

Building a Virtual Research Environment for the Humanities

(Ruth Kirkham and John Pybus)

The Oxford eResearch Centre,
University of Oxford

Abstract

In recent years Virtual Research Environments have become prominent in the drive to present researchers with online tools and services to support their research. We describe a project to determine what a VRE would need to provide to meet the needs of a diverse range of researchers across one division of a complex university. This was achieved by the conduct of a wide-ranging user requirements survey, followed by the implementation of a number of small scale demonstrator tools designed to elicit feedback from the research community and subsequently led to the development of a project which will build and deploy a pilot Virtual Research Environment for the Study of Documents and Manuscripts.

1. Introduction

The 'Building a Virtual Research Environment for the Humanities' (BVREH)¹ project began in June 2005. The aims of the project were to build on a preliminary survey of ICT use in research projects within the Humanities Division carried out in summer 2004 and to scope the technical needs and requirements of researchers across the division in order to establish where electronic research tools would be useful both for researchers at Oxford and in a wider context. The outcomes of the survey were used to build a number of demonstrators, or early prototypes designed to show how a VRE will eventually accommodate different kinds of tools, information sources and research activities across the spectrum of research represented in the Division.

The demonstrators resulting from the BVREH project included a Research Discovery Service², which adapted a prototype already under construction in the Medical Sciences Division; a design for an Eighteenth Century Workspace which would allow cross-searching across a range of datasets related to the period; a trial of digital paper and pen technologies and the 'Virtual Workspace for the Study of Ancient Documents' (VWSAD) which is currently being developed into a fully functioning 'Virtual Research Environment for the study of Documents and Manuscripts'.

We started out with no pre-conceived notion of what a VRE for the humanities should be either in terms of functionality or technology. Instead the objective was to identify areas where VRE tools would be of maximum benefit for the arts and

humanities.³ As such the overall strategy of the project continues to focus attention on end-user support rather than implementing demonstrators within pre-existing portal frameworks. As the user requirements become more solid and the VRE tools more substantial, we will decide on the most suitable framework, keeping interoperability with other VRE projects as the most essential criteria for selection.

2. Understanding User Needs

The Humanities Division at Oxford is a diverse community encompassing all levels of technical interest and ability from those who believe themselves to be completely non-technical right through to those who have significant involvement in digital projects and who are actively seeking collaborations across subject and institutional boundaries.

The priority of the BVREH survey was to uncover the needs of the community without enforcing or imposing technology. Interviews and meetings were carried out in a fairly informal fashion, allowing interviewees to talk openly and honestly about their research and to establish ways in which a Virtual Research Environment might add benefit to their working life.

2.1 Survey Methodology

A wide variety of research projects based within the Humanities Division at Oxford were interviewed along with many individual researchers, to develop a thorough understanding of the 'research life-cycle' within the humanities. It was important to discover

exactly how, when, and where researchers carry out their work and to analyse a 'day in the life'⁴ of each individual. The basic framework was to establish the interviewee's specific research interests and to discover to what extent technology fits into their average working day.

In total we interviewed 32 researchers and research projects across 11 of the 14 faculties within the humanities division. Largely the choice of interviewee came through word of mouth and from recommendations born out of an initial survey carried out in summer 2004. At each interview the individual/s were asked to name others both within their own faculty and within the division who might provide valuable contributions, or who would not be averse to being part of the interview process.

Meetings with the research projects comprised unstructured interviews carried out in person with the lead individual/s on the project. The focus was on finding out about the specifics of each project, its content, scope, technology, funding, audience and possibilities for VRE benefit and collaboration.

The aims of the meetings were to discover how a VRE might enable access and perhaps provide greater visibility for humanities research projects but also to gain an understanding of the projects themselves and to gain insight into the support and functionality that might be provided to them through a Virtual Research Environment. It was also the intention to discover what tools and functionality the projects might already have which would be useful both to the VRE and to each other ensuring that no duplication of effort was undertaken at any stage.

Although we acknowledge that the face-to-face approach took a great deal of time to organise and arrange over such a diverse range of projects, it was clear that this approach was much appreciated. One project commented that 'it's lovely to be asked' and the intention was that through relatively informal meetings, excellent relations between the VRE project staff and the research projects could be built up and sustained through continued communication and collaboration beyond the scope of the initial survey. This goodwill extended through the development of the demonstrators and participant's subsequent willingness to take part in user testing to the current stage at which we begin to build VRE prototypes and tools.

Interviews with individual researchers were approached from a more structured starting point. The aims of these interviews were to enable the VRE team to develop a thorough understanding of the 'research life-cycle' within the humanities and to discover exactly how, when, and where researchers carry out their work and to analyse a 'day in the life' of that individual. The basic framework was to establish the interviewee's specific research interests and to discover to what extent the following elements fit into their research and their average working day:

- Current work practice; use of/access to research materials (libraries etc), collaborative aspects, dissemination
- Current IT usage
- Collaboration specifically using IT
- Local and external tools/processes used to carry out research
- Interaction with all aspects of university services and administration such as support for grant applications, funding etc

We also held a Focus Group at which attendees were introduced to a range of potential research tools and asked to comment on their usefulness by asking themselves 'what's the applicability of this for humanities research and for my own research?', 'what would the required features be if it were applicable?' and 'what is the minimum functionality before such a tool or service would be of use to me?'

2.2 Survey Outcomes

Through the survey we have highlighted the need for a simple to use, easy to access interface which underpins all stages of the research life-cycle, highlighting the need for tools and services to support:

- Research administration;
- Resource discovery;
- Data creation, use and analysis;
- Collaboration and communication;
- Publication, curation and preservation

Research administration: Interviewees wanted seamless access to information about events, including conferences, lectures and seminars; research and researcher interests of individuals within the institution and beyond; and information regarding grants and funding

opportunities, all of the above accessible in one, easy to navigate place.

Resource discovery: Interviewees wished to make the process of finding resources and research material more efficient by linking datasets and databases so that they can be cross searched with a single search term and the results returned to personal area for future reference.

Data creation, use and analysis: Interviewees wanted a secure area in which they can store material such as saved searches, images and texts; create notes and annotations and use tools to enhance, manipulate and compare items.

Collaboration/Communication: Interviewees wanted tools to enable them to work collaboratively on documents, to share material with collaborators and to view material simultaneously with colleagues wherever they might be based. At the same time interviewees wished to communicate either through video conferencing or real time chat facilities and to enable a collaborator to point/highlight and annotate items throughout the discussion.

Publication, curation and preservation: Interviewees wished to store, publish and archive their work both on personal web pages for open access or in a more secure area for academic material. Interviewees wanted to be sure that their work and the work of those around them is preserved and made available both within Oxford and externally to promote the division and its work.

As a result of the survey findings the project built a number of demonstrator tools each designed to address specific needs highlighted by the survey outcomes. Each of the demonstrators is briefly described below, but this paper will go on to focus attention on the 'Virtual Workspace for the Study of Ancient Documents' which is currently being extended to become a 'Virtual Research Environment for the Study of Documents and Manuscripts'.

2.3 Demonstrators

To create the lightweight demonstrations and mock-ups, needed to gain fast, useful feedback from the humanities community it was decided that we would not, at that stage implement our demonstrators using existing portal frameworks such as Sakai, uPortal or any of the other portals

supporting JSR168 Portlets. It should be noted that this decision was based on the need for extremely light, agile demonstrator tools with an emphasis on step-through demonstrations designed to illustrate potential and to elicit feedback from the user community. For future development, the VRE community continues to inform the project on the potential of portal frameworks for more substantial VRE deployments.

Within the Medical Sciences Division, the Research Discovery Service (RDS) was already being developed to answer a need for more information about research and researcher interests within the Medical Sciences Division. The mock-up demonstrator, for the humanities was conceived to provide further transparency in finding resources and information about those working in the humanities community and beyond as the survey found that humanities research is often multi-disciplinary and ways of finding other researchers with similar interests leading to the possibility of collaboration was most useful.

The Eighteenth Century Workspace was conceived as an environment that would integrate resources related to a project in the English faculty working with Jane Austen's manuscripts. The environment would eventually enable cross searching of a number of data sources including ECCO (Eighteenth Century Collections Online), Chadwyck-Healy Literature Collections, Samuel Johnson 'A Dictionary of the English Language and British Fiction 1800 – 1829'. The discussions and mock up versions of the system, together with talks with ECCO and other data sources, have proven that although there is a great deal of interest in providing cross-searching functionality, there still a number of obstacles to delivering such a system.

Physical tools such as communication and novel user interface devices such as digital pen and paper (Anoto)⁵ and the 'Personal Interface to the Access Grid' (PIG) were trialled to see where such tools might be beneficial on an ongoing basis.

Finally the 'Virtual Workspace for the study of Ancient Documents' on which we will concentrate more here is an interface allowing browsing and searching of multiple image collections, including tools to compare and annotate the researcher's personal collection, all linked together with communication and

collaboration tools. The Workspace has already demonstrated that the collaborative study of documents and manuscripts is an area of great mutual interest across humanities research, which would benefit from the implementation of VRE tools and services.

3. A demonstration of a ‘Virtual Workspace for the Study of Ancient Documents’

The project worked with the Centre for the Study of Ancient Documents (CSAD)⁶ to build a workspace aimed at constructing a proof-of-concept implementation of a virtual workspace for research involving decipherment and textual analysis of ancient documentary texts written on a variety of materials. The demonstrator was constructed specifically to address the needs of research on ancient documents but it deployed technologies that are likely to be applicable across other Humanities disciplines in which textual analysis is important.

3.1 Background

In 1996 an important demotic papyrus from Rifeh (now housed at University College, London), was brought to the Centre for the Study of Ancient Documents (CSAD) to be scanned and digitised. Previously the text had only been readable through the use of a magnifying glass with many days spent poring over the text. Once the papyrus had been scanned, the team of papyrologists assembled in front of a computer to see what benefits digitisation might bring. Far from the barely legible original, the images on screen made the “script come to life” and readings and suggestions from the group flowed freely leading one of the group, Dorothy Thompson, to conclude⁷:

“I have seen the future and this future works - at least so far. Work in different countries on the same text at the same time can now take place without problem and for a long and difficult text, where the writing is small and faded; the possibility of working on the image on the screen is in itself a great advance.”

This enthusiastic response details how advances in technology and digital imaging are considerably enhancing research in the field of ancient documents. Not only do the texts

become more legible, but age old problems of time and access are significantly reduced.

Whilst the above scenario describes a number of researchers travelling to the CSAD to gather around a single computer screen, the ‘Virtual Workspace’ demonstrator has been designed to take the possibilities further, enabling Professor Thompson’s vision of the future to become a reality. To “Work in different countries on the same text” simultaneously allowing researchers to collaborate and form ‘virtual gatherings’ without being unnecessarily constrained by time, money or travel.

3.2 Functionality

In order to address the needs highlighted in the scenario above, the intention of the workspace was to provide direct access to dispersed research resources such as images of original documents, lexica, corpora of texts; and to mediate collaboration between researchers in remote locations, within an integrated environment with the following capabilities:

- to enable researchers to select, store, organize and annotate image and textual data in a ‘personal workspace’
- to support collaboration by allowing multiple researchers in separate locations to share a common view of the workspace, in conjunction with real time communication via Chat, VoIP and desktop integration with the Access Grid.
- to search across multiple, distributed data sets

3.3 User Testing Methodology

The project began by working with documentary specialists to draw up a detailed list of user requirements. Initially we carried out unstructured one-to-one interviews with a small number of individual researchers aiming to get a thorough understanding of the nature of their research and the way in which they work with ancient documents and digitised images day to day. Through unstructured interviews it was possible to allow the course of conversation to develop naturally, allowing interviewees to discuss their processes of work openly, suggesting ways in which a Virtual Workspace might provide benefit to their research and to consider what elements they might like to see integrated into the demonstrator.

During these interviews there were a number of key issues that came up repeatedly. In order to decipher a text a scholar must be able to read a text accurately and to zoom in and out, adjusting the brightness and contrast, cropping and resizing images as they work. Each of these processes would currently be carried out in Adobe Photoshop, however interviewees agreed that functionality to compare texts; to view texts together on the screen, zooming in and out of each independently; to view partial transcriptions next to an image and to annotate individual sections of that image would make the workspace an extremely valuable resource over and above what is currently possible.

The Workspace (and the subsequent VRE for the Study of Documents and Manuscripts) was also informed by earlier research carried out by the CSAD in collaboration with the Engineering Sciences department at the University of Oxford. Working with texts from Vindolanda,⁸ the team used multi-spectral imaging techniques to read the damaged, often fragmented, degraded texts. A full account of the work can be found in Terras 2006⁹, describing the process by which Papyrologists read ancient texts and the development of a specific tool designed to aid historians in this process.

3.4 Testing the Interface

The initial interviews were followed by ongoing testing of implementations of the demonstrator by researchers at the Centre for the Study of Ancient Documents (CSAD). These sessions generally took the form of a gathering of up to three researchers, each of whom specialized in a different area of papyrological or epigraphical documentary scholarship. Viewing documents of particular relevance to their own work, researchers were asked to view the images as though they were working to decipher a text, imagining that they might utilise any of the electronic resources which would usually be at their disposal. In this way they tested the usefulness of the workspace, commenting on what they did and did not like and what might be useful and what could be discarded. At the end of each session the project team prioritized the outcomes of the session and endeavoured to add the most useful elements to the workspace for the next round of testing.

Many requirements were derived from this type of testing. Added to the initial thoughts and

ideas from the early interviews, the researchers were now able to see a tangible demonstration of the workspace and as such the ideas and thoughts for moving the workspace forward became richer. Through these 'testing' sessions the team got a good idea of the types of practical tools which should be integrated into the VWSAD as highly desirable requirements. Rulers, grids, brightness and contrast for highlighted sections, rulers to drag over a space in a text to work out how many characters might be missing, or to flush fit an image that may have been badly mounted prior to digitization were all suggested as elements that should be standard within the workspace.

Suggestions for using the Virtual Workspace over the Access Grid, enabling groups of researchers in separate locations to work together through the workspace was felt to be highly desirable, together with an online community allowing researchers to discuss an image or edit a document collaboratively. There was also a good deal of discussion in user testing sessions as to which external datasets should be included in the earliest version of the workspace and how they might be integrated within the viewing environment. A number of datasets were prioritized including the Lexicon of Greek Personal Names (LPGN), the Duke Databank of Documentary Papyri, Perseus, Heidelberger Gesamtverzeichnis, the Gazeteer of Papyri in British Collections, the Vindolanda Tablets and Database of Greek Inscriptions.

3.5 Expanding the scope of the Virtual Workspace

As the Virtual Workspace was built for demonstration purposes, we were unable to build in all of the functionality desired by the users and it quickly became apparent that the expansion of the Workspace could benefit a much broader variety of disciplines.

Earlier work on the Eighteenth Century demonstrator had proven a great deal of overlap between the needs of ancient documentary scholars and those of researchers working in the English Faculty on the manuscripts of Jane Austen, both requiring tools to view, manipulate and annotate texts, together with the functionality to communicate, collaborate and cross-search a variety of datasets.

Alongside this, discussions with the Silchester Roman Town VRE led us to consider how

archaeological ‘small finds’ which naturally include documents (inscriptions, tablets, papyri, graffiti etc) may be viewed in one context as a small find and in another as a documentary text. We believe that the tools used to identify and define these separate aspects can be combined to provide a fuller picture for both the documentary historian and the archaeologist. This area of overlapping perspectives allows us to explore the ways in which separate and self-standing VRE implementations might compliment and build interfaces to one another.

4. Further Work

The possibilities for the expansion of the Virtual Workspace led us to the idea of building a ‘Virtual Research Environment for the Study of Documents and Manuscripts’, which would allow us to address the user needs of documentary, textual and manuscript scholars across the humanities. Continuing to focus in the first instance on the requirements of ancient documentary specialists and extending the capability of the Virtual Workspace for the Study of Ancient Documents, the pilot is adapting Open Source tools to enable annotation and sophisticated document viewing making use of existing VRE tools to facilitate communication and collaboration between scholars. The current intention is to embed the above within the uPortal¹⁰ framework. This offers interoperability with other VRE/VLE systems and provides the ability to reuse JSR-168 portlets from other projects whilst making our components easier for others to make use of.

Whilst the project is extending the capabilities of the Virtual Workspace it is also implementing two features of much broader significance. First, that the tools and structure of the environment are suitable for the study of a wide variety of types of documents and second that the project will extend the context by treating documents not as disembodied texts but as artifacts with an original archaeological or physical context which can, in a significant number of cases (in antiquity and later periods) be recovered or reconstructed.

The significance of this approach is that the construction of a VRE appropriate for texts as artifacts opens up the possibility for the archaeologist and the textual scholar to work both separately and together within a unified environment. This aspect of the project is being

developed in close collaboration with the Silchester Roman Town VRE which has already developed a sophisticated system for registering, tracking and analyzing data recorded in the field to allow efficient recovery of information on any given artifact or in its original environment.

5. Conclusion

The BVREH project started out with no preconceived ideas as to what a ‘Virtual Research Environment’ should be. Instead, we started from the point of view of current research practice across the Humanities Division at Oxford University and attempted to determine what tools and services would be required to augment existing provision and practices.

This bottom-up approach supported by an extensive user requirements survey led us to build a number of tools which were intended to demonstrate how particular needs could be met through eResearch technologies. As these demonstrators were designed to validate the outcomes of our survey and to show researchers how their needs might be addressed, no particular attention was paid to making these prototypes fit within any of the portal frameworks used by other VRE projects. This approach was successful in engaging parts of the user community without requiring the large amounts of development effort that would have been necessary to build more advanced prototypes within a VRE framework.

Having validated some of our initial ideas we are now constructing a ‘VRE for the Study of Documents and Manuscripts’, a fully-fledged prototype which can provide production services to support an important area of research across a number of humanities disciplines. As this will be a more substantial prototype than the demonstrators we have currently developed, experience gained from the VRE community informs our decisions for integrating the VRE into a Portal framework, with interoperability and rapid, iterative development being the essential criteria for a framework decision. By implementing this model within a traditional, but ICT-literate humanities research community, it is hoped that a contribution can be made to the further development and refinement of the e-Framework and that the project will provide an exemplar for the construction of Virtual Research Environments

across the broader humanities research community.

6. References

¹ Building a Virtual Research Environment for the Humanities
<http://bvreh.humanities.ox.ac.uk/>

² A Research Discovery Service, Medical Sciences Division
<http://www.ict.ox.ac.uk/strategy/events/bowtell/index.html>

³ Anderson, S., Dunn, S. and Hughes, L. 2005: VREs in the Arts and Humanities, Proceedings of the All Hands Meeting 2005,
<http://www.allhands.org.uk/2005/proceedings/>

⁴ Fraser, M. 2005: Virtual Research Environments: Overview and Activity. Ariadne 44, July 2005
<http://www.ariadne.ac.uk/issue44/fraser/>

⁵ Anoto Digital Paper and Pen Technology
<http://www.anoto.com/>

⁶ The Centre for the Study of Ancient Documents <http://www.csad.ox.ac.uk/>

⁷ The full account of Dorothy Thompson's experiences regarding the Demotic Papyrus can be found at:
<http://www.csad.ox.ac.uk/CSAD/Newsletters/Newsletter2/Newsletter2a.html>

⁸ Vindolanda: A Roman fort on the Stanegate, near Hadrian's Wall
<http://vindolanda.csad.ox.ac.uk/exhibition/>

⁹ Terras, M., Image to Interpretation: An Intelligent System to Aid Historians in Reading the Vindolanda Texts. 2006, OUP: Oxford.

¹⁰ <http://www.uportal.org/>