



# Harold B. Lee Library

ACCESS SERVICES DEPT.

3445 HBLL, PROVO, UT, 84602

PHONE: (801) 422-8663

FAX: (801) 422-0471

EMAIL: [ill@byu.edu](mailto:ill@byu.edu)

Thank you for using the Harold B. Lee Library! The document you requested is attached. If there is a problem with the content/quality of this document, please contact us with the following info:

ILL Number	
Your OCLC Symbol	
Problem Description	<p><input type="checkbox"/> This is the wrong article/material</p> <p><input type="checkbox"/> The document is unreadable/illegible and should be resent</p> <p><input type="checkbox"/> Some pages were missing: pp. <input type="text"/> to <input type="text"/></p> <p><input type="checkbox"/> Some edges were cutoff: pp. <input type="text"/> to <input type="text"/></p> <p><input type="checkbox"/> Other (explain):</p>

**NOTICE: This material may be protected by copyright law Title 17 U.S. Code**

# Digital Humanities in Practice

Edited by

Claire Warwick, Melissa Terras  
and Julianne Nyhan



facet publishing

in association with



© This compilation: Claire Warwick, Melissa Terras  
and Julianne Nyhan 2012  
The chapters: the contributors 2012

Published by Facet Publishing,  
7 Ridgmount Street, London WC1E 7AE  
[www.facetpublishing.co.uk](http://www.facetpublishing.co.uk)

Published in association with UCL Digital Humanities Centre.

Facet Publishing is wholly owned by CILIP: the Chartered Institute  
of Library and Information Professionals.

The editor and authors of the individual chapters assert their  
moral right to be identified as such in accordance with the terms of  
the Copyright, Designs and Patents Act 1988.

Except as otherwise permitted under the Copyright, Designs and  
Patents Act 1988 this publication may only be reproduced, stored  
or transmitted in any form or by any means, with the prior  
permission of the publisher, or, in the case of reprographic  
reproduction, in accordance with the terms of a licence issued by  
The Copyright Licensing Agency. Enquiries concerning  
reproduction outside those terms should be sent to Facet  
Publishing, 7 Ridgmount Street, London WC1E 7AE.

*British Library Cataloguing in Publication Data*

A catalogue record for this book is available from the British  
Library.

ISBN 978-1-85604-766-1

First published 2012

Text printed on FSC accredited material.



Typeset from editors' files by Flagholme Publishing Services in  
10/14 pt Palatino Linotype and Frutiger  
Printed and made in Great Britain by MPG Books Group, UK.

## CHAPTER 1

# Studying users in digital humanities

Claire Warwick

### Introduction



Until relatively recently, it was unusual to study users in digital humanities. It was often assumed that the resources created in digital humanities would be used by humanities scholars, who were not technically gifted or, perhaps, even luddites. Thus, there was little point asking them what they needed, because they would not know, or their opinion about how a resource functioned, because they would not care. It was also assumed that technical experts were the people who knew what digital resources should look like, what they should do and how they should work. If developers decided that a tool or resource was a good one, then their opinion was the one that counted, since they understood the details of programming, databases, XML and website building. The plan, then, was to provide good resources for users, tell them what to do and wait for them to adopt digital humanities methods.

Frustratingly, potential users seemed stubbornly to resist such logic. The uptake of digital resources in the humanities remained somewhat slower than in the sciences. As I have argued elsewhere, the numbers of articles in journals, such as *Computers and the Humanities* (CHUM) and *Literary and Linguistic Computing* (LLC) in the 1990s and early 2000s, complaining about why traditional humanities scholars did not use digital humanities techniques or suggesting techniques they might use, grew heavily to outnumber those reporting on the actual adoption of such techniques in the mainstream (Warwick, 2004). Lack of knowledge was sometimes advanced as a possible reason for lack of engagement. During this period, very large amounts of money were spent on initiatives to publicize digital resources for humanities research and teaching. In the UK, this included the *Computers*

and Teaching Initiative (CTI) (Martin, 1996), the Teaching and Learning with Technology Programme (TLTP) (Tiley, 1996) and the Teaching and Learning Technology Support Network (TLTSN) (Doughty, 2001); the Arts and Humanities Data Service (AHDS: [www.ahds.ac.uk](http://www.ahds.ac.uk)) also had an advice and outreach role, as well as its core function of data preservation. None of these are now in existence; funders did not feel they had proved sufficiently successful to continue supporting them. Many university libraries and computing services also offered training courses in the use of digital resources for humanities scholars. Yet, the rate of change remained stubbornly slow. Funding bodies also supported digital resources for humanities scholars, with little thought to, or predictions about, levels of possible use because they did not know how such predictions might be made. Such resources often cost hundreds of thousands of pounds, so there was a risk of a severe waste of money and of academic time and energy if a funded resource was then not adopted.

In the late 1990s, a few of us began to wonder if there might be another cause for the lack of adoption of digital humanities resources. Could it be that users did not adopt resources because they were not useful or did not fit what they would like to do as scholars? Could there be other reasons to do with design, content, presentation or documentation? Initially, I suggested that digital resources available in the late 1990s did not fit the predominant research method of humanities scholars, which is complex reading (Warwick, 2004). Later, empirical studies on the way humanities scholars interact, or fail to interact, with digital resources allowed us to test this hypothesis. This chapter presents an overview of the findings of such work, arranged thematically.

## **What we know about humanities users**

Despite some erroneous perceptions in both digital humanities and the computer industry, we know a significant amount about how humanities scholars use information, whether digital or not. Since Stone's pioneering article in the early 1980s (Stone, 1982), numerous studies of information needs, and some of information behaviour, have been published, both of the humanities as a field and of individual disciplines (Warwick, Terras, et al., 2008). As we have argued in more detail elsewhere (Warwick, Terras, et al., 2008), these suggest that humanities scholars are not luddites; they simply behave differently from scientists, and many social scientists, when interacting with physical and digital information. Humanities scholars tend

to avoid performing systematic keyword searches, although most information systems and digital resources assume this. Instead, they will follow footnotes in texts they are reading (what Ellis calls chaining (1993)) or browse for information. They may even do what Bates calls 'berry picking' – in other words, select interesting pieces of information that are particularly germane to the argument they want to make, rather than citing everything written on the subject (Bates, 1989, 1). (We might speculate that this may also become more common in science in the future, when the sheer number of articles published every year exceeds the researcher's ability to read them all.) They also need a greater range of information, in terms of publication date and type: instead of reading journal articles from the last five years, they may need to consult printed books or manuscripts that are hundreds of years old, as well as images, film, music, maps, museum artefacts and various different types of historical source material (Barrett, 2005). They do not expect to solve a research question comprehensively, but to reinterpret the sources and revise the findings of others: after Crick and Watson, no one tried to redefine the structure of DNA, but articles about *Hamlet* will probably always be written. They often reread or re-examine sources in a complex, immersive way, rather than searching digital documents for factual information.

It is evident, therefore, that humanities scholars have different information needs, both on- and offline, than scientists. They are a problematic population to design for, and the field lacks the financial clout of Science, Technical Engineering and Medicine (STEM) subjects, so funding to create resources for their needs is less plentiful and may seem less profitable for commercial publishers. It is, therefore, not surprising that, until recently, most resources have been designed for the majority of users who are not from the humanities. Yet, we might argue that the way they use digital resources is, in fact, closer to the way that the average, non-academic user interacts with digital or printed information. Most of us read for pleasure, may consult a wide range of information resources and don't conduct systematic keyword searches of recently published scientific literature; thus, a study of humanities user needs may also produce important results relevant to non-professional digital resource use.

## How to study users

There are numerous methods for studying users, most of which have been developed in the fields of Human-Computer Interaction and Information Studies. There are also many excellent texts describing, in detail, how these

may be carried out, for example, Shneiderman and Plaisant (2009), Blandford and Attfield (2010) and Ruecker, Radzikowska and Sinclair (2011). Our approach at UCLDH has been to use a variety of methods, most of them designed to be as naturalistic and unintrusive as possible. Our overall approach is to study use in context; that is, to study what people do in their real life or work activities. This means that we prefer to visit someone in their office (or, in one case, an archaeological dig) and ask them to carry out a real research activity using a digital resource, rather than asking them to perform a set task in an interaction lab. We have used task-based lab testing for some research projects, but, in general, prefer to adopt as naturalistic an approach as possible to avoid the user's behaviour being prejudiced by unfamiliar conditions.

Our approach to studying users is to involve them, if possible, from the beginning of the project. Too often user testing, both in academic projects and industry, is left until late in the project; users are only asked for their opinion when the resource is built and a prototype is being tested. This may work, if the users like what has been built for them. However, if they do not, and feedback suggests radical change is necessary, there may not be sufficient funding, time or goodwill from developers to make such modifications. In such cases, the resource either remains unmodified or different researchers may be called in to conduct other tests, in the hope that they will find what the developers want them to discover, not report what users actually need. This is a very dangerous strategy, for reasons that I shall discuss below.

Thinking about use before a resource is built means studying the users, not the resource; this may be achieved using various methods. We have used interviews to determine what scholars like and dislike about digital resources and how they use information, and we have observed them using existing digital resources. We have asked them to keep diaries of their use of information and digital technologies over periods varying from between a day and a week (Warwick, Terras et al., 2009). This allows us to identify patterns of, and problems with, information usage, about which we can subsequently interview users. We have used surveys and questionnaires about the use of existing resources. We have interviewed the creators of existing, successful resources to see whether it is possible to identify any common features, in terms of design, creation or documentation (this is an unusual approach, and we believe we are the only team to have employed it in digital humanities; but it is an approach that we found very instructive during the Log Analysis of Internet Resources in the Arts and Humanities

(LAIRAH) Project) (Warwick, Galina, et al., 2008). All of these methods allow us to build up a picture of what users like and dislike, what they want to do and what they currently cannot achieve. This is then fed back to design teams to inform initial design and prototype ideas.

When initial ideas are being developed, it is also possible to use models, such as Ruecker's affordance strength model (Ruecker, Radzikowska and Sinclair (2011): Chapter 3), which allows us to test the potential functionality of a prototype design against some possible uses. At a slightly later stage in development, we can use wire frames and design sketches to run user focus groups. We have also conducted workshops, where users are asked to investigate different digital resources, record their views on paper and then take part in a subsequent focus group discussion. During the LAIRAH project, for example, we presented users with a mixed sample of resources that were either known to be used or neglected, without identifying them, asked them to speculate on which ones were used and comment on their reasons for saying so. This was then followed by a focus-group discussion. This proved a useful way to limit the bias inherent in focus groups, when one or two vocal members of the group may dominate and, thus, skew results. Subsequent examination of the written responses showed that users were willing to be more positive about some resources in writing, than they were in group discussions.

This variation between what people may say to others and what they will record in private is the reason why it is important to use a variety of different methods in user studies. It is well known that interviewees may say what they think someone may wish to know; thus, they may be more forthcoming if asked to fill in a survey or write down responses to a hands-on workshop session (Smith and Hyman, 1950). This is also why we have used quantitative data from web log analysis, since reported use may differ from what logs record, which may also be attributable to the interviewer effect. In the days before logging software, such as Google Analytics, was routinely used, very few projects or, even, institutions, such as libraries, had any reliable indication of which resources were used. Log data allowed us to determine that up to one-third of digital resources in the humanities remained unused (a very similar level to that of printed material in libraries) (Warwick, Terras et al., 2008) and to indicate the kind of material most commonly searched for. Log analysis can also indicate whether certain parts of a resource are used more often than others and whether this is related to content or design problems (the more clicks away from the index page, the less likely it is that users may find material, for example) (Huntington et al., 2002).



Of course, conducting user studies adds to the cost of developing digital resources. The time required to undertake such activities, especially if they last throughout the project, is considerable. Some projects have, instead, chosen to make use of personae or use cases. Some designers create indicative personae of typical users, giving them names, ages and occupations, and thus suggesting the uses that such a person might make of a resource (Jane is a secondary school teacher in her 30s. She wants to use a museum website to construct some new assignments about Roman food for her year 11 class on classical civilization, for example) (Grudin and Pruitt, 2002). Personae can be a useful tool, if they are constructed as a result of the kind of user studies mentioned above. However, if they are used as a substitute, there is a danger of a kind of self-fulfilling prophecy of use, where functionality is designed for the kind of users the designers want or can imagine. Yet, they cannot be sure that this is the kind of user that the resource will actually attract or that these predicted difficulties are the kind of difficulties that imagined users might face.

Use cases consist of reports of how a user, or small group of users, is using a given resource or one that is very similar. These are often used to make the case to develop something new or to argue that certain types of interface or functionality may be useful. Once again, these may be used as part of a multimethod user study, as evidence of real usage (Keating and Teehan, 2010). However, if used in isolation, the picture of use may be very partial, unless a very large number of use cases are collected. The behaviour of expert users or early adopters may also be very different from that of a majority of users, yet it is often the interested experts who furnish the use cases. As a result, the need for complex, specialist functionality, or the general enthusiasm in the user population for the resource, may be overstated. Use cases and personae should, therefore, be used with care in a multimethod user study, and should never be a substitute for other, more time-intensive methods.

## Luddites or critics?

Despite the popular image of the luddite humanities scholar who does not know what they need or how to use it, we have found that users have very complex models of their information needs and environment; they are thoughtful and critical about the affordances of physical and digital resources. This may help to explain why e-journals have been such a success, and e-monographs are still not widely used. Users are aware that a journal

article and a book are used in very different ways, even if they do not articulate this until asked. Thus, most of us still prefer to read a book in print, because it is more convenient, but are happy to read a short article on screen or print out a longer one. We also found that humanities users had complex ways of evaluating physical information resources and could tell, simply from the design, publisher or even size of a book, whether it was likely to be useful. It is still difficult for users to find digital analogies for such skills, however, and it remains an important challenge for creators of large digital resources to design tools that will allow users to orientate themselves digitally as well as they can in a physical library. This is the reason for tools such as Amazon's 'user recommendation systems' (users who bought this, also bought ...), but it is far more difficult to deploy such metaphors in an academic setting. Even the question of extent of collections is problematic; physical library users can see how big the shelf is that they are looking at and how many of them there are in a library. It is still very difficult for users to estimate how large a digital resource is and, thus, how comprehensive the results set from their search may be and how much further they need to explore. This is important for humanities users, who value recall over precision and expect to find about 90% of the results from a given search familiar. Nevertheless, we should not assume that humanities users always prefer physical to digital information resources. Users we have studied have found the convenience of digital information delivery as important as those in any other discipline and expressed considerable enthusiasm for the use of digital resources and methods. Difficulties caused by a badly designed interface to a digital collection were no more significant than a library or archive that was cold, cramped, dark or uncomfortable or an unhelpful member of staff. However, they were more likely to put up with difficult physical conditions than persist with a disappointing digital resource. It would seem ridiculous to a humanities scholar to refuse to return to a physical library if a book they hoped to find was not stocked, yet I have often heard digital resources dismissed outright if the contents were not as expected. It is difficult to tell whether this is something inherent in the nature of physical and digital information resources or whether, like the question of transferring information skills from physical to digital libraries, it is a problem of relative unfamiliarity on the part of users and signals the need for further refinement of the digital resource design. We may only find the answer to this question by repeating studies over time and trying to determine whether, and how fast, attitudes change. It is not, however, necessarily a function of being a digital native or immigrant; indeed, a recent

study suggests that there is no empirical basis for such assumptions. Rowland's research suggests that the information literacy of even, what he calls, the 'Google generation' is relatively unimpressive (Rowlands et al., 2008, 1). We found that even students who have been trained in information-seeking skills will give up as soon as they have a minimal level of information to complete a task and that they use their relative expertise to determine how little searching is necessary in a given situation, rather than conducting more complex searches to find a more complete result set. If the results of a search seem too complex to evaluate, they may even alter their query to achieve a simpler, less demanding answer. We cannot, therefore, assume that once a younger generation of scholars arrives, their ability to interact with complex digital information will necessarily improve.

Finding and using digital information seems to have something to do with how important it is to users. Students may not gain expertise gradually, however. The difference noted in the skill levels of young legal professionals, who may only be a few years older than our student sample, is probably because the information tasks they faced at work were more complex and urgent and forced them to suddenly acquire more expertise. However, it does help to explain an interesting phenomenon we found during the LAIRAH project, when we discovered that humanities scholars could be very easily deterred from using digital resources. Numerous factors caused this: confusing interfaces, problems with navigation or searching, a need to download data and use it with another application, content that was incomplete, not extensive enough, of poor quality or not as expected (for example, if a literary resource did not contain appropriate editions, it was considered unacceptable to many users). Yet, we found that if a research task is vital to the individual, and they are convinced that a resource will deliver high quality information, they will persist with a digital resource and force themselves to learn new skills or struggle with a difficult interface or functionality. Thus, we found that some linguistic resources were reported to be very useful, even if poorly designed, dated and difficult to use, because there was nothing better available for specialists in that field. The problem is that the proportion of such determined and persistent users appears to be quite small.

It has become clear to us, however, that most users will be quick to abandon resources whose quality they are concerned about. This is partly as a result of minor problems that could be relatively easily avoided. Our study of successful digital resources, during the LAIRAH project, suggested that even the name of the resource could make a difference to use. If someone is searching for census data, they may not also think to use the term

'enumerator returns', and, unless there is very complex metadata, or semantic searching is possible, a resource with a confusing or unusual title may not, therefore, be found. The possible uses of digital resources designed by technical or academic experts were often not evident to potential users. Not everyone, for example, knows what Geographical Information Systems (GIS) are or how they might be employed. This is not a problem for the dedicated expert user, but may mean that a, potentially, much larger audience fails to understand the potential use of some resources (Warwick, Terras, et al., 2008). This is part of what we have called the 'designer as user problem' (Warwick, Terras et al., 2008). If digital resources are created by academic or technical experts and user testing is not carried out, the assumption tends to be made that the users of the resource will be just like the creators. The academic creator may assume that everyone will understand what the resource is for and what it contains without much explanation, because it is obvious to them. They may also assume (possibly abetted by technical staff) that complex functionality and search capability is needed to make the resource usable and if they can learn to use such functionality, then anyone can. It may be, however, that most people do not need, or perhaps even like, the complicated functionality and, perhaps, difficult interface necessary to make this possible (Warwick, Galina, et al., 2008). The simple, Google-like search box has become a standard way that users expect to interrogate most collections of information; this is partly because it works. Most users, especially humanities academic users, do not want to have to be trained to use digital resources, regarding it as a waste of time. Some librarians have even alleged, strictly off the record, that they suspect academics do not want to admit ignorance, especially in front of their students, and that this may be a more profound reason for their antipathy to training (for obvious reasons, a source for this cannot be cited). Most technicians, librarians and commercial publishers who market resources at librarians seem to believe that it is important that all resources must have an advanced search function. In fact, numerous studies have shown that most people never use this function (Rieger, 2009). It is, therefore, clear why the model of designer-as-user is not advisable. It may lead to the creation of a resource that is needlessly complex, expensive in developer time, potentially not what users want, and, therefore, at serious risk of being under-used as a result.

.....  
**CASE STUDY** The LAIRAH Project: log analysis of internet resources in  
 the arts and humanities  
 .....

*Claire Warwick, Melissa Terras, Paul Huntington, Nicoleta Pappa, and Isabel Galina, UCL Department of Information Studies*

The aim of the LAIRAH survey is to discover what influences the long-term sustainability and use of digital resources in the humanities through the analysis and evaluation of real-time use. We utilized deep log analysis techniques to provide comprehensive, qualitative and robust indicators of digital resource effectiveness. Our aims were to increase understanding of usage patterns of digital humanities resources, aid in the selection of projects for future funding and enable us to develop evaluation measures for new projects. Until we carried out our research, evidence of actual use of projects was anecdotal; no systematic survey had been undertaken, and the characteristics of a project that might predispose it for sustained use had never been studied.

## **Methods**

### *Phase 1: log analysis*

The first phase of the project was deep log analysis: we were the first team ever to analyse web transaction logs to measure user behaviour within digital humanities resources. Transaction and search log files were provided by three online archives that were supported by the Arts and Humanities Research Board (AHRB) (now the Arts and Humanities Research Council (AHRC)): the Arts and Humanities Data Service (AHDS) Arts and Humanities Collection, Humbul Humanities Hub and the Artefact Database for the Creative and Performing Arts. These provided rich data for comparing metrics between subject and resource type. The search logs showed which resources users were interested in and which ones users subsequently visited.

We analysed at least a year's worth of transaction log data (a record of webpage use automatically collected by servers) from each resource. This data provided a relatively accurate picture of actual usage, providing: information on the words searched (search logs), the pages viewed (user logs), the website that the user has come from (referrer logs) and basic, but anonymous, user identification tags, time and date stamps.

### *Phase 2: case studies*

We selected a sample of 21 projects that the log analysis indicated to have

varying levels of use – chosen to give us coverage of different subject disciplines – to be studied in greater depth. We classified projects as ‘well used’ if the server log data from the AHDS and Humbul portals showed that they had been repeatedly and frequently accessed by a variety of users. We also mounted a questionnaire on these sites and asked which digital resources respondents found most useful. Although most users nominated information resources, such as libraries, archives and reference collections, such as the eDNB, three publicly funded UK research resources were mentioned, and, thus, we added them to the study. We also asked representatives of each AHDS centre to name which resources in their collections they believed were most used. In the case of Sheffield University, the logs showed that a large number of digital projects accessed were based at the Humanities Research Institute (HRI). We therefore conducted interviews about the HRI and its role in fostering the creation of digital humanities resources.

The projects were studied in detail, including any documentation and reports that could be found on the project’s website, and a representative of each project was interviewed about project development, aims, objectives and their knowledge of subsequent usage. We analysed each project’s content, structure and design. We asked whether it undertook any outreach or user surveys and how the results of surveys were integrated into project design. We also asked what kind of technical advice the project received, whether from institutional support people, from humanities computing centres or from central bodies, like the AHDS. All these measures are intended to determine whether there are any characteristics shared between ‘well used’ projects.

We also studied projects that appeared to be neglected or underused. A small group of humanities users were asked to investigate a sample of digital resources: half were well used and the others neglected, but their status was not initially revealed. A hands-on investigation was followed by a discussion of factors that might encourage or deter future use of such resources. We aimed to find out whether their lack of use was because users had not heard of a given resource or whether there were more fundamental problems of design or content that would make the resource unsuitable for academic work.

## Findings

We found that roughly one-third of all projects appeared to be unused. When asked to evaluate unused resources, users were able to identify several problems with design and content. They were deterred from use because of unintuitive interfaces, the need to download data for use in another application, confusion

as to what the content might be used for and even a confusing name. They also needed more information about the content of resources, how and why it had been selected and the expertise of the project team.

Well used projects did share common features that predisposed them to success. The effect of institutional and disciplinary culture in the construction of digital humanities projects was significant. We found that critical mass was vital, as was prestige within a university or the acceptance of digital methods in a subject field. The importance of good project staff and the availability of technical support also proved vital. If a project is to be well used, it was also essential that information about it should be disseminated as widely as possible. The single most common factor in use of a project was a good dissemination strategy. Even amongst well used projects, however, we found areas that might be improved: these included organized user testing, the provision of, and easy access to, documentation, and the lack of updating and maintenance of many resources.

## Recommendations

Digital humanities projects should undertake the following actions:

1. Keep documentation and make it available from the project website, making clear the extent, provenance and selection methods of materials for the resource.
2. Have a clear idea of whom the expected users might be; consult them as soon as possible and maintain contact through the project, via a dedicated e-mail list or website feedback.
3. Carry out formal user surveys and software and interface tests and integrate the results into project design.
4. Have access to good technical support, ideally from a centre of excellence in digital humanities.
5. Recruit staff who have both subject expertise and knowledge of digital humanities techniques, then train them in other specialist techniques as necessary.
6. Maintain and actively update the interface, content and functionality of the resource, and do not simply archive it.
7. Disseminate information widely, both within specialist subject domains and in digital humanities. ■

## Trust

As we have seen, users need as much information about a resource as possible to understand what it might be useful for. However, underlying much of our research on users is the issue of trust in digital resources and technologies. The more information users can find about a resource, the more they are likely to trust it. As discussed above, humanities scholars have a complex repertoire of information skills that allow them to evaluate traditional information resources. These have grown up over several hundred years of the development of printed academic resources (Vandendorpe, 2009). A prestigious journal name or book publisher tells us that the content has been peer reviewed by other academic experts. Footnotes or references in the text reassure us that the writer has compared their findings with other work in the field and researched other sources. The academic affiliation of the author tells us about their expertise and standing in the field. The methodology of an article tells us how the work has been conducted, for example, how data was selected, sampled and analysed. Digital resources are only beginning to find ways to provide such information. In the LAIRAH report recommendations, we suggested that all digital resources should have a top-level link called 'About this Project', or something similar, under which creators should provide as much information as possible about its purpose and how it might be used; what its contents are and how comprehensive they are; if selections have been made from a larger corpus, how this has been done, why, and who has done so; who created the resource and where they are based; how technical decisions were made, for example, about the markup or metadata schema. The more effectively this is done, and the more easily it can be accessed, the more users are likely to trust digital resources. This is likely to become even more important in the near future. The UK's Research Excellence Framework will now allow digital resources to be submitted in all subject areas and not simply the publications written about them (Higher Education Funding Council for England, 2011). As a result, it will become even more vital that we gain a sense of the rationale for the choices made in the course of digital resource construction, so that assessors can make informed decisions about resource quality and impact in the wider world.

At present, however, trusted brands are very important. Many digital resources that are most familiar to users, such as e-journals or large digital reference collections, are produced by commercial publishers, who make significant investments in testing the appearance and functionality of their resources. This is also usually the case with digital resources in major



cultural heritage organizations, such as museums and galleries. This means that the standard of resource that academic users expect is often higher than most academic projects can manage, especially for interface design. This, coupled with the brand identity of museums and major publishers, reassures users about the quality of the content.

A pioneering study showed that visitors to websites make judgements about them in fractions of a second. We appear to make up our minds about digital resources too quickly to perform a conscious critical evaluation of it: our gut instinct tells us whether it looks 'right'. If users sense that something looks 'wrong' – which may simply mean that the interface looks unfamiliar, is difficult to use or lacks information about its creation and provenance – users may regard it as untrustworthy, neglect it and revert to more familiar resources, whether printed or digital (Warwick, Terras et al., 2008). This demonstrates why those creating digital projects must design a resource that works easily and looks as impressive as possible. The only way to do this, other than being lucky, is to carry out proper user testing.

One of the reasons users think that resources look 'wrong' is if they seem dated. If they try to use a resource and parts of it no longer work – links are broken, for example – they will lose yet more trust. Commercial resources are updated constantly, to make sure that information is current and the interface functional and consistent with current design trends. The problem for many digital resources based in the academic and cultural heritage sectors is that there may no longer be any funding to perform such updating if the content is freely available and was funded by a fixed-term grant. As we have seen, if users do not feel that a resource is to be trusted, because it appears to be dated, they are reluctant to use it. This is a waste of the (probably) very large amount of money that was spent on its creation. Institutions have only recently begun to develop strategies to deal with this problem.

This is especially serious for resources that involve crowdsourcing or web 2.0 technologies, where users become an integral part of the research process. For example, the award-winning Transcribe Bentham project, discussed in Chapter 2 of this volume, was funded by a short research grant. However, at the end of its funding period, over 1000 people had already taken part in transcribing manuscripts and become part of a thriving user community. Since this project is an important vector for engagement between UCL researchers and the public, to have closed it and locked out all our volunteers would have been disastrous and contrary to everything that UCL believes in, in terms of outreach and openness. As a result, short-term

internal funding had to be provided, while further external funds were sought. UCL recognizes the need both to maintain the infrastructure and to continue the activity with which it is engaging volunteers.

## Longitudinal studies

Change over time is not something that is very often considered in terms of user studies. They are often carried out at particular points in time, and significant longitudinal studies are relatively rare. There tends to be an assumption, therefore, that user views of digital resources are somewhat fixed. In digital resource creation, one important principle that should be followed, if at all possible, is a cycle of user testing and feedback. Once tests have been carried out and modifications made, it is important to feed back to users what has been done in response to their views. This can either be done by direct communication, in the form of a change log or development blog on the website; an end of project workshop; or another written form of communication with the user community, such as an online newsletter or progress report. An iterative development cycle is, in itself, a useful way to communicate with users. If, for example, a focus group has been carried out to ascertain users' views of wireframes or design sketches, then a hands-on session with a prototypical system not only helps to indicate whether views initially expressed are true in a working version, but shows that the development decisions taken reflect initial users' views. Users like to be able to see that changes have been made as a result of their input and will often be very supportive of something that they helped to create. Our work on the VERA Project was an excellent example of this. The following case study gives the full details of the project, our part in which was to study the way that archaeologists use digital technologies in the field, especially to record what they have found.

---

### **CASE STUDY** The VERA Project

---

*Claire Fisher, British Museum, Melissa Terras, UCLDH, and Claire Warwick, UCLDH*  
 The Virtual Environments for Research in Archaeology (VERA) Project was funded as part of the Joint Information Systems Committee (JISC) and involved the Department of Archaeology and the School of Systems Engineering at the University of Reading, the York Archaeological Trust and the School of Library, Archive and Information Studies at UCL. The project was based around the

University of Reading's well established excavation at the Roman site of Silchester. The Silchester *Insula IX* project ([www.silchester.rdg.ac.uk/index.html](http://www.silchester.rdg.ac.uk/index.html)) provided the ideal test-bed for a Virtual Research Environment (VRE) project, because key to the smooth running of this large, urban excavation was the Integrated Archaeological Database (IADB, [www.iadb.org.uk/index.htm](http://www.iadb.org.uk/index.htm)). Used for recording, analysis, archiving and online publication of archaeological finds, contexts and plans, the IADB allows integrated access to all aspects of the excavation record. It was used to populate the VRE with data generated by a complex urban excavation.

The VERA Project set out to:

1. Investigate how digital input could be used to enhance the process of documenting, utilizing and archiving excavation data.
2. Create a suitable Web portal to provide an enhanced user experience.
3. Develop tools that could be integrated with existing practices of research archaeologists unfamiliar with VREs.
4. Test the tools in a real world setting.

UCL's role was to ensure that the needs of the archaeologists and associated specialists remained at the heart of developments.

The VERA IADB usability study was carried out at the 2007 VERA winter workshop at Reading. The development of the IADB has always been driven by its users and has developed alongside their working practices. However, this was the first time that user reactions to the IADB had been formally documented. Participants at the workshop were divided into two groups:

- those with no (or little) experience of using the IADB, designated 'novice users'
- those who have experience of using the IADB in their work, designated 'experienced users'.

The usability study provided the team with useful information about user perceptions, plus details of the typical tasks carried out by archaeologists and associated specialists. The novice users felt that they could quite quickly get to grips with the system; the experienced users carried out a wide range of tasks using the IADB and used it at (almost) all stages of various projects.

The Silchester project utilizes the skills of a large and geographically dispersed group of specialists. Each specialist uses the IADB for varying purposes, and one of the aims of the VERA Project was to enhance the ways in which each

researcher uses it. Interviews were carried out to explore how the existing users organize their work; to discuss their experiences of working with the IADB; to find out to what extent the IADB met their needs; and if any changes might make their work easier. The results from these interviews were used and fed into IADB development.

Excavation data has traditionally been entered into the IADB through manual digitization, usually once the excavation season is over. A key aim of the VERA Project was to investigate the use of information technology (IT) within the context of a field excavation and to ascertain whether it may be appropriated to speed up the process of data recording, entry and access. From 2005 onwards, a number of field trials had been carried out at the Silchester excavation, using a variety of digital recording methods, including digital pens and notebooks and hand-held internet browsers. The 2008 field season focused on the use of digital pens for direct digital data gathering. We used fieldwork observations, user needs discussions, a diary study and an end-of-season questionnaire to analyse user reactions to the digital pens.

We aimed to observe how well the digital pens fitted in to the workflow of the site and to record user feedback. The discussions provided the framework for creating the end-of-season review for the digital pens. A diary study was used to gather information about different archaeological roles and the way that they are supported by both digital and analogue technologies. These studies allowed the VERA team to explore how the implementation of new technology affected the workflow of the site. Lessons learnt included the central role of the traditional context-recording sheet and the need for any new technology to integrate with existing workflows. ■

## Responses to the new technology

Introducing new ways of working into well established systems can be problematic, especially if the changes include the introduction of unfamiliar technology. UCL's involvement in the VERA Project illustrates how user case studies, analysis and feedback were used to develop recording systems and virtual research environments that fit into the current workflow of archaeologists and associated specialists.

The digging season was short – six weeks in the summer of each year – and we studied the dig for three years. This gave us an unusual opportunity to study change over time. Initially, digital methods were only trialled in a small part of the site. We found that they therefore seemed risky and abnormal to most people, and, thus, the methods, and we, were treated with

suspicion. Most people were relatively negative about the use of digital technologies – such as digital pens and paper – for finds recording, as opposed to traditional printed context cards. They also felt they had suffered from a lack of training. The following year, digital technologies, predominantly digital pens, were used throughout the site, and we provided training in their use, as well as feedback on what we had learnt the previous year. Diggers became more positive and began to understand the aims of the study, becoming more open to possible changes. In the final year, further improvements were made to the way digital data was entered and maintained as a result of user feedback, and, when other technologies, such as GPS, were introduced, they were adopted much more readily than we might have expected. Users could understand how their feedback had been integrated into the use of technology and that while systems were not perfect, they had improved, and we had made every effort to act upon user comments as far as possible. As a result, they became noticeably more positive about the use of digital technologies in each year of the study. This shows how important it is that users can see how their feedback has been used to improve a system: if they can see progress, it appears that they will make an effort to support the system they created. If it is not exactly what they would have wished, they will make an effort to deal with pragmatic decisions, if they can understand the reasons for them. In the case of Silchester, they understood that the cost of producing a fully digital recording system would have been prohibitive and were, thus, willing to work with a compromise – a semi-digital solution, which, nevertheless, resulted in faster, more accurate data entry than had been possible using manual recording.

## Conclusion

It is clear, therefore, that we cannot, and must not, try to tell users what they ought to like, need or use. We also cannot expect people to abandon working practices instantly when they have suited them well over many years and, in some humanities fields, generations. As we saw at Silchester, if users are consulted, and researchers take the time to understand their working culture and how digital resources fit into it, there is the possibility that attitudes to, and levels of, digital resource use may change. However, we must ensure that users know what they need, to complete their work successfully. If digital resources fit well with what they want to do with them, users will adopt them. Attitudes to digital resources have changed massively in the last

decade, with far greater use of the internet for information seeking and the widespread uptake of resources, such as digital reference resources and e-journals. This is surely because they fit well with what humanities academics would like to do. For example, e-books have recently become more popular, because a new generation of digital reading devices are as light as a paperback, with screens that are more comfortable to read from than earlier e-readers. Thus, users are far more likely to adopt them, because they fit well with their usual reading behaviour and have notable advantages, such as the ability to carry several hundred 'books' in a small, light device.

The aim of those of us designing resources in digital humanities, therefore, remains analogous to this. We must understand the needs and behaviours of users. As a result of this understanding, we must design resources that fit well with what our users already do, while providing advantages in terms of convenience, speed of access, storage capacity and innovative information tools that digital publication affords. If we do so, there is every chance that such resources will be used and will help to make possible new kinds of scholarship that would be inconceivable without digital content, tools and delivery mechanisms.

## Bibliography

- Attfield, S., Blandford, A. and Dowell, J. (2003) Information Seeking in the Context of Writing: a design psychology interpretation of the 'problematic situation', *Journal of Documentation*, **59** (4), 430–53.
- Barrett, A. (2005) The Information Seeking Habits of Graduate Student Researchers in the Humanities, *The Journal of Academic Librarianship*, **31** (4), 324–31.
- Bates, M. (1989) The Design of Browsing and Berrypicking Techniques for the Online Search Interface, *Online Information Review*, **13** (5), 407–24.
- Bates, M. (1996) Document Familiarity in Relation to Relevance, Information Retrieval Theory, and Bradford's Law: the Getty Online Searching Project Report No. 5, *Information Processing & Management*, **32** (6), 697–707.
- Bates, M. (2002) The Cascade of Interactions in the Digital Library Interface, *Information Processing & Management*, **38** (3), 381–400.
- Blandford, A. and Attfield, S. (2010) *Interacting with Information*, Morgan & Claypool.
- Blandford, A., Rimmer, J. and Warwick, C. (2006) *Experiences of the Library in the Digital Age*, <http://www.ucl.ac.uk/annb/docs/abjrcwCCDTpreprint.pdf>.
- Brown, S., Nonneke, B., Ruecker, S. and Warwick, C. (2009) Studying Orlando's

- Interfaces, paper presented at *SDH/SEMI 2009*, Carleton University, Ottawa, Canada.
- Brown, S., Ross, R., Gerrard, D. and Greengrass, M. (2007) *RePAH A User Requirements Analysis for Portals in the Arts and Humanities*, Arts and Humanities Research Council.
- Doughty, G. (2001) *Teaching and Learning Technology Support Network*, [www.elec.gla.ac.uk/TLTSN/](http://www.elec.gla.ac.uk/TLTSN/).
- Ellis, D. (1993) A Comparison of the Information Seeking Patterns of Researchers in the Physical and Social Sciences, *Journal of Documentation*, **49** (4), 356–69.
- Grudin, J. and Pruitt, J. (2002) Personas, Participatory Design and Product Development: an infrastructure for engagement, *Proceedings of Participation and Design Conference (PDC2002)*, pp. 144–61, <http://research.microsoft.com/en-us/um/people/jgrudin/>.
- Higher Education Funding Council for England (2011) *Assessment Framework and Guidance on Submissions*, HEFCE.
- Huntington, P., Nicholas, D., Williams, P. and Gunter, B. (2002) Characterising the Health Information Consumer: an examination of the health information sources used by digital television users, *Libri*, **52** (1), 16–27.
- Jones, C., Ramanau, R., Cross, S. and Healing, G. (2010) Net Generation or Digital Natives: is there a distinct new generation entering university?, *Computers and Education*, **54** (3), 722–32.
- Keating, J. and Teehan, A. (2010) Appropriate Use Case Modeling for Humanities Documents, *Literary and Linguistic Computing*, **25** (4), 381–91.
- Lindgaard, G., Dudek, C., Fernandes, G. and Brown, J. (2005) Attention Web Designers: you have 50 milliseconds to make a good first impression, *Behaviour & Information Technology*, **25** (2), 115–26.
- Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C. and Buchanan, G. (2007) A Library or Just Another Information Resource? A case study of users' mental models of traditional and digital libraries, *JASIST*, **58** (3), 433–45.
- Martin, J. (1996) The Computers in Teaching Initiative, *Ariadne*, Issue 5, [www.ariadne.ac.uk/issue5/cti/](http://www.ariadne.ac.uk/issue5/cti/).
- O'Hara, K., Smith, F., Newman, W. and Sellen, A. (1998) *Student Readers' Use of Library Documents: implications for library technologies*, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Los Angeles, CA*, pp. 233–40, New York, ACM Press.
- Rieger, O. (2009) Search Engine Use Behavior of Students and Faculty: user perceptions and implications for future research, *First Monday*, **14** (12).
- Rimmer, J., Warwick, C., Blandford, A., Gow, J. and Buchanan, G. (2008) An Examination of the Physical and Digital Qualities of Humanities Research,

- Information Processing and Management*, **44** (3), 1374–92.
- Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., Withey, R., Jamali, H., Dobrowolski, T. and Tenopir, C. (2008) The Google Generation: the information behaviour of the researcher of the future, *Aslib Proceedings*, **60** (4), 290–310.
- Ruecker, S., Radzikowska, M. and Sinclair, S. (2011) *Visual Interface Design for Digital Cultural Heritage*, Ashgate.
- Shneiderman, B. and Plaisant, C. (2009) *Designing the User Interface*, Pearson.
- Smith, H. and Hyman, R. (1950) The Biasing Effect of Interviewer Expectations on Survey Results, *Public Opinion Quarterly*, **14** (3), 491–506.
- Stone, S. (1982) Humanities Scholars: information needs and uses, *Journal of Documentation*, **38** (4), 292–313.
- Tiley, J. (1996) *TLTP: The Teaching and Learning with Technology Programme*, <http://www.ariadne.ac.uk/issue4/tltp>.
- Vandendorpe, C. (2009) *From Papyrus to Hypertext: toward a universal digital library*, University of Illinois Press.
- Warwick, C. (2004) Print Scholarship and Digital Resources. In Schreibman, S., Siemens, R. and Unsworth, J. (eds), *A Companion to Digital Humanities*, Blackwell.
- Warwick, C., Terras, M., Huntington, P., Pappa, N. and Galina, I. (2007) *The LAIRAH Project: log analysis of digital resources in the arts and humanities. Final report to the Arts and Humanities Research Council*, [www.ucl.ac.uk/infostudies/claire-warwick/publications/LAIRAHreport.pdf](http://www.ucl.ac.uk/infostudies/claire-warwick/publications/LAIRAHreport.pdf).
- Warwick, C., Galina, I., Terras, M., Huntington, P. and Pappa, N. (2008) The Master Builders: LAIRAH research on good practice in the construction of digital humanities projects, *Literary and Linguistic Computing*, **23** (3), 383–96.
- Warwick, C., Terras, M., Huntington, P. and Pappa, N. (2008) If You Build It Will They Come? The LAIRAH Study: quantifying the use of online resources in the arts and humanities through statistical analysis of user log data, *Literary and Linguistic Computing*, **23** (1), 85–102.
- Warwick, C., Rimmer, J., Blandford, A., Gow, J. and Buchanan, G. (2009) Cognitive Economy and Satisficing in Information Seeking: a longitudinal study of undergraduate information behaviour, *Journal of the American Society for Information Science and Technology*, **60** (12), 2402–15.
- Warwick, C., Terras, M., Fisher, C., Baker, M., O’Riordan, E., Grove, M. et al. (2009) iTrench: a study of user reactions to the use of information technology in field archaeology, *Literary and Linguistic Computing*, **24** (2), 211–24.



Copyright © 2012. Facet Publishing. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.