The master builders: LAIRAH research on good practice in the construction of digital humanities projects

Claire Warwick, Isabel Galina, Melissa Terras, Paul Huntington, and Nikoleta Pappa

School of Library, Archive and Information Studies, University College London, UK

Abstract

Although many digital humanities resources are being developed for online use, there is little understanding of why some become popular, whilst others are neglected. Through log analysis techniques, the LAIRAH project identified twenty-one popular and well-used digital humanities projects, and in order to ascertain the factors they had in common, which predisposed them to be well used, conducted in-depth interviews with the creators of these resources. This article presents the findings of the study, highlighting areas that developers should be aware of, and providing a set of recommendations for both funders and creators, which should ensure that a digital humanities resource will have the best possible chance of being used in the long term.

Correspondence:
Claire Warwick, School of Library, Archive and Information Studies, University College London.
E-mail:
c.warwick@ucl.ac.uk

1 Introduction

The creation of digital resources is an important activity in the digital humanities and has produced a large number of materials currently available online. These are funded by universities, governments and philanthropic bodies: the Arts and Humanities Research Council (AHRC) in the United Kingdom has funded over 250 digital humanities projects since 1998. What happens to these resources once available is poorly understood. Anecdotal evidence suggests that some resources are ignored whilst others are frequently used but the reasons for this are not clear. There have also been no systematic, evidence-based studies of the use and non-use of digital humanities resources.

The LAIRAH (Log analysis of Internet Resources in the Arts and the Humanities) project (http://www.ucl.ac.uk/slais/LAIRAH) was based at University College London's (UCL's) School of Library, Archive and Information Studies. The fifteen-month study, which ran from July 2005 to September 2006, was funded by the AHRC's ICT strategy project scheme to discover what influences the long-term sustainability and use of digital resources in the humanities. This article presents the results of a study of the producers of digital humanities resources, in particular those that are particularly well used. Our aim was to discover whether certain practices in the construction of digital humanities projects had an effect on its subsequent use.

1.1 Previous work in the area

There are numerous studies about the information needs and information-seeking practices of

humanities scholars (Herman, 2001; Talja and Maula, 2003; Barrett, 2005; British Academy, 2005). These have added to our understanding of what kind of resources humanities scholars use, and those they may like to see in future (Bates, 2002; Palmer and Neumann, 2002; Dalton and Charnigo, 2004).

Recent studies, including those produced by other projects that were part of the ICT Strategy scheme, show that many humanities scholars are enthusiastic users of digital humanities resources, although they prefer generic informational resources (web pages of libraries and archives, or large online reference collections), to the kind of digital object which is comparable to a scholarly book (British Academy, 2005; Brown *et al.*, 2007; Huxley *et al.*, 2007). They need a wide range of resources, of both age and type, including printed materials, manuscripts, personal knowledge collections and face-to-face information gathering. They usually reinterpret ideas rather than creating or discovering new data or facts (Stone, 1982; Barrett, 2005). Humanities scholars will use technologies if they fit well with what they do (Bates, 2002), especially if they save time or effort (Wiberley, 2001). However, they prefer not to have to undertake any training to enable them to use digital resources or applications (Brown *et al.*, 2006).

Research has only recently been conducted into the actual online behaviour of humanities scholars: by our CIRCAh research group at UCL, of which LAIRAH formed a part, and by Duff and Cherry at the University of Toronto (Duff and Cherry, 2004). (Although Bates' work on the Getty project predates web publishing; Bates, 1996.) This research shows that humanities researchers have sophisticated information skills and mental models of their physical information environment, although they find these difficult to apply to the digital domain (Makri *et al.*, 2007). They are aware of the affordances and problems of digital resources, being concerned with accuracy, selection methods, and ease of use (Warwick *et al.*, 2008a). They require information about the original item when materials are digitised (Duff and Cherry, 2000). They expect high quality content: anything that makes a resource difficult to understand—a confusing name, a challenging interface, or data that must be downloaded—will deter them from using it (Warwick *et al.*, 2008a). They may even be affected by emotional responses to the practical environments in which they use digital or analogue resources (Rimmer *et al.*, 2008).

Thus it is incumbent on producers of digital resources not only to understand the working practices of the scholars for whom they design, but to produce a resource that is attractive, usable and easy to understand. However, perhaps surprisingly, there appears to be no research that assesses how well digital humanities resources are performing in these respects. This benchmarking approach, previously applied to web design (Nielson and Tahir, 2002), new to the study of digital humanities, allows us to observe the features of successful digital resources and their effect on usage (e.g. is a link to documentation available? Do logs indicate that it is used?). In a previous article, we discussed the reactions of users when introduced to a sample of digital humanities resources (Warwick *et al.*, 2008a), and while we refer briefly to certain findings on the views of users, this article studies digital resources from the point of view of their producers. This is designed to complement our study of users and enables us to understand the reasons for certain practices and strategies in resource creation, and whether they have any impact on use.

2 Methods

A group of twenty-one digital humanities projects from different subject disciplines were selected for the study. Although these resources had varying levels of use, our aim was to select those projects that were perceived as particularly useful or that we knew were heavily used. We selected projects to represent different subject disciplines using a combination of criteria including log analysis, online questionnaires and recommendations from Arts and Humanities Data Service (AHDS) subject centres.

We analysed the server logs of the AHDS and Humbul portals. Server log analysis can serve as an important first step towards measuring the use of electronic resources. This unfiltered data can be

used in conjunction with other more qualitative data, such as interviews with users, in order to build up a picture of electronic resource usage. Although deep log analysis has been used extensively by the CIBER research group at UCL SLAIS, in areas such as health information and commercial publishing (Huntington et al., 2002; Nicholas et al., 2007), this technique has not been previously used to assess use levels of digital resources in the arts and humanities. Our use of deep log analysis is discussed in more detail elsewhere (Warwick et al., 2008b). However, in essence it involved the analysis of the search and transaction logs of the two portals web servers. The records of the IP numbers of machines accessing these portals enabled us to determine which resources were being accessed, and patterns of data concerning frequency of access, nationality of users, and which parts of the site were accessed. Questionnaires were also mounted on both the AHDS and Humbul portals asking respondents as to which digital resources they found most useful. The questionnaires were available for four months and we collected 149 complete responses.

Finally, we selected projects based on expert recommendations. Although funding for the AHDS has subsequently been discontinued, at the time of the study, the AHDS consisted of subject centres in Archaeology, Visual Arts, Literature, Language and Linguistics, and Performing Arts. We asked representatives of each centre as to which resources in their collections they believed to be most frequently used. The logs also indicated that a number of digital projects based at the Humanities Research Institute (HRI) at Sheffield University were frequently accessed. We therefore conducted interviews about the role of the HRI in fostering the creation of digital humanities resources.

The projects were as follows:

- Old Bailey online[1]
- Andre Gide Editions project[2]
- French Stars Project[3]
- The English Monastic Archives Project[4]
- The Survey of English Usage[5]
- The London College of Fashion Archives[6]
- Excavations at Eynsham Abbey[7]
- Toronto Dictionary of Old English Corpus
- The Ave Valley Project[8]
- The Avant Garde Project[9]
- The DIAMM Project[10]
- The Channel Tunnel Rail Link Archives
- Designing Shakespeare
- Exeter Cathedral Keystones and Carvings
- The Suffrage Banners Project[11]
- The Jeremy Bentham Project[12]
- PARIP[13]
- The Powys Digital History Project[14]
- The Celtic Inscribed Stones Project[15]
- The Imperial War Museum Concise Art Collection
- GIS of the ancient Parishes of England and Wales, 1500–1850

We studied any documentation and reports that could be found on the project's website, and conducted a series of semi-structured interviews with a representative of each project, either the principal investigator (PI) or a research assistant (RA). The interview guide is available as Appendix 1. However, we asked about the history of the creation of the resource, and how it was funded, about the use of technical standards and any technical support available. We asked about whether any documentation had been kept or any user testing carried out. We asked whether the resource had been as envisaged, whether any unexpected problems had arisen and what solutions have been reached. Interviews were recorded and later transcribed, the text was then analysed using a grounded theory style approach. In addition to the themes we had identified in the interview guide, this approach revealed new areas of importance, such as researcher training and PI career progression.

We also carried out workshops with academics and MA students to determine the reactions of users to the sample of projects, both used and neglected. These findings are discussed in detail elsewhere (Warwick, 2008), but where relevant we refer to and make comparison with the findings of the user research in the following discussion.

3 Findings

The interviews suggested a number of factors that may affect the creation, use and long-term

sustainability of a digital resource. These include institutional context, staffing, dissemination, age of resources, user contact, documentation, access to documentation, and sustainability issues, which we discuss in detail subsequently.

3.1 Institutional context

Institutional context had an important impact on whether or not digital humanities research was to thrive. Sheffield University in the United Kingdom is a good example of this. Their HRI is highly valued by the University and the success of early research leaders has underlined the prestige of work in digital humanities, not least because of the availability of grant funding for humanities research. Where academics have gained prestige and promotion, other junior colleagues have been inspired to follow their example.

However, in another university where a successful digital project was seen as outside the core research of a rather traditional department, and the acquisition of grant funding had not been perceived as especially prestigious, no other digital projects had begun and few colleagues or graduate students showed interest in the area. We therefore found a clear correlation between institutional encouragement of digital humanities research and the creation and use of digital resources.

Related to these issues of prestige is that of recognition for scholars who engage in digital humanities research. One PI, a relatively junior scholar, who has become well known due to her digital humanities research, also raised these concerns:

> This is one of my difficulties because my academic profile is quite different from my colleagues and in fact it actually does map onto a science model much more. I publish in journals, I haven't got a monograph because [...] the minute I start typing it's out of date and I go to a lot of conferences
> [several intervening paragraphs]
> I think I am doing an interesting job and productive work but I think we have to rethink some of those expectations if we want other people to join us [...] Invariably I work with people who are full professors and men, I am usually the only woman and the only person

under 50 in the room and that's a weird thing for me and I do think that its one of the reasons I persist but it does put me in a very difficult position in terms of the responsibility I am given versus the experience I have and the, you know academic cloak that I have because I don't have the years and years of scholarship behind me to make the kind of claims that I am making in some policy level things and yet on one else is in a position to make any decisions at all. So I find I alternate between thinking it's terribly exciting being at forefront and being pushed out in front and thinking wait a minute, can't I just do a normal job you know teach and write books like everybody else? I am part of the new group of people but I think it's quite difficult to advise even PhD students who want to work in this area that they may not get promoted, they may not get jobs so I think it's something that needs to be taken on board. I don't know by whom or under what circumstances but crediting this work properly is really important. (P17)

The research culture of particular disciplines affects the production and use of digital resources. As the PI makes clear, digital scholarship tends to assume a more scientific model of scholarly production. In Archaeology, this is recognized and rewarded, and the production of a good digital resource is regarded as similarly prestigious as important print publications. Her comments about advice to PhD students are unfortunately not unusual. A recent MLA report (MLA, 2007) has shown a disquieting lack of knowledge amongst traditional humanities scholars about how digital scholarship should be valued. However, it is vital that such issues are addressed, given the importance of prestige and recognition, within both universities and subject disciplines, for the acceptance of digital humanities research.

The advocacy of key individuals who are respected equally for their scholarship and digital knowledge may have a galvanizing effect on the production of digital resources. If such scholars are seen to have attained recognition and promotion as a result of their digital research activities, the products of such

digital research ought themselves to acquire prestige, in an analogous fashion to the respect accorded to books written by distinguished scholars.

3.1.1 Technical support

Most PIs of successful projects were very positive about the level of support and information provision received. Most projects were supported by local IT services or more expert colleagues. But those who had contact with a centre of digital humanities excellence, such as the Centre for Computing in the Humanities, Kings College London or the HRI at Sheffield were especially well advised. All projects interviewed had received advice from the AHDS and some had also been advised by national bodies (such as HEDS and JISC). However, the lack of the AHDS suggests that in future there will be an even greater reliance on local IT services and potentially greater advantages for those projects fortunate enough to have access to digital humanities expertise, which remains relatively rare at least in the United Kingdom.

Most projects had been carefully planned, and PIs reported that the finished product turned out much as expected or better. Nevertheless, most had encountered unexpected technical difficulties, which had taken longer than expected to solve. Problems occasionally occurred during the planning stage if the PI had insufficient technical knowledge, or insufficient IT advice had been available, and in a few cases this meant that the resource could not be implemented in the form that had originally been anticipated.

> Well originally we did want [the database to be delivered via the web] but it turned out to be excessively expensive and actually the programming that would have been involved would have been far too complex for [the RA] and the data team at the University wanted to charge about £20,000 or something so, and there was no money left in the budget for that and [. . .] we weren't aware of how big a job that would be to make it searchable on the web. (P21)

The more detailed planning that was undertaken, and the more technically informed the planners,

(whether PI or IT support staff) the less likely it was that the projects would encounter such problems. This underlines once again the importance of sufficient technical support, both at local and national level.

3.2 Staffing

Recruiting a competent RA for a project is vital and many PIs interviewed viewed this as key to the success of the project. However, finding an RA with both subject expertise and good knowledge of digital techniques was described as difficult.

> She is a very special breed of person because she is herself a [subject specified] scholar, among other things and an IT person and you know I think that's where the future lies. Not to the exclusion of other practitioners doing other things but you know you need new breed and people like that are rare. (P22)

Training was therefore indispensable for RAs, since PIs preferred to employ finding a person with subject expertise that could be trained in technical matters than the opposite, as understanding the subject-specific requirements of a digital project was considered vital. One PI explained the difficulties that could arise.

> I did try having the website designed internally by our own computer science department and it was a total disaster. The [features of the resource] are all circular in shape, when they had finished processing the images they were all oval, as if they had been sat on by an elephant and I said "But the image is oval." And they said, I quote them exactly "Oh it won't matter people won't know". (P14)

It was also important that the RA should have good communication skills in order to collaborate effectively with computer support staff, acting as a type of translator between the different groups.

> [the RA] actually built the database in Access and she consulted with the IT people at the university, IT team, and she has also been in touch with the AHRC data team and so she [. . .] comes out with all these specialists about

how to go about it but she actually learnt to use Access from scratch and build it all herself and so she has kind of been responsible for the technological side. (P21)

Despite the need for most RAs to learn new (usually technical) skills, both PIs and RAs thought the training available unsatisfactory. Many had been left with no option but to train themselves in the relevant technical skills—a time-consuming process.

There was no provision within the grant [...] to provide me with any more advanced training to do what I should be doing.
I had had database training but not at that sort of level [...] I have people come to me and it's like well you know they won't pay for us to learn how to do this because we ought to know how to do it. [...] I have done a MSc in Computing and Archaeology [...] I can program [...] but I have never done any Windows programming before, [...] so I was having to teach myself the Windows system on a live dataset, [...] that people were trying to use and that did sort of create problems. (P19, RA interviewee)

The absence of technical knowledge on the part of some PIs therefore meant that RAs might be expected to undertake tasks for which they had insufficient knowledge, or were not given enough time to train themselves in new techniques or to keep up with technical developments.

It also proved difficult to continue to employ RAs once individual projects had finished. Many PIs expressed frustration at training an RA to the desired level and then being unable to retain them for future projects.

[It is] very hard to find good assistants. I mean one of the bottlenecks in this kind of work is that it takes a certain odd combination of skills in order to be able to do it well, so you need to be very detail oriented, you need to know [the subject] and have a analytical certain mind and you have to be computationally literate. [...] I mean we have tried to get other people to do this kind of work [but] it's difficult to find people with

the right combination of skills and get the money to pay them [...] because the money comes and goes. I have been lucky because I have been doing this for a long time so I have developed all the appropriate skills [but] I am dependant on funding and funding or I have to go and do something else. [...] what you really need is a secure funding scheme so people can actually develop these kind of skills. But you can't just like take them off the street. (P1, RA interviewee)

In future, such problems might be addressed by providing more training in digital humanities techniques as options for students undertaking graduate courses in humanities subjects. For those who did not progress to work in academia, this would also provide important transferable skills for students in the employment market.

The unpredictable nature of arts and humanities funding usually means that experienced researchers cannot be retained and it is necessary to train new, inexperienced staff if the project is granted further funding. This means precious time is lost on new grants, and hard-won expertise is lost. This problem, however, was lessened for some PIs who worked at departments, or digital humanities centres where a critical mass of digital humanities projects were developed, allowing RAs to move from one project to another. Here a member of the Sheffield University French Department describes a rare example of career progression, where an RA has progressed from being a post-doctoral researcher, to faculty membership:

It's a whole new career structure that doesn't really exist in the humanities. How do research associates build on what they have done to a proper career? And recently [the RA] has had her post confirmed so she really is a fully fledged lecturer. So that's a happy ending and a sort of collaboration really of the potential for developing a whole set of innovative research schemes in a non-conventional career path which leads it into a conventional one: because you couldn't of course bring someone in with expertise which is just, well like gold dust really, very rare. (P22)

This kind of recognition of the work done by digital humanities scholars is still rare in the United Kingdom, but it is vital for their career development. It is also significant that this has happened in the context of a university where digital humanities research is highly valued. Thus, the issues of prestige, promotion and the recognition of digital humanities scholarship are evidently linked.

3.3 Dissemination

The strongest correlation that we found between a single activity and resources that were well used was in the area of dissemination of results. All the projects that were interviewed spent considerable effort in disseminating information. For the majority of academics, this was a completely new area, as most of them were used to writing books or other textual materials that would then be marketed by publishers. However, it was not surprising that some of the best known projects had the most varied and determined dissemination strategy.

Dissemination activities were varied and included sending out flyers to departments, libraries and archives and using email lists and the web. Most PIs gave papers at workshops, conference and seminars; in subject-specific and digital humanities domains, this was deigned to disseminate information to as broad a community as possible. The most unusual form of dissemination was reported as follows:

[...] one of the women from [the] village phoned me up a few months back and said, we want to make a tea-towel out of one of the pages of your book about the [archaeological site], do we have your permission? I said yes, if you give me a tea-towel, so it is still generating a product itself. Yeah, not many of our projects end up as a tea-towel. (P13)

Interviewees stressed that both producers of digital resources and funding agencies must realize the key role of dissemination for a project's success. When planning new projects, successful producers planned to allow time and money to make it possible for PIs to make presentations at conferences and workshops, and for RAs to publicize the resource on appropriate email lists or by other more traditional means. Enthusiastic promotion of a resource was a new but vital role that the producer must undertake on to ensure its continuing use.

3.3.1 Age of resources

Many of the well-used projects also tended to be relatively long-lived. It may be that early adopters have remained loyal to resources which they first used when little else was available It may also be that when little else was available, it was easier for producers to disseminate information about their resource, and thus it became well known. Early adopters and promoters of resources also tended to become experts to which people turned to for recommendations on digital resources.

[...] increasingly what people want is guidance through the huge number, [of digital resources] people are just bewildered by the amount of information that's out there and what to do with it. So I find that people have gone from just sort of saying, "Wow that's great that you have done this" to, "Yes that's great that you have done this but how does that work with the X collection or how do I incorporate that with these other things that are going on?" Basically give me a list of [...] your top ten. (P17)

The persistent use of older digital resources, even when newer, perhaps better ones become available may be explained by a commercial phenomenon known as 'switching costs'. This suggests that even when a better product is available, users will remain loyal to products or services that they know because the cost of switching to a new product is too great, either monetarily or in terms of effort (Yanamandram and White, 2006). In the case of digital resources, users may be unwilling to spend the time or the effort to learn how to use a new resource if the older resource continues to fulfil their needs adequately.

3.4 User contact

Very few projects had undertaken any type of user testing or had kept in contact with their users. Most projects were unaware of the amount and type of

use that their resources were experiencing. This is important, since our research with users showed that they had numerous concerns about the content, presentation and usability of digital resources, many of which might have been relatively easily remedied, had producers been aware of them. For example, they often failed to find even simple information about the purpose of the resource, and its intended use. This kind of requirement would have emerged through user testing, and such information would be easy to provide, and would significantly add to the user experience (Warwick *et al.*, 2008a).

Yet, all projects were interested in finding out more about how their resource was being used and by whom, and in receiving user feedback and had made some efforts to consider their users. We discuss these subsequently.

3.4.1 Designer as user

This was the most common method used. PIs believed that because of their subject expertise they understood the needs of users and could infer user requirements from their own behaviour. Although this method may uncover some user needs, it is not advisable. Since it is only possible to truly to know what users may need by asking them (Schneiderman and Plaisant, 2005) some projects also discovered that their audience consisted of a much more diverse group of users than the academic subject experts they had expected.

> Its impact is not what I expected. I remember for NOF we had to do a projection of visitor numbers and we kind of we just pulled them out, we had absolutely no idea and of course they were just tiny in comparison to what's actually happened and it has really taken off. You know I mean this is obviously just to learn how effective the internet is at spreading information but it just turns up in all the most unlikely places like you know the last week came an email saying ''Did you know that your websites been cited in arguments before the United States Supreme Court''. (P6)

A complex user population, including non-expert users, imposes new requirements of producers. Users at both of our workshops found that guidance for non-expert users was lacking in most projects, and would have welcomed more information about how to use resources. This does not mean dumbing down a resource, since, as this quotation shows, a non-expert in digital resource use may be highly expert in their subject area, but require some introduction to the use of a new digital resource.

A more formal user survey can also help to avoid making resources more complex than necessary.

> Looking back at it we should have really done a proper user needs sort of survey thing before hand rather than just launch into this thing. [. . .] a lot of it was complicated luxury that you didn't really need and so I think we could have saved ourselves a lot of work by going to the users to start with and saying you know ''What is it you would really like out of this?'' (P19)

This was confirmed by our workshop users who preferred a simple, uncomplicated interface to resources that required significant effort to learn how to use. This kind of design is therefore not advisable since it is difficult to design a resource based on the producer's own patterns of use, as this can lead to unexpected difficulties for potential users and ultimately lead to its neglect (HEFCE, 1999).

3.4.2 Informal user feedback

Some projects collected information from users from the comments and questions when presenting conference papers or at workshops. This method was especially common in projects concerning history. However, for this to be effective, the project must be at least partially developed and it can be difficult to make significant changes at this stage. It can also be difficult to derive any concrete design requirements from such informal feedback, where there is little active interaction between users and producers.

3.4.3 'Contact us'

Several projects made it possible for users to make comments, give feedback or to ask questions by e-mail feedback from the project website. The use made of this information varied. Four projects

used e-mail lists to communicate with their users. One project also used user comments to inform yearly updating of the resource. Another maintained a bibliography of publications written after use of their resource. As a result, they had an unusually clear idea of how their resource was used. This kind of feedback tended to be used by well-established projects, as a way of keeping the resources updated and staying in contact with the user community.

3.4.4 Direct user feedback gathering

Only two projects carried out organized user testing in the initial phases of project design. PARIP had carried out a user needs survey, and the Channel Tunnel project conducted focus groups. Three projects also took an informed decision not to carry out user testing. They adapted interfaces and systems that had worked well on similar projects or they were associated with a large digital humanities research centre and based their work on previous experiences.

Six other projects collected user feedback by testing a pilot version or a new software version release at workshops or by sending out e-mails to a known group of users for beta-testing. One PI commented that as a result of the funding body's insistence on user testing they had carried out more rigorous tests than they had planned to do.

Testing early releases of software or pilot projects on users is to be welcomed as later versions can be adapted in response to feedback. The disadvantage of not including users from the initial design phase, however, is that if significant changes are required to design or functionality at a relatively advanced stage of development, it may prove too expensive or time-consuming to make them. These types of tests usually concentrate on more technical aspects, whereas humanities users are also concerned with the provenance and the selection of the contents, as we found at our user workshops (Warwick et al., 2008a).

3.5 Documentation

Documentation was available for all but one resource. However, the degree of sophistication of this documentation varied considerably between projects (Warwick et al., forthcoming). The best documented resources tended to be older ones where the documentation was a vital part of the project's collective memory. Documentation practice also varied by subject. For example, archaeologists expect a resource to be documented, whether it is produced in physical or digital format.

> Yes well that's the sort of scientific paradigm, in a sense that if you are given a pile of Roman pottery then saying what you are doing while you are doing with it and documenting it is seen as part of the, you know, the rigour of the study. (P20)

This also applied to projects produced by archivists and linguists where documenting decisions is an expected part of the research process.

In many cases, however, documentation was fragmented and partial, consisting of e-mails, minutes of meetings, planning documents, and progress log books, most of which would only be meaningful to internal project staff. In some cases, original plans and documents had been lost. The resulting documentation therefore would generally be very difficult for someone not directly involved in the project to understand and would usually only cover certain aspects.

Although most projects were aware of the importance of documentation, they blamed time constraints for lack of better documentation. Since it was not a project deliverable or submitted for peer review, it tended to be neglected.

> I do remember it was quite fraught latterly because there were [...] the publisher's deadlines to meet and so on and ironing out the bugs. It was very much a seat of the pants business really. So it was very much operational really rather than, we didn't have the time [...] we were in new territory for us we were so anxious to get the thing done that we didn't really have the leisure or indeed the foresight to plot what we were doing. (P22)

3.5.1 Access to documentation

Documentation was usually difficult to access, especially when decisions were recorded in an informal way. Most documentation was kept by the PI or the

institution, and whether in paper or electronic format, its availability was generally not advertised. At one university, documentation was deposited with the library, but would still have been relatively difficult for an external researcher to access, first requiring him/her to know that such a despot had been made, and second requiring a visit to the library and permission to use the material.

In some cases (notably in Archaeology, linguistics, and archive studies), projects made documentation available either through the AHDS Archaeology website or from the project website itself. This represented especially good practice, as it was easy for users to find and access it. This is important, since users at our workshops repeatedly requested easy access to information about the content and provenance of a resource typically provided by just such documentation.

3.6 Sustainability

Maintenance and long-term sustainability was another area of concern. At the end of development, resources were either archived with the AHDS or backed up on the university server, yet few were actively updated. Most PIs seemed unaware that updating is vital in order for a resource to remain functional despite possible changes in software systems and delivery interfaces. A stark demonstration of the potential problems caused by this was provided by one former RA.

> **Male Speaker:** I was very concerned right from the very beginning ten years ago about who was going to maintain this and how it was going to stay available and how it was going to be updated that never really got resolved and as a result we are in this very unfortunate situation where the AHRC spent £200,000 whatever it was employing two of us for three years and within ten years of the start of the project half of it doesn't work anymore.
>
> **Interviewer:** Did you talk with anyone about trying to maintain it further?
>
> **Male Speaker:** Well I talked to my bosses about it yes and they worried about it a bit but in the end they just decided that the easiest way of doing that was to give it to these people in

Glasgow but I have no idea whether it's just sitting on a hard disc in Glasgow and nobody has touched it or whether there is actually anybody working on it.

> They weren't helped by the fact that the man [...] who did the web design for us then took early retirement. [...]
>
> Every so often I have guilty pangs in the back of my brain that I really ought to try and find out why the web interface has stopped working and whether we can actually get it up and running again but I am so involved in my own projects and this wasn't my project and it doesn't contribute to anything that counts to anything as far as I am concerned. It's not going to add to my RAE rating, it's not going to give me any value points in the [academic organisation] but I come out in a cold sweat every time I think about it and even the webpage of the site hasn't changed in six years. (P19)

Lack of funding to update the resource has, in this case, resulted in a digital resource that is almost unusable, despite the large investment made in its creation.

A few interviewees, however, did recognize the importance of updating and maintaining the website. Because a website that is not updated may indicate a resource that is no longer fully functional, as in the example above, users have rightly become wary of interfaces that appear dated, and use this criterion as a way of assessing the relevance and trustworthiness of a site. One of the few PIs to realize this observed that:

> I think it is important that you [update] partly because when you look at a website and it says last updated more than 12 months ago you just immediately think this is being allowed to wither on the vine and you don't trust it. So I want to be able to if nothing else to say on our homepage, last updated or we have the version number 4.2 you know date July 2006 is a way of assuring the users that we are still paying attention. (P6)

Another important issue is related to the life cycle of a digital resource. When the AHDS deposit

system was created in the mid-1990s it was modelled on a digital production system where social scientific data sets were deposited with a data archive once research was completed. Static data sets could be downloaded and used again by other researchers. Early digital humanities resources were also designed for CD, which is also a static medium. Now, however, most digital resources are delivered over the web and this old model of deposit seems to be no longer sufficient. In the case of most large digital resources, the data is no longer independent of the software or the interface that delivers it, and the changeable nature of web delivery means that a static resource produced at the end of a research project will become outdated relatively quickly and may become unusable, although project creators appear not to realize this. The same problem applies to archiving resources in digital repositories, since although the data may be deposited, most institutions do not have the resources to update interfaces to that data.

4 Discussion

We have shown that well-used digital resource projects in the humanities share a number of common features. They rely upon good technical support, researchers that combine expertise in technical matters and humanities disciplines, and a supportive institutional environment that fosters digital humanities research and promotes scholars who engage in it. Not surprisingly, many successful projects were, as a result, associated with digital humanities research centres. Remaining problems include the intractable difficulties of adequate recognition for the digital humanities model of team-based scholarship This includes both the problem of rewarding scholars who do not conform to the monograph producing single scholar research norm, and that of how to retain skilled researchers and ensure that their careers progress. We found notable success in these areas at the University of Sheffield, but this remains an isolated example in the United Kingdom.

However, the single most common uniting feature of well-used projects was their enthusiastic dissemination of information about themselves. This is a new demand on humanities scholars, but our study suggests that it is vital that projects promote themselves in as many ways, and at as many different fora as possible.

Our study also suggested that even in the case of well-used resources, some areas of practice could be improved, including organized user testing, the provision of and access to documentation and more effective methods for updating and maintaining resources.

User consultation was relatively rarely undertaken, despite the fact that it helps projects to design effective resources, and to avoid developing in ways that users may find over complicated or confusing. However, user testing, like disseminating information, is a skill that most humanities scholars have not acquired. It is therefore important that digital projects should be willing to work with those who already have expertise in this area, for example, researchers from Human Computer Interaction, Library and Information Studies, or practitioner librarians.

Funding agencies can have a vital role in this area by encouraging or even requiring digital humanities projects to include user testing within their planning, including allocating the time and funding required. One very positive result of the LAIRAH projects is that the UK JISC funding agency now requires a much greater component of user consultation in proposals to create or digitize resources.

Documentation was another relatively neglected area. Most projects understood its importance, but assigned it a low priority because it was not a deliverable. Only in disciplines where documentation was usual in general scholarship, such as Archaeology and Linguistics, was it routinely kept. Yet, documentation is vital, to maintain the institutional memory of a project and preserve the rationale of the design process. If scholars planning new digital resources can understand the decision taken by successful projects, they can avoid the repeated solving of the same problems, thereby saving time and money. Users also need to find the information that documentation contains about provenance and the selection of source material, yet it was often difficult to locate. If more projects adopted the

excellent practice of providing top-level links form their website to documentation, then users would be greatly reassured about the quality of resources with which they are unfamiliar.

Funding councils might also play an important role in this area, and might consider requiring at least a minimum level of documentation to be a deliverable of all funded projects. An agreed template would be helpful to producers from subjects not accustomed to keeping documentation and would provide intelligible information in a standard format for users.

Sustainability remains an intractable problem given the current models of funding and archiving digital resources. Yet, we have seen that money spent developing resources is wasted if they are allowed to deteriorate, because of a lack of funding to update and maintain them. The *de facto* solution is that individual institutions have become responsible for the electronic resources produced by their staff. However, although they may be willing to archive a static version of a resource in a repository and provide web server space, it is far more difficult for them to provide resources for active updating, since few institutional repositories have the expertise or personnel to ensure that the functionality of the resource is maintained. As a result, it seems likely that the slow decay of once functional digital resources will become more rather than less prevalent in future, at least in the case of the UK-based digital resources.

Some funding bodies, for example the Getty foundation, have begun to require electronic resource producers to find a commercially viable option for their project's sustainability once the initial funding has been used, by charging for all or part of their use. However, while this may work in the case of popular resources with a large user base, projects that are of interest to a smaller, more specialized user community are likely to find this model challenging. This model also seems likely only to work where resources are uniquely valuable, since attempts to make money out of web resources that are initially free, such as online newspapers, suggests that users will pay for digital content only if it is vital and cannot be found elsewhere for free (McCarthy, 2003).

The ideal solution is that funding bodies should make available small sums for continued updating. Although this is likely to prove unpopular, the alternative is that digital resources will slowly decay and become unusable for want of funds to update them. This would be a waste of the initial funding which is usually much larger than the cost of timely updating. Yet, the decision taken by the AHRC has forced institutions in the United Kingdom to take sole responsibility for sustainability of digital resources. The consequences of this may not be apparent for some years, but it is to be hoped that other national bodies may make more enlightened decisions.

5 Recommendations

We have developed a series of recommendations whose aim is to highlight the areas where digital resource projects have proved successful and those where changes could greatly aid their success. These have already been used to evaluate, and aid in updating one digital humanities project (Warwick *et al.*, 2007), and we hope that they will be of use to projects and funding agencies alike.

(1) Documentation: Projects should keep documentation and make it available from the project website, making clear the extent, provenance, and selection methods of materials for the resource. Funding bodies might consider making documentation a compulsory deliverable of a funded project. Discussions could be held between relevant stakeholders and the funding bodies, with the aim of producing an agreed documentation template. This should specify what should be documented and the level of detail.

(2) Users: Projects should 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, website feedback or other appropriate method. They should carry out formal user surveys, software and interface tests and integrate the results into project design. Applicants for funding should show that they have consulted documentation of other relevant projects and discuss what they have learnt from it in their case for support. The results of such contact could

then be included in the final report as a condition of satisfactory progress.

(3) Management: Projects should have access to good technical support, ideally from a centre of excellence in digital humanities. Projects should recruit staff who have both subject expertise and knowledge of digital humanities techniques, then train them in other specialist techniques as necessary. Funding bodies might consider requiring universities to offer more training to graduate students and RAs in digital humanities techniques.

(4) Sustainability: Ideally projects should maintain and actively update their interface, content and functionality of the resource, and not simply archive it with a data archive such as the AHDS. However, this is dependent on a funding model which makes this possible.

(6) Dissemination: Disseminate information about itself widely, both within its own subject domain and in digital humanities. Information should be disseminated widely about the reasons for user testing and its benefits, for example, via AHRC/AHDS workshops. Project should be encouraged to collaborate with experts on user behaviour.

6 Conclusion

By identifying successful and well-used digital humanities resources, and ascertaining the common approaches and problems faced by the producers of these resources, the LAIRAH project has highlighted possible areas that may predetermine whether a digital humanities resource will be well used by its constituent community (and beyond). By adopting the recommendations presented here, both funders and creators of digital humanities resources should have a greater chance of seeing the resource become usable, used, and known within the academic community.

References

Bates, M. J. (2002). The cascade of interactions in the digital library interface. *Information Processing & Management*, **38**(3): 381–400.

Bates, M. J. (1996). The Getty end-user online searching project in the humanities: report No. 6: overview and conclusions. *College & Research Libraries*, **57**: 514–23.

British Academy (2005). *E-resources for Research in the Humanities and Social Sciences*. http://www.britac.ac.uk/reports/eresources/report/ (accessed 29 May 2008).

Brown, S., Ross, R., Gerrard, D., Greengrass, M., and Bryson, J. (2007). *RePAH A User Requirements Analysis for Portals in the Arts and Humanities*. Arts and Humanities Research Council. http://repah.dmu.ac.uk/report/pdfs/RePAHReport-Complete.pdf (accessed 29 May 2008).

Barrett, A. (2005). The information seeking habits of graduate student researchers in the humanities. *The Journal of Academic Librarianship*, **31**(4): 324–31.

Dalton, M. S. and Charnigo, L. (2004). Historians and their information sources. *College & Research Libraries*, **65**(5): 400–25.

Duff, W. M. and Cherry, J. M. (2000). Use of historical documents in a digital world: comparisons with original materials and microfiche. *Information Research*, **6**(1) http://InformationR.net/ir/6-1/paper86.html (accessed 29 May 2008).

Duff, W., Craig, B., and Cherry, J. (2004). Historians' use of archival sources: promises and pitfalls of the digital age. *Public Historian*, **26**(2): 7–22.

HEFCE (1999). *Use of TLTP Materials in UK Higher Education* Report 99/39. http://www.hefce.ac.uk/pubs/hefce/1999/99_39.htm (accessed 29 May 2008).

Herman, E. (2001). End-users in academia: meeting the information needs of university researchers in an electronic age. Part 2. Innovative information-accessing opportunities and the researcher: user acceptance of IT-based information resources in academia. *ASLIB Proceedings*, **53**(10): 431–457.

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.

Huxley, L., Mullings, C., Hodos, T., and Jones, D. (2007). *Gathering Evidence: Current ICT Use and Future Needs for Arts and Humanities Researchers*. Arts and Humanities Research Council.

Makri, S., Blandford, A., Buchanan, G., Gow, J., Rimmer, J., and Warwick, C. (2007). A library or just another information resource? A case study of

users' mental models of traditional and digital libraries. *Journal of the American Society of Information Science and Technology*, **58**(3): 433–45.

McCarthy, K. (2003). This is the future of online newspapers, *The Register*. 9 July 2003. http://www.theregister.co.uk/2003/07/09/this_is_the_future/ (accessed 29 May 2008).

MLA (2007). *MLA Task Force on Evaluating Scholarship for Tenure and Promotion*. http://www.mla.org/tenure_promotion (accessed 29 May 2008).

Nicholas, D., Huntington, P., Jamali, H., and Dobrowolski, T. (2007). Characterising and evaluating information seeking behaviour in a digital (consumer) environment: spotlight on the ''bouncer''. *Information Processing and Management*, **43**(4): 1085–102.

Nielsen, J. and Tahir, M. (2002). *Homepage Usability: 50 Websites Deconstructed*. Indianapolis: New Riders Publishing.

Palmer, C. L. and Neumann, L. (2002). The Information work of interdisciplinary humanities scholars: exploration and translation. *Library Quarterly*, **72**: 85–117.

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 & Management*, **44**(3): 1374–92.

Schneiderman, B. and Plaisant, C. (2005). *Designing the User Interface: Strategies for Effective Human Computer Interaction*. 4th Edn. Reading, MA/Harlow: Pearson-Addison Wesley.

Stone, S. (1982). Humanities scholars — Information needs and uses. *Journal of Documentation*, **38**(4): 292–313.

Talja, S. and Maula, H. (2003). Reasons for the use and non-use of electronic journals and databases- a domain analytic study in four scholarly disciplines. *Journal of Documentation*, **59**(6): 673–691.

Warwick, C., Terras, M., Galina, I., Huntington, P., and Pappa, N. (2007). Evaluating digital humanities resources: The LAIRAH Project Checklist and the Internet Shakespeare Editions Project. *ELPUB2007. Openness in Digital Publishing: Awareness, Discovery and Access – Proceedings of the 11th International Conference on Electronic Publishing held in Vienna, Austria, 13–15 June 2007,* Vienna: OEKK-Editions, pp. 297–306.

Warwick, C., Terras, M., Huntington, P., and Pappa, N. (2008a). 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., Terras, M., Galina, I., Huntington, P., and Pappa, N. (2008b). Library and information resources, and users of digital resources in the humanities. *Program*, **41**(2): 5–27.

Warwick, C., Galina, I, Rimmer, J., Terras, M., Blandford, A., Gow, J., and Buchanan,G. (Forthcoming). Documenting digital resources in the humanities: the importance of technical and procedural documentation, *Journal of Documentation*.

Wiberley, S. E. (2001). Time and technology. *College and Research Libraries*, **61**(5): 421–31.

Yanamandram, V. and White, L. (2006). Switching barriers in business-to-business services: a qualitative study. *International Journal of Industry Management*, **17**(2): 158.

Notes

1 http://www.oldbaileyonline.org/ (accessed 29 May 2008).

2 http://www.shef.ac.uk/hri/projects/projectpages/gide.html (accessed 29 May 2008).

3 http://www.shef.ac.uk/hri/projects/projectpages/frenchstars.html (accessed 29 May 2008).

4 http://www.ucl.ac.uk/history/englishmonasticarchives/ (accessed 29 May 2008).

5 http://www.ucl.ac.uk/english-usage/ (accessed 29 May 2008).

6 http://vads.ahds.ac.uk/collections/LCFCA.html (accessed 29 May 2008).

7 http://ads.ahds.ac.uk/catalogue/projArch/eynsham_OAU/index.cfm?CFID=370757&CFTOKEN = 91009870 (accessed 29 May 2008).

8 http://ads.ahds.ac.uk/catalogue/search/fr.cfm?rcn = AVE_MILLET_BA-1 (accessed 29 May 2008).

9 http://www.arts.ed.ac.uk/europgstudies/rprojects/avant-garde/index.html (accessed 29 May 2008).

10 http://www.diamm.ac.uk/ (accessed 29 May 2008).

11 http://vads.ahds.ac.uk//collections/FSB.html (accessed 29 May 2008).

12 http://www.ucl.ac.uk/Bentham-Project/ (accessed 29 May 2008).

13 http://www.bris.ac.uk/parip/ (accessed 29 May 2008).

14 http://history.powys.org.uk/ (accessed 29 May 2008).

15 http://www.ucl.ac.uk/archaeology/cisp/ (accessed 29 May 2008).