work relating to the VRE. Close cooperation between researchers and the development team was seen as an important success factor for the project. Another key factor was that the UNSW principal investigator in the research cluster took a keen interest in the project and is now championing e-research methods amongst the researchers at the university. MemRE was eventually launched about a year ago and has now over 50 registered members from the research cluster. The virtual collaboration within this community is mostly indirect, as the main purpose of the VRE is to share information on outputs and activities of the different groups in the cluster. The research mostly takes place within the separate research groups, with MemRE being the place to collect the outputs and work on the dictionary etc. that brings them together.

Ethical and legal issues

One of the aims of the project is to create a publicly accessible digital library of 30 years of membrane research; to be able to do that the project has to clear the copyright for many publications and get permission to publish digitised materials.

Research outputs

MemRE holds a variety of research outputs already and will be expanded considerably, however these outputs come from the 'normal' work of the research cluster as the VRE itself is mostly used for publication and dissemination. It has had an impact on the work of the cluster insofar as it improved dissemination of results and the speed of communication.

Future plans

At the end of the funding for the research cluster, the materials in MemRE will be made publicly accessible to create a digital library for the wider community. Funding for further technical development of MemRE is currently under negotiation. Plans include: further development of the metadata schema, especially in relation to the wiki as that does not yet have the same level of metadata as the repository; application of text mining techniques to help users to find related materials; developing a way to automatically feed readings from instruments into the repository instead of having to upload them manually.

Sustainability

Both the researchers and the library are interested in continuing to use and support MemRE after the cluster funding comes to an end. The system is will continue to be supported and the library sees MemRE as a test-bed for developing infrastructure.

Lessons learned

- From the library's perspective, work on this project was simplified because the researchers had a clear idea of what they wanted.
- Requirements from the researchers should drive the development of VREs and research infrastructure projects in general.
- The line between a digital library, a repository and a VRE can be blurry.
- Ideally, libraries and information specialists should be involved in research processes insofar as to get a better understanding of how researchers work so that they can support them better.
- Institutional and academic champions, especially senior academics, are crucial for the uptake of e-research infrastructure.

5.2.8 MyExperiment

Region	Europe (UK)
Main project partners	University of Manchester, University of Southampton
Subject domain(s)	Currently, mainly biomedical, but developing in the domains of chemistry, social sciences and humanities
Technologies	Workflow management system
Project started	2007
Funded by	JISC
Website	http://www.myexperiment.org/

Introduction

myExperiment is a web-based resource supporting the sharing of workflows and related resources for scientific experiments. The resource exploits many Web 2.0 features in order to harness the power of user-generated content to provide interactive and contextual background required for meaningful sharing of workflows to occur.

Origin and motivation of the project

The project originated from a previous project, Taverna, which developed a tool with which scientists could build, run and edit workflows. It soon became clear that scientists also needed to share workflows, but also that shared workflows circulate through social networks and communities. The approach taken by the project was to put in place several social networking features, and many Web 2.0 principles: reuse, user-generated content, interactivity, responsiveness to user needs (the perpetual beta) and ease of use.

Features and technology

myExperiment has used a federated approach, so that scientists need not deposit their workflows if they do not wish to do so and can work with a local myExperiment version. There is also a public site, which can be used in order to conduct searches across all available workflows. myExperiment can be integrated into other tools and sites, such as wikis and blogs, and even Facebook. myExperiment also allows users to develop their own applications on top of the service through the use of a RESTful API. Thus myExperiment sees itself as providing functionality rather than being a website.

myExperiment depends on workflows as user-generated content, but also many other forms of user-generated content, such as tags, reviews, ratings, and the forming of groups and friend lists. This results in a high degree of interactivity around workflows, which also makes them more usable.

User community and development

myExperiment currently has more than 2,000 registered users, and growing. It was initially developed for bioinformaticians, but is also being rolled out for chemistry and other disciplines.

Ethical, legal and institutional issues

Evolving out of myGrid, the team in myExperiment has long experience in the ways in which scientists do and do not share data. The reality of career imperatives that are based on article publication means that it is not always in the interests of individual researchers to share data and workflows until they are sure that they have extracted the full value from them. There is therefore

no point in pursuing a policy of completely open access to everything. Instead the approach taken is to reserve some rights:

> *Some Rights Reserved _ users require protection as well as sharing, but the environment must be designed for maximum ease of sharing to achieve collective benefits. Initiatives such as Science Commons provide a useful context for this. Meeting the particular needs of scientists in terms of ownership, attribution and licensing is one of the things that distinguishes myExperiment from other social web sites.*

Future plans

The project has an ongoing Enhancement Work Plan[71] which includes the development of research packs. These packs reflect how workflows are not isolated entities but need to be associated with other workflows, and to data, results, provenance information, tags, documentation, and so on. A semantic web approach is being taken in order to allow for the integration of the different research objects into packs which contain everything that is needed in order to replicate an experiment.

Sustainability

One route to sustainability is to create a larger community using myExperiment and all contributing to its ongoing innovation and feeding back into it. myExperiment is already supporting multiple communities, the biggest being Taverna. One challenge that this brings is that myExperiment was designed in close partnership with specific user groups, and each project that uses it – 6 or 7 until now – adapts it and introduces innovations that are specific to their own user groups. The challenge that this creates is how to bring innovation back in and to see how all of these different innovations feed into the future of the project. The project is not considering advertising as a route to sustainability, although there have been discussions about business models. Another route to sustainability is via the tie in with institutional repositories, and repository integration which is what the project is currently working on. There is a planned roll out of new features which have to do with curation. The project has also had engagements with Microsoft which is trying the myExperiment software; it is also being used by Meander, which is a big humanities community in the states.

Lessons learned

- It is important to think of sustainability not only in terms of software but also in terms of content and social networking.
- It is important to think of sustainability not only in terms of software but also in terms of content and social networking.
- Designing for integration into the existing working systems of scientists.
- Understandings the different needs of scientists with respect to data sharing, and their different motivations.
- Understanding the complex socio-technical nature of a VRE, and appreciating it as a system where technologies and users co-evolve.

71 Further details can be found here: http://wiki.myexperiment.org/index.php/EnhancementWorkplan [Accessed 18/12/2009].

5.2.9 NceSS: Sakai as a VRE for Social Science

Region	Europe (UK)
Main project partners	Manchester eResearch Centre (MeRC), Science & Technologies Facilities Council (STFC) Daresbury Laboratory
Subject domain(s)	Economics and social sciences
Technologies	Sakai; Portal frameworks and portlet technologies; Web 2.0 applications
Project started	2007
Funded by	Several sources, most notable the ESRC (UK Economic and Social Research Council) e-infrastructure project
Website	http://portal.ncess.ac.uk/

Introduction

Until it was discontinued in 2009, the National Centre for e-Social Science (NCeSS) in Manchester was operating a Hub of social science research projects across UK universities. To support the Hub and other projects, a VRE based on the Sakai platform was set up to support exchange of data and project management and to serve as a repository. In addition, it was/is used in a more wide-ranging exploration of what e-research tools and infrastructure could be provided for social scientists, through a set of projects taken forward jointly between NCeSS and the STFC Daresbury Laboratory. The VRE, or portal as it is called, is now used for projects of MeRC, the successor of NCeSS in Manchester, and continues to be supported by STFC Daresbury. Feedback generated through the project also contributed to the general development of the Sakai platform.

Origin and motivations

Initially, the VRE was developed to support collaboration in the ESRC funded e-Infrastructure project, for which a virtual environment was needed that could serve as a repository of project data and support project management and communication. Once the platform was established, it was also increasingly used to support collaboration within the geographically and institutionally dispersed NCeSS nodes and research projects. Within this conglomerate, some projects had special requirements for e-research tools, particularly relating to modelling and visualisation. Funding through the e-Infrastructure project allowed NCeSS to take this forward, using the Sakai experience of a team at the STFC Daresbury Lab. Together, the partners set out to develop a VRE for the NCeSS; explore what e-infrastructure could usefully be developed for social science research; and contribute to the general development of the Sakai platform.

Features and technology

The NCeSS portal builds on the Sakai platform that, despite its origins in e-learning, offers a range of tools that can be used for research collaboration as well (such as wikis, blogs, calendar, forums, repository, scheduling, RSS news etc.). The project uses several of these, notably the wiki. The NCeSS portal also offers specific work sites for research projects. As Virtual Organisations, these sites allow fine grained access permissions to control access to materials. Some tools were specifically developed, or at least adapted, with a particular focus on simulation and modelling. In addition to Sakai, some of the NCeSS nodes also use their own Web 2.0-style applications as tools with a suitable API can be difficult to integrate into a portlet framework. The repository holds a wide range of documents, the number of documents varies between the various worksites.

User community and development

The NCeSS portal continues to serve a community of social scientists (including related fields such as geography) and computer scientists in several projects. Apart from the e-Infrastructure project, the group mainly consisted of members of NCeSS with the NCeSS hub using it as a daily collaboration tool, but it also came to include project partners. Usage slowly expanded as people who had used the portal for one project then also used it for another one. Some 340 users are currently registered on the site. The number of active users fluctuates with the number of active projects the portal supports and it is currently about 70. The development of the system started with a consultation and an analysis of the features needed, including an analysis of the roles people might have within a project and how they accessed content. Other means of gathering requirements were also used, such as surveys. Implementation and further development of the system progressed through a dialogue between the researchers and the development team and continues that way. The way in which the system has been and is used depends on the respective project and the individual researchers. There are around 60 active Virtual Organisations or work sites, with activity moving from older to newer sites as projects evolve. The main usage is in project management, research and on-line training, with data sharing a particular focus. For some researchers the functionalities the system provides are very useful for their daily practice, while others sometimes prefer to use more specific tools for a particular purpose (such as Dropbox[72] for more instant sharing of files with partners for writing a report). The key features for users are the repository and the wiki the portal provides, but collaboration and communication also frequently happens without the environment (email is still predominant across the board). The portal has become an integral part of work practices in the community it supports.

Ethical and legal issues

As the portal was almost exclusively used internally, there were no legal issues as such. Even so, some of the demographic data used in NCeSS projects could not be put into the repository because of the strict policies regulating their use.

Research outputs

The use of the portal has contributed to several research projects, mostly through facilitation of collaboration and sharing of information and data. It could be argued that the various lessons learned through the software development constitute research outputs in themselves, including documentation, strategy and requirement documents and training material

Future plans

MeRC and STFC Daresbury are involved in several Sakai related projects, including the JISC funded OneVRE to Join Them All project[73] that looks at the integration of VREs and as one part at the integration of Access Grid audio-visual collaboration into Sakai and the NeISS (National e-Infrastructure for Social Simulation) project.[74] Further development of Sakai as a platform will happen through these and possible future projects. Additional tools and services, for instance relating to workflows and simulation, are under consideration.

Sustainability

Despite the demise of NCeSS in September 2009, the newly formed MeRC and STFC Daresbury continue to support and maintain the Sakai installation. As the environment is still embedded in several active projects, it will be supported until 2011. Beyond that the sustainability depends on its use in future projects.

---

72 http://www.dropbox.com/ [Accessed 18/12/2009].

73 http://www.rcs.manchester.ac.uk/research/OneVRE [Accessed 18/12/2009].

74 http://www.neiss.org.uk/ [Accessed 18/12/2009].

Lessons learned

- A platform such as Sakai offers a wide range of basic functionalities for a VRE, but it may still have to be adapted to suit the needs of a specific community even as these more general features are concerned.
- Even then the use of such an environment may not suit everyone and at least some researchers may prefer to use external tools and applications that are specifically focussed on one particular task or that suit their way of working.
- Users expect an ease of use for VRE projects that is based on their experience of using commercial web platforms developed at a much higher cost than academic projects. Even so, usability appears to be the key factor for the uptake of an environment.
- Sharing and archiving of data is a key request that researchers make for VREs; how far 'just' a repository can address researchers' needs depends on the particular usage, but a VRE can provide history and context of how the data was created, which may be very valuable.
- Project teams should be aware that because of both personal preferences and specific requirements researchers may want to continue to use tools and services outside of a VRE instead of the ones provided through the VRE; also, not everything can be integrated. As long as the use of external applications works in a project context, it should be accepted.

5.2.10 Noguchi Memorial Institute for Medical Research (NMIMR) and Virtual Research Collaboration in West Africa

Region	Africa (Ghana)
Main project partners	Noguchi Memorial Institute for Medical Research (NMIMR)
Subject domain(s)	Medical sciences
Technologies	General website and wiki environment
Funded by	Mostly own resources
Website	http://www.noguchimedres.org/ (VRE not publicly accessible)

Introduction

Based at the University of Ghana, Legon, a suburb of the Ghanaian capital city Accra, the Noguchi Memorial Institute for Medical Research (NMIMR) is a centre for biomedical research with a particular focus on diseases that have public health importance. To support its staff of 250 as well as partners in international collaborations, the Institute is developing a virtual environment that, initially, aims to facilitate project planning and access to data. This case study looks into the challenges of providing virtual research support in a part of the world that has comparatively limited access to communication infrastructure.

Origin and motivations

There are two main motivations for setting up a VRE at the Noguchi Memorial Institute. Firstly, staff in the rapidly growing institute need a space to share documents and data amongst themselves and also to access administrative documents and materials that support them in developing grant proposals. Secondly, the Institute is engaging in national and international collaboration, involving countries in Europe as well as North and South America, but staff neither have the time nor budget to constantly travel to meet partners. The Institute also hosts several centres that have an international remit and rely on being able to work collaboratively, such as the West African Centre for Parasite Control (WACIPAC) and the Lymphatic Filariasis Support Centre for Africa (LFSCA). In Africa alone, the collaboration encompasses a dozen countries. For all of this research, tools for communication, sharing of data, project management and also collaboration are needed and actively requested by the researchers.

Features and technology

Currently, the institute uses a combination of well established technologies, including proprietary software such as Microsoft Outlook and, at least on an individual basis, Skype. The website also serves as a data storage, with a section for staff to access and share documents currently being set up. This storage facility will also support backing-up of crucial project data from staff computers. Access to bandwidth proves to be a critical factor for all online activities at the Institute. Files attached to email have to be very limited in size, video connections via Skype are not possible at all, while audio cannot be used during normal office hours because of bandwidth issues. Staff at the Institute commented: 'We're really hungry for bandwidth and speed'. The Institute is connected to the internet via satellite, but the connection can be affected by weather. Not unusually in Africa, most telephone communication takes place using mobile phones. The communication infrastructure dictates that the website has to be hosted with a European hosting company. As a next step, wiki functionality to support collaboration is being considered. However, even more limited bandwidth at some of the African partner institutions of the NMIMR heavily restrict the possibilities for further development, including technically relatively simple ways of sharing data.

User community and development

The impetus for developing support for virtual collaboration at the NMIMR partly came from the administration, as it wanted to make documents and support for grant writing and project management available online so that it can be accessed by staff, including remote access (although not for all types of documents), and selected partners. Communication and sharing of data are also the two key requests from researchers and they are taking an active interest in this. Especially because of the collaborative and international nature of the work at the NMIMR, researchers are well aware of developments in other countries and require access to features such as synchronous audio-visual communication. Access to such facilities is seen as crucial to save both time and money on travel. Researchers already participate remotely in events, as far as their connection allows, for instance in regular webinar online seminars with colleagues at Columbia University in New York. The Institute could not host something like this though and they can only participate via audio, not video; most of its African partners have even more limited resources. While academic institutions in African capital cities usually have broadband access, mobile phones, limited wireless and dial-up connections dominate in other areas.

Ethical and legal issues

The Institute is currently considering setting up a wiki environment to facilitate collaborative biomedical research in West Africa, but partners in other countries are concerned that data might be stored on servers in foreign countries. For this reason, cloud computing is also seen as problematic, especially as far as sensitive medical data from governments is concerned. Also, accessing resources that theoretically are freely available on the internet can be an issue in less developed countries where no reliable broadband is available.

Research outputs

Virtual research support at the NMIMR is in an early stage and is initially focussed on supporting grant applications and project management, for reasons outlined above.

Future plans

Implementing a wiki environment for sharing data and collaboration between partners in West Africa is currently under consideration; further development depends on access to funding and bandwidth.

Sustainability

The NMIMR uses a conservative approach for development as it does not want to implement features or systems it cannot support in the longer term. So far, development has mostly been funded through its own resources, but some external funding was available too. As support and funding in Ghana is limited, partnerships and founding foreign funders such as the European Union (especially Framework 7) are seen as important.

Lessons learned

- Limited access to bandwidth and computational resources is a key barrier for virtual research collaboration.
- Storing and sharing data as well as communication are key requirements of researchers, while support for grant development and project management are for the institution.
- To facilitate international collaborative research, investment in basic infrastructure in less developed countries is needed to enable more participation.

5.2.11 Orlando Project Document Archive and Canadian Writing Research Collaboratory

Region	North America (Canada)
Main project partners	University of Guelph, University of Alberta
Subject domain(s)	Arts and humanities
Technologies	Repository management system and publication technology
Project started	Orlando: 1996 / CWRC: 2009
Funded by	Orlando Project: Major Collaborative Research Initiative (MCRI) grant from the Social Sciences and Humanities Research Council of Canada (SSHRC); CWRC: Canadian Foundation for Innovation
Website	http://nifflheim.arts.ualberta.ca/wwp/

Introduction

Both of these projects are collaborations between the Universities of Alberta and of Guelph in Canada, including a partnership with the University of Alberta Library. The Orlando Project grew out of the print publication of the *Feminist Companion to Literature in English*[75]. There was much material that had to be excluded from the Companion because of space constraints and it was felt that a web-based publication was one way to do justice to the complexity of relations among women writers, their texts, circumstances and contexts. Over 100 researchers were involved in writing marked-up content for the Orlando Project.[76] Although this is a large number of collaborators, it was a closed community. After it came to an end and was published as an online resource by Cambridge University Press, it was clear that there is a lot of scope for collaborative writing in the humanities. Funding was obtained from the Canadian Foundation for Innovation for The Canadian Writing Research Collaboratory or CWRC. This will be a collaboratory for a variety of writing projects, with the infrastructure clearly distinguished from the different projects that will use it[77].

Features and technology

CWRC, like the Orlando Project, will be a combination of archive and collaborative writing tools, with a stress on using digital resources to create innovative outputs, that are 'born digital' and not only consist of faster or more efficient access to traditional humanities sources. The Orlando Project used existing mark-up tools. CWRC aims to build an open access infrastructure combining a number of resources as needed for archiving and collaborative writing, with Web2.0 components; however, no specific technology has been decided upon yet.

Users, community and research outputs.

The Orlando Project is now part of the Cambridge University Press catalogue and is sold mostly to libraries. CWRC is only just starting out, with one group of early adopters in the form of a network of researchers interested in Canadian women's writing. There is no restriction of topic or

75 BLAIN, Virginia, Patricia CLEMENTS & Isobel GRUNDY, eds. (1991). The Feminist Companion to Literature in English: Women Writers from the Middle Ages to the Present, London & New Haven: Batsford/Yale UP.

76 See Brown, S., et al, Published Yet Never Done: The Tension Between Projection and Completion in Digital Humanities Research, Digital Humanities Quarterly, vol 3, no 2. http://digitalhumanities.org/dhq/vol/3/2/000040.html [Accessed 18/12/2009].

77 http://www.hastac.org/blogs/cathy-davidson/scholarship-20 [Accessed 18/12/2009].

country in CWRC, with the hope that it will be used internationally too. The main research outputs will be publications, both in traditional paper and also new media/digital formats.

Ethical, legal and institutional issues

The Orlando Project consulted a lawyer for advice on copyright and attribution. However, there are ethical as well as legal issues:

> *When people have contributed to a project, it is an ethical imperative that that work gets acknowledged in an appropriate way. One of the challenges for CWRC is how do we, in a more open environment, where people would not necessarily be owning the projects that they' are contributing to, how can they receive appropriate credit for the work that they've done.*[78]

For these kinds of projects, which deal with representations of individuals and groups, there is also a more deep-seated ethical issue in the very subject matter dealt with. Representations of class, gender and race and similar concepts are a substantive issue in literary studies, thus this brings together the content, the form and the infrastructural arrangements (such as authorship) of these collaboratories. For example, the Orlando Project decided against coming up with a preset vocabulary of race, class and gender tags that writers could simply use; instead it encouraged writes to be more self-conscious about the tags they were using and creating in order to highlight these issues. This is just one of the ways in which these issues emerge in a humanities collaboratory.

Sustainability

The sustainability of the Orlando Project has been addressed through the fact that it is a licensed product, owned by Cambridge University Press. As part of the funding from the Canadian Foundation for Innovation, there is built-in four years of continued funding for sustainability. In addition, discussions with the University of Alberta Library to establish a degree of long-term stability and endurance for the project have been positive. Ultimately the open access infrastructure to be developed by CWRC needs to become an indispensable resource so that there is a commitment to it by the communities that use it.

Lessons learned

- VREs can be ideal platforms for larger publication and editing projects, because other than a printed edition the digital publication practically puts no restrictions on the amount of material that can be published and it allows interlinking of content and better representations of complex relationships.
- Scholarly editions, which have been large, collaborative projects for a long time, may be particularly suitable for using social networking approaches within a VRE.
- There is an increasing demand from the users to more open, interlinked Web 2.0-style systems as opposed to digital silos and one-way/single task tools.

78 Interview with Susan Brown.

5.2.12 Research Information Centre (RIC)

Region	Europe (UK)
Main project partners	The British Library and Microsoft
Subject domain(s)	All domains; initially focussed on medical sciences
Technologies	Microsoft Office SharePoint Server (MOSS)
Project started	2007
Funded by	Project partners
Website	http://research.microsoft.com/en-us/projects/ric/ http://www.bl.uk/reshelp/experthelp/science/ric/ric.html

Introduction

The Research Information Centre (RIC) is a virtual research environment framework being jointly developed by Microsoft External Research and The British Library. It will offer an integrated suite of tools for finding, creating, managing, sharing and disseminating all the types of information associated with a research project. The RIC provides a set of core functionalities supporting the whole research lifecycle that can be used to build domain specific VREs, into which additional modules can be added. A particular focus is to support researchers as 'extreme information workers' by providing easy access to relevant information resources. The first specific VRE layer is focussed on biomedical research, with additional layers in preparation, for instance for digital humanities. The RIC is built using the Microsoft SharePoint platform and while SharePoint is necessary to run it, the actual source code will soon be released under an open source license. Much more than building a specific VRE, the aim of the project is to create a VRE development community and to understand better how libraries can support future research.

Origin and motivations

Early on, the British Library engaged with the JISC VRE programme because it wanted to learn more about how such environments would be underpinned by libraries – not just in relation to curation of research outputs – and how this would in turn impact on the libraries. Because of this, the library was looking for a general framework that could support collaborative work throughout the whole research lifecycle. This in turn interested Microsoft, who wanted to explore how the SharePoint platform could be used to build such an environment and to demonstrate the compatibility of commercial Microsoft software and community-developed open source software. Microsoft and the British Library started the RIC as a joint project to develop a virtual research environment framework. RIC aims to reduce the time researchers spend on administrative tasks, to support collaborative research, provide easy access to relevant information, to facilitate networking and to help preserve not only project outcomes, but the whole process of research. A particular focus is to reduce 'pain points', inefficiencies in the knowledge management over the research lifecycle. One aspect of this is that researchers are not necessarily always aware of relevant resources, a problem that could be reduced by providing domain specific access to information.

Features and technology

The RIC is built using SharePoint, Microsoft's collaboration platform, and makes use of a range of out-of-the-box features such as collaboration, document management and search services. While this means that SharePoint is needed to run the RIC, the RIC software itself will soon be released under an open source license and the environment can be accessed through a web browser. The RIC supports a research lifecycle organised around the four phases of idea discov-

ery, funding, experimentation and dissemination. It addresses key areas such as content and knowledge management, social networking and online collaboration. Templates for projects can be created and specific project sites set up based on those templates. The RIC offers a range of features such as access control; workflows; sharing and annotation of resources; RSS feed integration; federated search over domain-specific literature sources and a full-text search over local resources; blogging; wikis; networking; creation of project groups; bibliographical support; and archiving of project sites. Deployed as an institutional VRE environment, the RIC could support the management of projects and facilitate sharing of information across the institution, while providing the researchers with a domain and project specific environment into which additional resources can be added. The prototype VRE for biomedical research, for instance, allows searches over a range of relevant databases and repositories such as UK PubMed Central or the British Library's Electronic Table of Contents database. The RIC can be deployed by an individual project or institution, but it can also be imagined in a shared hosting situation, for instance as a service offered through the cloud.

User community and development

The initial phase of development was guided by previous research, including the BL's experience from participation in the JISC VRE programme, and community engagement mostly through presentations and discussions. In May 2008, after the core features had been developed, the system was opened to a beta testing by 24 groups from the biomedical domain, with some 50 researchers from universities and companies in the UK, the US, Australia and South Africa. The groups were deliberately kept fairly small as members of the project team were personally in contact with all users during the testing phase. Following positive feedback, it was decided in December 2008 to continue the project. Version 1 is due to be released in December 2009. It remains to be seen how the RIC will be used in 'live' research.

Research outputs

As the RIC has not yet been used in an actual research context, there are no research outputs beyond the research aspect of developing the platform.

Future plans

After the release of version 1, the project will focus on creating a sustainable community that can take the development of the RIC forward. To enable this, the software will shortly be released under an open source license and Microsoft is committed to a community development process. The project is also involving the community into the development of domain specific layers and site templates. The biomedical VRE will be expanded, for which the library is consulting with several partners in academia and industry. New domain-specific layers will be set up, the next one being digital humanities, in consultation with Trinity College Dublin. Other collaborations are also ongoing, and one of the projects funded as part of the JISC VRE 3 programme, the Cancer Imaging VRE, will make use of parts of the RIC platform.

Sustainability

Both Microsoft and the British Library will continue to commit resources to the project. The development of a RIC community will be a key factor and further activities will depend on the progress made here.

Lessons learned

- As the concept of a VRE is still fairly new, it can conceptually be difficult for researchers to articulate their needs; because of this it is very beneficial to integrate them into the actual development process so that they can experience an environment and formulate requirements in response to that.

- While it is important to preserve the outputs of projects, a VRE should ideally also capture the process that led to them, so that later on it is possible to understand how certain ideas came to exist.
- Many aspects of information management are generic and can be shared across domains; however, the domain-specific parts add real value to a VRE and need to be developed in close collaboration with the researchers.
- In relation to information management, VREs have the potential to reduce time spent on administrative tasks for the individual researcher and also to make data management easier for the institution.

5.2.13 TextGrid – Virtual Research Environments in the e-Humanities

Region	Europe (Germany)
Main project partners	Göttingen State and University Library, coordinator of a consortium of ten partner institutions in Germany
Subject domain(s)	Arts and humanities
Technologies	Grid; Web Services
Project started	2006
Funded by	BMBF, the German Federal Ministry of Education and Research
Website	http://www.textgrid.de/

Introduction

TextGrid is a large VRE development project taken forward by a consortium of ten German libraries and higher education institutions, coordinated by SUB, the Göttingen State and University Library. TextGrid is one of the first grid-based humanities projects and aims to create an infrastructure for the collaborative editing, annotation, analysis, and publication of textual resources. The project was initiated in response to a funding call from the German D-Grid initiative that aims to develop a 'distributed, integrated resource platform for high-performance computing and related services to enable the processing of large amounts of scientific data and information'.[79]

Origin and motivations

The SUB had a key role in taking the project forward, as it was already involved with D-Grid. The idea for the project, however, came from a group of researchers working in the field of linguistics, languages and literature and the requirements of the research community have driven project development from the start.

The research community involved with TextGrid had certain requirements to start with, partly through their use of the TUSTEP software, a suite of text processing tools that was widely used.[80] The SUB partly saw its role in translating these requirements into a workable development plan that made use of existing standards and cutting edge technologies. However, the library did not see itself as purely an infrastructure provider, but it also has an interest to understand the changes in research practice to be able to better support the research of the future. The researchers were interested in tools that would support their whole research lifecycle in a shared environment, which also supported collaboration across institutions. An important aspect of this was that the VRE should combine tools and data with an infrastructure that would address the issue of preserving the research outputs and data.

Features and technology

The backbone of the TextGrid architecture is the TextGrid Repository, a large (hundreds of terrabytes) grid-embedded long-term archive for research data; this part of the project aims to ensure long-term availability and access to research data as well as interoperability. On top of that infrastructure sits the TextGrid Laboratory, a single point of entry to the virtual research environment, which will provide integrated access to both new and existing tools and services. TextGridLab is currently in beta and can be downloaded from the website. It includes components such as an XML editor, Lemmatizer, Tokenizer, Metadata Annotator, a Dictionary Search Tool and facilities for project and access control management. The architecture was developed to be as flexible

79 http://www.d-grid.de/index.php?id=1&L=1 [Accessed 18/12/2009].

80 http://www.zdv.uni-tuebingen.de//static/skripte/tustep/ [Accessed 18/12/2009].

and standards compliant as possible to ensure interoperability; the software is OpenSource. This is important as TextGrid provides the technical infrastructure into which services and tools can be integrated. Project development has been informed by the idea that it is possible to build both a data management infrastructure and generic modules that can be used across subject disciplines, but that a VRE needs to be flexible enough so that discipline specific tools can be integrated and others customised in order to be of use for specific research tasks.

User community and development

Customisation of tools and services is also important because the TextGrid community has been expanded since the project started and the new subject disciplines such as Musicology have their own tools and ways of working that need to be integrated into the project. Many members of the original community were early adopters with a strong interest in digital humanities and the project will have to expand beyond those groups to include more researchers who do not have a particular interest or agenda with regards to digital technology. Increasing the stability and usability of the TextGrid software is an important aspect of this work and the project will only start to systematically expand the user community once the environment is seen as stable enough. Even so, TextGrid has already hosted various events to engage with the community, including workshops focused on user testing. Because of the current stage of the development process, TextGrid is not yet fully embedded into the everyday work of the participating researchers.

Future plans

Apart from adding new features (such as an OCR module for Gothic characters) and increasing stability of the software, developing plans for the sustainability of TextGrid are key aspects of the second project stage (funding until 2012). Additional features will also partly be supported through new funding bids, such as an application to build communication facilities into TextGrid.

Sustainability

As far as funding is concerned, TextGrid approaches sustainability on several levels. Members of the TextGrid consortium have agreed to support specific parts of the project with their own resources. This not only includes research data generated through the project, but also through the environment. The Competence Centre for Electronic Text Processing and Publication in the Humanities, hosted by the University of Trier, for instance, is particularly committed to addressing the sustainability of TextGrid tools. The consortium is also working with funders to develop concepts for sustainability. Furthermore, TextGrid is part of 'WissGrid, Grid for Science'[81], an initiative whose aim is to establish long-term organisational and technical grid structures for the academic world. Being part of this interdisciplinary group allows TextGrid to represent the humanities in the Science dominated grid environment and ensure sustainability through sharing of resources.

Lessons learned

- a dialogue between developers, researchers and information managers should start with or even before the planning of a funding bid
- even with such a dialogue in place, it takes time for all the participants to develop a shared language
- institutional champions, particularly senior academics, are important for the acceptance of a VRE project
- development needs to be driven by the needs of the users
- it is not only possible, but also beneficial (especially in terms of sustainability) to develop a shared infrastructure across disciplines

81 http://www.wissgrid.de/ [Accessed 18/12/2009].

- a VRE as a framework may be developed as a generic platform, but the architecture needs to be flexible so that specialist modules suited to the needs of specific disciplines can be embedded and customised
- even universities and libraries with technical expertise should not underestimate the challenges of larger scale development projects, for which they will have to learn new skills and find staff who can bring together the academic and technical worlds

5.2.14 Virtual Knowledge Studio Collaboratory

Region	Europe (Netherlands)
Main project partners	Virtual Knowledge Studio
Subject domain(s)	Mostly arts and humanities and social sciences
Technologies	Microsoft SharePoint Server (MOSS)
Project started	2008
Funded by	SURFfoundation
Website	http://www.virtualknowledgestudio.nl/

Introduction

The VKS Collaboratory is a VRE for collaboration between staff and partners of the Virtual Knowledge Studio, a group based in the Netherlands that both studies and supports researchers in the humanities and social sciences in the creation of new scholarly practices such as e-research. The Collaboratory is a further development of SURFgroepen, a collaborative platform based on Microsoft SharePoint and supported by SURFfoundation, a Dutch funding body.

Origin and motivations

Distributed over several locations in the Netherlands and with a remit to work collaboratively across institutions and subject domains, the VKS needed a virtual environment for collaboration, communication and sharing of data both for internal and external projects. Such a system was even more important as some staff are also tele-workers that rely on communication infrastructure, and the set of individual tools used by VKS staff were not seen as ideally suited. The VKS responded to a call by SURFfoundation, who were interested in further testing, evaluation and development of SURFgroepen, a collaborative platform built on Microsoft SharePoint. This led to the VKS implementing and testing the system for their own use. At the same time, the VKS also had a second Collaboratory project, which had the subtitle: Understanding Scholarly Collaboration in Practice; it was focused on the psychosocial aspects of online scholarly collaboration in the humanities and social sciences.

Features and technology

The Collaboratory is mainly used to support sharing of digital objects and the coordination of projects. A web-based environment, it has features such as fine-grained access control, online discussion and limited interfacing to outside resources. The design process at the VKS has mostly been dominated by deciding which features not to use in order to ensure the environment is more user friendly. While SharePoint proved to be a potentially very powerful platform, working on it was sometimes a little heavy going, because of the many features and configuration options. While most relevant aspects of collaboration are now supported through the Collaboratory, VKS staff are still using Google Apps for time coordination and scheduling tasks. In a similar way, Zotero appears to be the first choice of researchers for reference management and because of that similar functionality has not been included in the VKS Collaboratory.

User community and development

The Collaboratory is used by all full-time staff of the VKS as well as by some of the partners and external researchers, numbering about 50 people. VKS already have experience with interdisciplinary work and used online tools beforehand, so a set of requirements existed, which were augmented by a research and communications plan. Overall development was content and task driven and an iterative process was used. Staff were involved in the development of the Collab-

oratory, for instance through workshops and the chance to explore the system during development. The VKS continued to encourage staff to use the system, but there wasn't a specific roll-out plan as it was also felt that the system had to prove itself by providing functionality that the users would actually want to make use of. Not only the discussions on the development of the VKS, but also project tasks such as the preparation of a book took place within the Collaboratory. Through this approach, the system became progressively integrated into the normal workflow of VKS to a point where it is now regularly being used by all staff. Despite what was seen as a comparatively good integration of the staff in the development process, more dedicated on-location support would have been welcomed to help staff to formulate questions for the help desk in the initial phase of the project. While collaboration so far is seen as working well, the use of the Collaboratory differs between disciplines and certain ways of working: even apparently fairly simple functionality such as sharing draft documents can have different meanings: in some disciplines it is more a formality, whereas in others it is part of a complex reflection process that is more than writing up an analysis.

Ethical and legal issues

The implementation of the Collaboratory has been accompanied by regular reflections on legal and especially ethical issues. For instance, staff were discussing what documents could and should not be shared with others, and at what stages – and whether comments that others make on these drafts should be shared too. As the VKS is a strong supporter of Open Source software, the use of proprietary technology was also seen as problematic. A legal issue was the integration of resources such as repositories or other library databases: because of the distributed nature of the VKS, staff at different locations would not have had permissions to access the same resource.

Research outputs

The VRE is used to support the ongoing research at the VKS; as a project in its own right, it generated several resources, including a study of the literature and websites relating to the organisational, social and technological components of collaboratories; generic and specific adaptations of SharePoint; a rights management and permissions document; and a report on the psychosocial aspects of online scholarly collaboration in the humanities and social sciences.

Future plans

While the first project phase was mostly about development, the second phase focussed more on the roll out of the system, which has just been finished. Further plans are under discussion.

Sustainability

The second phase of project funding through SURFfoundation is just about to end. One of the deliverables will be a sustainability plan that also links in with the general plan for the next phase of VKS funding.

Lessons learned

- Online collaboration is first and foremost a social and organisational problem. Technical issues are of less importance.
- Dedicated, on-location support is very helpful for users, especially during the phase of first contact with a system.
- Reducing functionality to key tasks increases usability.
- Doing a few things very well may be better than trying to do too much.
- Having a focus on tasks instead of functionalities helps to develop a VRE that is actually used. However, it is important to be aware that the same task can actually be approached in a very different way across different disciplines.

- Access to database and information resources is often very fragmented; at least in terms of licenses this should be approached on a greater, national level as this access needs to be seen as part of the research infrastructure.
- Generic social network tools are not always of high interest to researchers; in smaller groups such as the VKS a traditional website listing projects can work as well.
- Rights management is a crucial success factor.
- As with every community, a virtual community needs a set of rules about behaviour and human interaction.
- Personal attention is not sufficient in stimulating use. Commitment of the team leaders is also important
- An essential is the appointment of a collaboratory management team different from the scientific staff involved.
- Physical meetings are essential for the success of online collaboration.

5.3 Appendices

5.3.1 Interviewees

1. Robert Allan, STFC e-Science Centre, Daresbury Laboratory, United Kingom
2. Stephen Andrews, British Library, United Kingdom
3. Anne Beaulieu, Virtual Knowledge Studio, Netherlands
4. Ann Borda, VeRSI, Australia
5. Susan Brown, Orland Project and Collaborative Writing project
6. Natasa Bulatovic, Max Planck Digital Library, Germany
7. Lou Burnard, Adonis, France / United Kingdom
8. Tim Clarke, Harvard Medical School and Massachussets General Institute for Neurodegenerative Disease, USA
9. Marin Dacos, Cleo, France
10. Martie Deventer, CSIR, South Africa
11. John Doove, SURF, Netherlands
12. Stefan Dormans, HubLab, Netherlands
13. Sigrun Eckelmann, Deutsche Forschungsgemeinschaft (DFG), Germany
14. Maude Frances, University Library, University of New South Wales, Australia
15. Gudmund Høst - Special adviser in the Research council of Norway
16. Jan Kok, HubLab, Netherlands
17. Cherie Ann McCown, Noguchi Memorial Institute for Medical Research, Ghana
18. Reagan Moore, Data-Intensive Computing San Diego Supercomputer Center (SDSC), United States
19. Heike Neuroth, Goettingen State and University Library, Germany
20. Luca S. Paderni, Google, United Kingdom
21. Meik Poschen, Manchester e-Research Centre, United Kingdom
22. Rafaella Santucci, Sapienza University of Rome, Italy
23. Jennifer M. Schopf, Office of Cyberinfrastructure, National Science Foundation, United States
24. Ulla Tschida, Max Planck Digital Library, Germany
25. Alex Wade, Microsoft External Research, United States
26. Nancy Wilkins-Diehr, Science Gateways, United States

5.3.2 Glossary

Adonis	Accès Unifié Aux Données et Document Numériques des Sciences Humaines et Sociales
Alzforum	Schizophrenia Research Forum
ANDS	Australian National Data Service
API	Application programming interface
ARCS	Australian Research Collaboration Service
BMBF	German Federal Ministry of Education and Research
BOS	Bristol Online Survey
BVREH	Building a Virtual Research Environment for the Humanities
Cléo	Centre pour l'édition éléctronique ouverte
CNRS	Centre national de la recherche scientifique
CSIR	Council for Scientific and Industrial Research
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CVE	Collaborative Virtual Environment
CWRC	Canadian Writing Research Collaboratory
DARE	Digital Academic Repositories
DFG	Deutsche Forschungsgemeinschaft
DPI	Department of Primary Industries
e-RC	e-Resource Centre
ESFRI	European Strategy Forum on Research Infrastructures
ESRC	Economic and Social Research Council
FIZ	Fachinformationszentrum Karlsruhe
HIPPA	Health Insurance Portability and Accountability
HSRC	Human Sciences Resources Council
IBVRE	Integrative Biology VRE Project
IdP	Identity Provider
IISH	International Institute of Social History
IPR	Intellectual Property Rights
IRODS	Integrated Rule-Oriented Data System
ITEM	Institut des Textes etManuscrits Moderne
JISC	Joint Information Systems Committee
KCL	King's College London
LFSCA	Lymphatic Filariasis Support Centre for Africa
LONI	Louisiana Optical Network Initiative
MCRI	Major Collaborative Research Initiative
MemRE	Membrane Research Environment
MeRC	Manchester eResearch Centre

MOSS	Microsoft SharePoint Server
MPG	Max Planck Gesellschaft (Max Planck Society)
NCAR	National Center for Atmospheric Research
NceSS	National Centre for e-Social Science
NCI	National Computational Infrastructure
NCSA	National Center for Supercomputing Applications
NeAT	National eResearch Architecture Taskforce
NeISS	National e-Infrastructure for Social Simulation
NICS	National Institute for Computational Sciences
NMIMR	Noguchi Memorial Institute for Medical Research
NorGrid	Nordic Data Grid
ORNL	Oak Ridge National Laboratory
PD	Parkinson's Disease
PSC	Pittsburgh Supercomputing Center
RIC	Research Information Centre
SAMI	South African Malaria Initiative
SDSC	San Diego Supercomputer Center
SNDC	San Diego Supercomputer Centre
SOA	Service Oriented Architecture
SSHRC	Social Sciences and Humanities Research Council of Canada
STFC	Science & Technologies Facilities Council
SWAN	Semantic Web Applications in Neuromedicine
TACC	Texas Advanced Computing Centre
TGE	Très Grands Équipements
UNSW	University of New South Wales
VCCAP	Victorian Climate Change Adaptation Program
VeRSI	Victorian eResearch Strategic Initiative
VKS	Virtual Knowledge Studio
VO	Virtual Organisation
VPN	Virtual Private Network
VRC	Virtual Research Community
VRE	Virtual Research Environment
WACIPAC	West African Centre for Parasite Control

5.3.3 References

- Blain, V., Clements, P & Grundy, I. eds. (1991). The Feminist Companion to Literature in English: Women Writers from the Middle Ages to the Present, London & New Haven: Batsford/Yale UP.
- Brown, S., et al, Published Yet Never Done: The Tension Between Projection and Completion in Digital Humanities Research, Digital Humanities Quarterly, vol 3, no 2. http://digitalhumanities.org/dhq/vol/3/2/000040.html
- Clark T, Kinoshita J. (2007) Alzforum and SWAN: the present and future of scientific web communities. Brief Bioinform. 8(3):163–71.
- De Roure, D., Goble, C., Stevens, R. (2009) The design and realisation of the myExperiment Virtual Research Environment for the social sharing of work flows. Future Generation Computer Systems. 25, 561-567.
- Kok, J. (2008): HubLab: Towards online collaboratories for global data gathering in social and economic history. Final report. IISG - Internationaal Instituut voor Sociale Geschiedenis.
- Page-Shipp, RJ, Hammes, MMP, Pienaar, H, Reagon, F, Thomas, G, Van Deventer, MJ, Veldsman, S 2005. 'e-Research support services: responding to a challenge facing the South African research and information communities', South African Journal of Information Management, vol. 7, no. 4. [http://www.sajim.co.za/]
- Pienaar, H. and van Deventer, M. (2009) To VRE or Not to VRE?: Do South African Malaria Researchers Need a Virtual Research Environment?, Ariadne, Issue 59, http://www.ariadne.ac.uk/issue59/pienaar-vandeventer/.

5.3.4 Literature Overview

- Allan, R. (2009) Virtual Research Environments. From portals to science gateways. Oxford: Chandos Publishing.
- Anderson, S., Dunn, S. et Hughes, L.M. (2006) VREs in the Arts and Humanities. Proceedings of the UK e-Science All Hands Meeting, Nottingham.
- Barabasi, A.L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A. et Vicsek, T. (2002) Evolution of the social network of scientific collaborations. Physica A 311, 590-614
- Barjak, F. (27 October 2007) M4 Final Report: Achieving transition to Virtual Research Organisation in Social Science (AVROSS). Issued by Information Society and Media Directorate General, Commission of the European Communities
- Berman, F., Fox, G.C. et Hey, A.J.G. (2003) Grid computing: Making the Global Infrastructure a reality. Chichester: J. Wiley and Sons.
- Blanke, T. (2008) From tools and services to e-Infrastructure for the arts and humanities. Proceedings of the International Symposium on Grid Computing (ISGC). Taipei: Springer.
- Blanke, T. et Dunn, S. (January/February 2008) Next Steps for E-Science, the Textual Humanities and VREs A Report on Text and Grid: Research Questions for the Humanities, Sciences and Industry, UK e-Science All Hands Meeting 2007. D-Lib Magazine Volume 14 Number 1/2.
- Bos, N., Zimmerman, A., Olson, J., Yew, J., Yerkie, J., Dahl, E., et al. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. Journal of Computer-Mediated Communication, 12(2). http://jcmc.indiana.edu/vol12/issue2/bos.html.
- Büscher, M., Mogensen, P., Agger Eriksen, M. et Friis Kristensen, J. (2004). Ways of grounding Imagination. Proceedings PDC (Participatory Design Conference) 2004.
- Carusi, A. et Jirotka, M. (forthcoming) Reshaping Research Collaboration: the case of Virtual Research Environments. In Dutton, W.H. and Jeffreys, P.W. (forthcoming) World Wide Research: Reconfiguring Access to Information, Expertise and Experience in the Sciences and Humanities. Cambridge, MA: MIT Press
- Carusi, A. et Jirotka, M. (28-30 June 2006) Building Virtual research Environments and User Engagement. Oxford University Computing laboratory. Proceedings of the 2nd International Conference for e-Social Science, Manchester University.
- Carusi, A. et Jirotka, M. (2009) From data archives to ethical labyrinths. Qualitative Research, vol 9, no 3, 285-298.
- Churchill, E.F., Snowdon, D.N. et Munro A.J. (eds.) (2001) Collaborative Virtual Environments: Digital places and spaces for interaction. London: Springer-Verlang.
- Clark, T. et Kinoshita, J. (17 March 2007) Alzoforum and SWAN: the present and future of scientific web communities. Briefings in Bioinformatics, vol. 8 n. 3.
- Cox, A. (2004) Building collaborative e-research environments. University of Loughborough technical report compiled for JISC.
- Dalle, J.M, David, P.A., Ghosh, R.A. et Wolak, F.A. (June 2004) Free and Open Source Software Developers and "the Economy of Regard": Participation and Code-Signing in the Modules of the Linux Kennel. OWLS: Oxford Worshop on "Libre Source". Oxford Internet Institute.
- OSI e-Infrastructure Working Group. Developing the UK's e-infrastructure for science and innovation. http://www.nesc.ac.uk/documents/OSI/ (Accessed 18/12/2009)

- Dormans, S. (March 2009) Collaboratories: from natural sciences to social sciences and humanities. Literature review for HUBLAB2: Toward successful implementation of the Liferay platform in historical research.
- Dunn, S. (February 2009) Dealing with the complexity deluge. VREs in the Arts and Humanities. Library Hi Tech, special issue on Virtual Research Environments, edited by Judith Wusterman.
- Dunn, S. (March 2007) Trustworthy Characters: common issues for archaeology, classical studies and VREs. Proceedings of First International Workshop on VREs, Edinburgh.
- Ehn, P. (1988) Work Oriented Design of Computer Artifacts. Arbetslivscentrum. Stockholm.
- Ehn, P. et Kyng, M. (1992) Cardboard computers: mocking-it-up or hands-on the future, Design at work: cooperative design of computer systems. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.
- European Strategy Forum on Research Infrastructures (2008) European Roadmap for Research Infrastructures. Luxembourg: Office for Official Publications of European Communities.
- Finholt, Thomas A.(2003) 'Collaboratories as a new form of scientific organization',Economics of Innovation and New Technology,12:1, 5-25.
- Fraser, M.A. (July 2005) Virtual Research Environments: Overview and Activity. Ariadne 44.
- Gavaghan, D., Whiteley, J., Pitt-Francis, J., Slaymaker, M., Lloyd, S., David Boyd, Mac Randal, D., Kleese van Dam, K., and Sastry. L. (2004) 'Gathering requirements on an Integrative Biology project'. In Proceedings of the 2004 UK e-Science All Hands Meeting, 2004.
- Gaver, W., Dunne, T., et Pacenti, E. (1999) 'Design: Cultural probes'. Interactions. Vol 6. Issue 1. Jan/Feb 1999. pp 21-29.
- Goguen, J. et Linde, C. (1993). Techniques for Requirements Elicitation. In Proceedings of IEEE International Symposium on Requirements Engineering. San Diego, CA.
- Goldberg, D.T. et Franklin, K.D. (May 2007) Socializing Cyberinfrastructure: Networking the Humanities, Arts, and Social Sciences. CTWatchQuarterly http://ctwatch.org/quarterly, vol. 3 n. 2
- Goodwin, C. et Goodwin, M. H. (1996) 'Seeing as a Situated Activity: Formulating Planes' in (eds) Engeström, Y. Middleton, D. Cognition and Communication at Work. Cambridge University Press.
- Grint, K. et Woolgar, S. (1997) The Machine at Work: Technology, Work and Organization, Cambridge, Polity Press.
- Grudin, J. (1988) 'Why CSCW applications fail: problems in the design and evaluation of organization of organizational interfaces' Proceedings of the 1988 ACM conference on Computer-supported cooperative work. Portland, Oregon, United States pp. 85 – 93.
- Guerrero, L.A., Collazos, C.A., Pino, J.A., Ochoa, S.F. et Aguilera, F. (2004) Designing Virtual Environments to support collaborative work in real spaces. Journal of Web Engineering, Vol. 2 No. 4, pp. 282-92.
- Heath, C. C., Sanchez Svensson, M., Hindmarsh, J., Luff, P et vom Lehn, D. (2002). 'Configuring Awareness' Journal of Computer-supported cooperative work. Vol 11 Issue 3-4, pp 317-347.
- Heffernan, M. et David, N. (August 2007) Legal and project agreement issues in collaboration and e-Research: Survey Results. Queensland University of Technology, Brisbane: Legal Framework for e-Research Project.

- Jackson, R.L. et Fagan, E. (2000) Collaboration and learning within Immersive Virtual Reality. Proceedings of the third international conference on Collaborative virtual environments. pp. 83-92.
- Keraminiyage, K., Amaratunga, D. et Haigh, R. (March 2009) Achieving success in collaborative research: the role of Virtual Research Environments. Journal of Information Technology in Costruction.
- Kok, J. (June 2008) Hublab. Towards online collaboratories for global data gathering in social and economic history. Final report. Internationaal Instituut voor Sociale Geschiedenis.
- Kraut, R. E., Galegher, J. et Egido, C. (Eds) (1990). Intellectual Teamwork: The Social and Technological Foundations of Cooperative Work. Lawrence Erlbaum Associates. Hillsdale, New Jersey.
- Kyng, M. et Mathiassen, L. (1997). Computers and Design in Context MIT press.
- Maron, Nancy L., Smith, K. Kirby, et Loy, Matthew (2009). Sustaining Digital Resources: An On-the-Ground View of Projects Today. Ithaka Case Studies in Sustainability. http://www.ithaka.org/ithaka-s-r/strategy/ithaka-case-studies-in-sustainability (Accessed 18/12/2009).
- Neubauer, B.J. et Brewer, G.A. (April 2004) Virtual Scholarly Collaboration: A case study. Journal of Computing Sciences in Colleges Volume 19, Issue 4, pp. 92-98.
- Okamura, K., Orlikowski, W., Fujimoto, M. et Yates, J. (1994). Helping CSCW applications succeed: the role of mediators in the context of use. In Proceedings of the 1994 ACM conference on Computer supported cooperative work table of contents Chapel Hill, North Carolina, United States, pp: 55-65.
- Olson, G.M, Zimmerman, A, and Bos, N, (2008), Scientific Collaborations on the Internet, MIT Press.
- Preston, M.N. (2000) Virtual organization as process: integrating cognitive and social structure across time and space. Michigan State University technical report.
- Randall, D., Hughes, J. et Shapiro, D. (1994). 'Steps towards a Partnership: Ethnography and System Design' In (eds) Jirotka, M. Goguen, J. (1994) Requirements Engineering: Social and Technical Issues. Academic Press.
- Redfern, S. et Naughton, N. (2002) Collaborative Virtual Environments to support communication and community in Internet-based distance education. Journal of Information Technology Education, Volume 1 No. 3.
- Sargent, M. (30 September 2005) An e-research strategic framework interim report of the e-Research coordinating Committee.
- Sillipigni Connaway, L. et Dickney, T.J. (October 2009) Common Themes identified in the analysis of JISC Virtual Research Environment and Digital Repository Projects.
- Stanley, T. (2007) Developing a VRE in a portal framework. Ariadne 51.
- Stefik, M., Foster, G., Bobrow, D.G, Kahn, K., Lanning, S., et Suchman,L. (1987) 'Beyond the chalkboard: computer support for collaboration and problem solving in meetings'. Communications of the ACM. January 1987. Volume 30, Issue 1. pp. 32 – 47.
- SURFnet (2009). Collaboration Infrastructure. Report. http://www.surfnet.nl/Documents/indi-2009-07-020%20%28Report%20Collaboration%20Infrastructure%29.pdf
- Tatar, D. G., Foster, G. et Bobrow, D. G. (1991) 'Designing for conversation: Lessons from Cognoter' International Journal of Man-Machine Studies. Vol 34. Issue 2, pp 185-209.
- Voss, A. et al. (2007) e-Research Infrastructure Development and Community Engagement.

- Weller, M. (2007) Virtual Learning Environments: Using, Choosing and Developing your VLE. Oxford: Routledge.
- Welsh, E., Jirotka, M. et Gavagan, D. (2006) 'Post-genomic science: cross-disciplinary and large-scale collaborative research and its organizational and technological challenges for the scientific research process'. Philosophical Transactions of the Royal Society. A. 364, 1533–1549.
- Working Group on Virtual Research Communities for the OST e-Infrastructure Steering Group, 31 March 2006, Final Report. http://www.nesc.ac.uk/documents/OSI/vrc.pdf (Accessed 18/12/2009)

A bibliography on collaboratories can also be consulted on the 'Science of Collaboratories' website: http://www.scienceofcollaboratories.org/Resources/biblist.php (Accessed 18/12/2009).

5.3.5 Project staff

The VRE Landscape Study project was conducted by the Centre for e-Research (CeRc) at King's College London (project lead) and the Oxford e-Research Centre (OeRC) at Oxford University. The original bid was written by Craig Bellamy and Torsten Reimer (CeRch), with significant input from Annamaria Carusi (OeRC) and contributions from Sheila Anderson (CeRch).

Centre for e-Research

- Dr Torsten Reimer (Principal Investigator)
- Dr Craig Bellamy (left the project before the interviews stage)
- Dr Valentina Asciutti (joined the project during the final phase)
- Lydia Horstman, MA (administration)

Oxford e-Research Centre

- Dr Annamaria Carusi (Co-Investigator)
- Dr Marina Jirotka
- Prof. Anne Trefethen (advisory role)