Arts & Humanities Research Council

SUSTAINABILITY OF AHRC-FUNDED DIGITAL RESOURCES

Summary

1) Starting from the recommendations of the recent Review of the Resource Enhancement (RE) scheme, this paper surveys a substantial number of RE outputs, and discusses the issues of sustainability that arise from them, together with the related results of two projects funded by the ICT Programme. It recommends a substantial development of the present relationship with the Arts and Humanities Data Service (AHDS) in order to solve future sustainability problems: instead of acting simply as a repository for data outputs, the AHDS should act as coordinator of a network of expert centres who can provide the technical development for future digital resource creation or enhancement projects. This should ensure that the AHDS can maintain the resulting outputs in their full form for the long term, and that these outputs are informed by the right level and range of technical expertise. It also recommends that the Research Committee consider providing a modest additional funding stream to enhance the sustainability of existing RE outputs, after reviewing the results of a questionnaire survey of these to be administered by the ICT Programme.

Context and aims

2) The main ideas in this paper are substantially those in the previous, interim report on sustainability seen by the Research Committee. This started from the recent Review of the RE Scheme, and was based on extensive discussions with the AHDS, the ICT Programme Steering Committee, members of the AHRC Research Committee, RE award holders, and a wide range of experts and stakeholders in digital creation and preservation. The present version has added to it the results of a survey of completed RE projects and the two related projects funded by the ICT Programme. As a result its recommendations have been developed and firmed up. Timing has prevented discussion of this new version by the ICT Programme Steering Committee, but I am confident that it is substantially consistent with the Committee's discussion of the earlier version. Its recommendations are intended particularly to serve the new Strategic RE Scheme, which has been the main consequence so far of the RE Review; but they also concern future digital resources funded through the Research Grants Scheme, and existing digital resources funded under the Research Grants and old RE schemes.

3) It is worth emphasizing at the outset that the existing arrangement between the AHRC and the AHDS, requiring that all AHRC-funded data outputs be offered to the AHDS for open access and long-term preservation, is a model for other Research Councils and internationally. At present only ESRC and NERC have comparable data preservation policies; it may come as something of a surprise that the other Research Councils are only now beginning to develop policies of a similar kind. Looking at the recent report by the Office of Science and Innovation's e-Infrastructure Working Group, *Developing the UK's e-infrastructure for science and innovation*, we can see that the AHRC is ahead of the game in this respect. The report sets out the requirements for a national e-infrastructure to help ensure that the UK maintains and enhances its global standing in science and innovation; it is available at http://www.nesc.ac.uk/documents/OSI/.

4) Nevertheless the RE Scheme Review noted that sustainability was an issue for a 'significant number' of projects. As a result it made a range of recommendations for changes to the application process, including the exploration of a common technical framework for adoption by projects, and for reviews of existing digital outputs of RE awards (i) to establish their current status (active or not, deposited or not with the AHDS), and (ii) to quantify the additional resourcing and support required from the AHRC or other sources in order to ensure sustainability.

5) Taking up these recommendations, the present paper (i) reports on a survey the ICT Programme has conducted of the nature and status of completed RE projects, in accordance with one of the recommendations above; (ii) explores a number of issues that relate to the matter of sustainability; (iii) reports on two related *ICT Strategy Projects* funded by the ICT Programme, on the use of of internet resources in the humanities, and on peer review and evaluation of such resources; and (iv) proposes a substantial development of the idea of a common technical framework that the RE Review proposed. If the Research Committee is willing to take these proposals further, we intend to administer a questionnaire to carry out the further recommendation of the RE Review, concerning a survey to quantify the additional resourcing and support previous RE award holders may require from the AHRC or other sources to ensure sustainability.

6) The RE Review made no recommendation concerning alternative funding models, for instance involving charging, to help ensure sustainability. It did however conclude that 'only a very few' projects would have a sufficiently large user base to make charging feasible. This conclusion seems compelling, and accordingly alternative funding models are not explored in the present paper. The system that will be proposed is based on existing funding arrangements for research grants and the AHDS.

Nature and status of completed RE projects

7) A desk survey of all RE projects completed by the end of 2005 has been carried out by Heather Haskins, the ICT Programme Administrator. This has been accompanied by selective contacts with individual award holders. The number of projects (based on the information we have received) is 74; a full list, with the current location of the related digital outputs, is given in the Appendix. As the first Table shows (**Type of Resource**), almost all RE outputs for the period have been digital; only six are in printed form. Of the digital outputs, by far the largest group (49) are in the form on on-line searchable databases; many of these are multi-media, and generally involve complex user interfaces. The remainder are mostly either incorporated into existing library catalogues (7), or downloadable (5). One is available only on CD. A few are, for one reason or another, not yet fully available, or we have not been able to find them (but they have been contacted by the AHDS).

8) We have made a preliminary categorization of the different sources from which the resources are available in their full form, since different categories are likely to give rise to different sustainability concerns. As the second Table (**Location Type**) shows, only three (all on-line searchable) are held in their full form at the AHDS. 10 are incorporated into *university library* catalogues; 18 others are hosted by university or other *specialist centres*, the latter including two at the British Library. The largest category (24), is that of *Other university sites*, mainly departmental or similar web pages; this is the category with which we should be most concerned, since in most cases the responsibility for maintaining the resource is likely to lie with one or two individuals, and may be fulfilled only for as long as they are present and willing to carry out the task. The small *commercial* category comprises four commercial sites which make the resources available without charge: none of the resources surveyed operates a charging model.

9) Since the vast majority of the RE projects are available in their full form elsewhere than in the AHDS, and mostly on-line, the requirement of AHDS deposit mostly involves only the underlying data tables, not the full on-line resource including the user interface or front end. This is an important point to which we shall shortly return: in a great many cases the AHDS is simply not able to take over, without substantial additional work, a front end it has not developed itself. Table 3 (**AHDS Deposit**) shows how deposit has progressed. *Deposit* has either occurred or is in progress in 34 cases, *negotiation* or *chasing* is taking place in 28, and in six cases (typically where a university library catalogue is involved) deposit has been *waived*. The relatively large proportion of still-to-be-completed deposits is explained in part by the complexity of the process, and in particular the need for the data to be ingested in an appropriate form; it also reflects the slowness with which some grant holders have

responded to the request for material. Table 4 (**AHDS Availability**) shows that of the 22 projects that have been deposited three (as we have already seen) are available at the AHDS in their full on-line form (*Online no restrictions*), seven can be *downloaded* (data only) with or without restrictions, and 12 are *not for release*, or appear in the AHDS catalogue only as a *link* to the main holding elsewhere.

1. Type of Resource	
Output not digital	6
On-line searchable	49
Library catalogue	7
Downloadable	5
CD	1
Project page only	2
Temporarily unavailable	1
Not found	3

2. Location Type	
AHDS	3
University library	10
University specialist centre	7
Other specialist centre	11
Other university sites	24
Commercial	4
N/a	11
Unclear	4

3. AHDS Deposit	
Deposited	22
Deposit in process	12
Under negotiation	12
Chasing	16
Waiver	6
Output Not Digital	6

4. AHDS Availability	
Online no restrictions	3
Download No Restrictions	3
Download Restrictions	4
Not for Release	8
Link only	4
Deposit in process	12
Not yet available	28
N/a	12

Issues relating to sustainability

10) We need to distinguish between two aspects of sustainability, *academic* and *technical*: it involves keeping a resource (i) up-to-date in terms of its *academic* content, and (ii) available and fully functional within the *technical environment* in which it has been created and presented.

11) Academic sustainability is an issue whose importance varies from case to case: some resources are static or 'frozen', as far as their content is concerned, at the point of delivery, for instance an electronic edition of a text. Others are essentially dynamic, and quickly lose value if their content is not updated; an example is The Royal Historical Society Bibliography (http://www.rhs.ac.uk/bibl/), which is updated three times per year and requires the employment of one salaried full-time person. These are in a sense taken care of under existing funding arrangements, to the extent that it is open to resource owners to make repeated applications for AHRC support. Most resources are somewhere in between these two extremes, and require only occasional updating to keep up with scholarly developments. They present a widespread if not very grave sustainability problem, since there is as yet no regular funding stream to cover the occasional costs involved; updating depends on the willingness of those responsible to carry it out unfunded, or to obtain small grants from a variety of sources. The likely consequence is that most such resources will lose value rather slowly, but lose it nonetheless. While the AHRC does not yet have a funding stream to cover the relatively modest needs of this category, I recommend that it should consider creating one.

12) The problem of academic sustainability could in principle be solved quite easily if funding streams are found to cover it. I shall be concentrating mainly on technical sustainability in this paper, since the solution to the problem is rather less straightforward. Technical, like

academic, sustainability is a problem for some but not all digital resources. Where the data is simple in form, for instance a single electronic text, the current preservation system may be sufficient for purposes of technical sustainability: for as long as it remains deposited with the AHDS, a single electronic text can generally be made available for downloading. As we have seen, however, most RE outputs take the form of more complex on-line systems, and these include user interfaces or front ends on which much of the resource's intellectual content, and its ease of use by the research community, depend. In the few cases where such systems have been co-developed and are maintained by the AHDS, then, as long as the AHDS continues to be funded to do so, it will maintain them in their full functionality. As the RE Review has already observed, the problem arises where these on-line systems have been developed and are maintained outside the AHDS, on university or other sites, as is the case with the great majority.

13) How realistic is it to expect host institutions to guarantee full availability and functionality of these systems after the end of the grant period? The RE Review proposes a requirement to maintain a resource for three years after the end of the award; given the amount of funds invested, however, the AHRC will surely expect a significantly longer period of technical sustainability than that. Estimates are that some form of upgrading is likely to be required after three years, major upgrades after five years, and the resource is likely to be unusable after ten years unless significant work is undertaken to upgrade and migrate systems. To undertake this work properly requires a serious long-term commitment.

14) If we return to the different categories of location type listed in Table 2 above, we can probably expect such commitment from university libraries, and from some of the specialist centres in our list, notably the British Library. But not all the specialist centres will remain in existence for the longer term, and, as I have already indicated, serious cause for concern must arise from the largest category (24), of other university sites, as well as from the small category of commercial sites, none of which are supported by a charging model. In these cases not only the commitment but the expertise and knowledge available are likely to diminish as staff - particularly research assistants involved in the resource's creation - move on. One grant-holder we approached for technical information about his AHRC-funded database claimed to be unable to provide it, on the grounds that his technical assistant had moved elsewhere. On the whole I think the realistic solution in these cases can only be to develop procedures that do not rely on grant-holders and their institutions to sustain AHRC-funded digital resources in the medium-to-long term.

15) What happens when a host institution is no longer able or willing to maintain an on-line resource that it has developed? The underlying content should available through the AHDS, where it should have been deposited. As I have said, however, it is no simple matter for the AHDS to take over the whole on-line system, including the user interface or front end, in order to provide the same access to and functionality of the content for which the original system was designed. To do this will often involve redoing at considerable expense work already carried out by the original developer, in order to mount the system on the AHDS's technical platform. Since the AHDS is not currently funded to do this, we need to give thought to two issues: (1) should we provide for the AHDS to carry such work out for *existing* RE outputs, and if so how? (2) how could we avoid this kind of problem in the case of *future* AHRC-funded digital resources, particularly under the new Strategic RE scheme? I shall focus on (2) first, since the solution I am proposing to it will also be relevant to the solution to (1).

16) The RE Review suggests the exploration of a common technical framework for future RE-type projects, which seems to me a necessary condition for a solution to our problem, but not a sufficient one. In theory it would be possible to impose sufficiently tight technical specifications on grant holders, so as to ensure that everything they develop could if necessary be taken over by the AHDS in its entirety. But this could on the one hand be

unduly restrictive, and more importantly would in practice be difficult to implement effectively. Nor would the problem be solved by a blanket requirement to use open-source solutions, as some have suggested. While an obvious reason for adopting open source is that it is free and thus reduces costs, it would be unduly restrictive in terms of technical solutions to exclude the use of proprietary systems altogether. The AHDS currently accepts material in both proprietary and open-source formats for good reason, and by the same token is not able to preserve every kind of open-source system.

17) I would argue that the implementation of a common technical framework cannot be adequately achieved simply through formal technical specifications, but needs to be based on concrete structures and processes. Before suggesting what these might be, however, there are some further issues that need to be addressed if the matter of sustainability is to be dealt with properly. These are *quality assurance* and *reusability*, and *harmonization* or *interoperability*. If effort and resources are to be put into making a resource sustainable, one needs to ensure that it is worth sustaining: the AHRC should be seeking ways to optimize the digital resources it funds, not just to keep them going. Issues of *intellectual property rights* are also critical to the general sustainability question, but lie beyond the scope of this paper.

18) The present system of AHRC end-of-award assessment provides some *quality assurance* for the digital outputs of research grants, but is not rigorous from the technical point of view, and may even leave something to be desired from the academic point of view as well. We need to develop a more rigorous process for the *publication* of digital resources, as opposed to simply making them available on the web: what procedures can we develop for digital resources that could provide guarantees of quality comparable to those that a reputable publisher provides for printed books? Among other things, digital data projects seem to vary greatly in the extent to which they take the needs of users into account, other than those of the users involved in the creation process. Many may only be concerned to cater for their own academic interests, without thinking much about the issue of *reusability*, that of the other potential uses to which the data could be put. To cater properly for reusability the data creator needs to be aware not only of the data's potential academic significance, but also of the full range of technical methods that can usefully be applied to it: digital output projects need to be informed by *methodologies of use* as well as of creation and preservation. We need to ensure that this area of expertise, currently represented by the AHRC's ICT Methods Network, remains available to data creators in the future.

19) There is a further aspect of reusability which is not central to the AHRC's mission, but to which thought ought surely to be given in the resource creation process: the potential use of the resource for teaching and learning. Given the cost of producing the typical AHRC on-line databases, and the general desirability of promoting links between research and teaching, there is a serious case for seeking to facilitate their usability as resources for education as well as research. This is not to say that the AHRC should allow significant funding to be allocated to this purpose; it merely makes sense to try to ensure that front-end design is informed by an awareness of teaching and learning needs.

20) *Harmonization* and *interoperability* are important because, in addition to ensuring that the resources the AHRC funds are sustainable on their own, we should also aim to maximize their value by providing the means to connect them together. Excellent though many existing resources are, their value is significantly limited by their technical separation from one another: the more related resources can be connected, the more useful they are likely to be. Considerable intellectual and technological effort is currently being spent on the solution to this problem, mainly in science, technology and medicine, but increasingly in the social sciences and the arts and humanities as well. This effort is a large part of what the UK e-Science programme is about, a programme to which the AHRC is now contributing through the AHRC-EPSRC-JISC Arts and Humanities e-Science Initiative. But the technology

developed through this and related initiatives will not provide instant solutions to the problem of data dispersion. The goals of interoperability and harmonization will need to be pursued from both above and below: from above through the development of software solutions that facilitate the interlinking of dispersed data; from below by developing the data in formats, and on the basis of standards, that makes it more capable of being linked, alongside the provision of metadata that supports integration and harmonization and facilitates rich, deep querying of distributed resources.

21) The conclusion these considerations should lead us to is that future digital resource creation projects will need to be informed by the right kind of expertise not just in matters of data curation and preservation, but also with regard to the uses to which the resources can be put.

Two related 'ICT Strategy Projects'

22) Two of the recent *ICT Strategy Projects* funded by the ICT Programme have much to say on the same topic of sustainability, and tend to confirm a number of the points that I have made above.

23) Claire Warwick's, 'LAIRAH: Log Analysis of Internet Resources in the Arts and Humanities', conducted a web log analysis of the Humbul, Artifact and AHDS portal sites to determine the scale of the use of digital resources in the humanities, whether resources that are widely used share any common characteristics, areas of good practice, and aspects of project design that might be improved to aid greater use and sustainability. The web-log was limited, as it turned out, by practical problems, but was supplemented by a user questionnaire, interviews and workshops focusing on a selection of arts and humanities data projects. While all this adds up to a fairly limited evidence base, a number of the findings merit serious attention. These are, notably, the following:

a) Their evidence suggests that 30-35% of arts and humanities digital resources remain unused.
b) The title of the resource affects whether it is used or neglected, and should therefore be as unambiguous as possible
c) Users require high quality resources, both in terms of interface and content. If in any doubt about a resource's quality or authority they tend to abandon it.
d) Barriers to access deter many users. These may include having to download data, copyright permissions forms or an interface that is not easy to use.
e) Non-expert users found it difficult to understand the purpose of several resources. As well as an unambiguous project title, they required information about the contents, scope and how it was selected; the purpose of the resource; and advice about how it might be used.
f) Few projects kept formal documentation or made it easily available.
g) Few projects carried out formal user testing, and thus have little idea of the needs of their user community. Those projects which had carried out user tests were amongst the most well-used in the survey.
h) Staff who are knowledgeable both about humanities research and ICT techniques were key to successful projects. However, a lack of appropriate training meant that they were difficult to find, and scarce funding made them difficult to retain from one project to another.
i) Few projects realised the importance of ensuring their resource remained sustainable and that both content and interfaces must be maintained and updated. They did not appear to realise that archiving a resource with the AHDS does not guarantee its future accessibility.

These findings are mostly consistent with the points made earlier in this discussion about considerations of quality assurance and reusability. Despite the limited nature of their evidence base, they seem sufficiently commonsensical to be worth incorporating in the future awards process and a future version of the Technical Appendix. They are also consistent with the selective direct contacts we have so far had with completed RE projects, which have evidenced a degree of worry about the issue of sustainability, and a feeling that more expert technical expertise would have helped at the projects' early stages.

24) The second relevant *ICT Strategy Project* suggests ways of meeting some of the requirements on quality assurance that the first project suggests: David Bates, 'Peer Review and evaluation of digital resources for the arts and humanities'. Led by the Institute of Historical Research and the Royal Historical Society, the project's findings and recommendations are based on an online survey (777 responses), focus groups, interviews, and benchmarking reviews by specialists of a range of web-based digital resources for history, archaeology and classics. Its main findings and recommendations are as follows:

a) 71 per cent of respondents considered peer evaluation and recommendation to be either important or extremely important in their selection of digital resources for use in their personal research.
b) A kite marking system should **not** be adopted for research projects with digital outputs for the arts and humanities. Any system of evaluation or review should not adopt a simple 'pass/fail' approach when considering a digital resource in its entirety.
c) Instead, a multi-staged assessment process should be adopted after the completion of a resource but before its 'publication'. There is a role for learned societies and subject organisations in such a process. Exchanges between reviewers and resource creators should be made publicly available
d) A strong need for adequate project documentation was identified by many of those consulted by the project, both to assist the user and to inform the assessment process.
e) The system of peer review and evaluation proposed by this project is one means of ensuring that digital resources are properly assessed, and consequently amenable to consideration in, for example, the Research Assessment Exercise.
f) Common and widely-publicised citation standards for digital resources should be established.

This model of quality assurance seems well worth commending to future resource development projects. Some related findings on the issue of sustainability correspond to some of the concerns voiced earlier in this paper, and will be taken up in the proposals I develop below:

g) The issue of sustainability is perceived to be of vital importance, both academic and technical, by the respondents to their survey. It is the view of this project that sustainability is of such importance that some substantial investment ought to be made, and a devolutionary strategy devised, by the AHRC itself.
h) The Technical Appendix is no longer a reliable indicator of the robustness of methodology or project planning.
i) Full collaboration with the specialist humanities computing centres, such as the Centre for Computing and Humanities at King's College London and the University of Sheffield Humanities Research Institute (HRI), should be encouraged by the AHRC and other funding bodies—as project partners, not simply providers of consultancy or advisory services.

A common technical framework: structures and processes

25) Our survey and these projects together show that while the AHDS has been taking the necessary steps to ensure that copies are kept of the data produced by RE projects, in many

cases this will not have been enough to ensure the technical sustainability of the full on-line resource in the medium-to-long term. If there has been a failure here, it is not the fault of the AHDS, which has simply done what has been asked of it. It would be fairer to say that the whole grant-awarding process has assumed a model of sustainability, that of depositing data in an archive, which has not fully kept up with the problems posed by on-line searchable sites with complex user interfaces. If we are to make sufficient provision for the sustainability of future data resource projects, we need to ensure either that they are developed and hosted by organizations that can provide a serious and robust guarantee to maintain them in their full functionality for the medium-to-long term, or that the resources can be taken over in their full functionality by the AHDS should their original host institutions fail to maintain them adequately. We have seen, moreover, that the data-creation needs to be informed by expertise relating to a number of considerations other than mere sustainability, if the resource is to be truly worth sustaining: more steps need to be taken to optimize the quality of the outputs that are produced. All of this points to a solution in which the AHDS is not just a receptacle for data, but a significant collaborator in the resource development process.

26) AHDS has already pioneered a small number of collaborative solutions through its Archaeology Centre, in the form of database systems hosted by the Centre and jointly developed by the Centre and an academic grant-holder. See the examples at

http://ads.ahds.ac.uk/catalogue/specColl/hominids_ahrb_2003/
http://ads.ahds.ac.uk/catalogue/specColl/atlas_ahrb_2005/

Another example is the Stormont Papers project (http://stormontpapers.ahds.ac.uk/), in which three members of AHDS staff in the London Executive worked with a project team based in Belfast. In this model, the AHDS would provide a technical infrastructure, into which content could be placed, and an adapted and customized interface would be created in conjunction with AHDS staff for each grant-holder's needs.

27) This model is in some ways similar to that operated by the ESRC with its equivalent of the AHDS, the Economic and Social Data Service (ESDS), which works in conjunction with the UK Data Archive. The ESDS is generally the primary destination for ESRC-funded data outputs and the primary site for dissemination, and the cost of developing data to the required specifications can be built into the project grant. However ESRC-funded outputs, unlike those of AHRC grants, are generally held as files of numeric data which users can download and import into their statistical package of choice; creation of dedicated websites presenting and delivering ESRC-funded economic and social data is rare. This contrast reflects a difference in the kind of data produced, typically large-scale social-science datasets for the ESRC, and smaller-scale multi-media systems for the AHRC. It also no doubt reflects a difference in the technical competence of users: social scientists are comfortable working directly with data tables, arts and humanities researchers usually are not. The ESRC thus does not share the problem I have identified above of the technical sustainability of the user front end.

28) While a collaborative solution is needed, however, I am not proposing that the AHDS should be the only technical partner for future data creation projects. Although the key requirement must be that the whole resource should be capable of being maintained by the AHDS, insisting that the AHDS should carry out all technical development work would unnecessarily curb the creativity or flexibility of researchers. Just as importantly, there are other centres in the UK that have expertise in developing data systems comparable to that of the AHDS, and it will surely be desirable to take advantage of their expertise in future AHRC-funded development work. The best solution therefore seems to be one that the AHDS itself is now proposing, in which the key technical development role would be shared by a network of existing expert centres including the AHDS and coordinated by it. Grant applicants would designate their preferred partner in their grant application and provide in their applications for the necessary funds to support the partner’s involvement in their project. At the same time

the coordinating role of the AHDS should ensure that projects are developed in a form that the AHDS is capable of maintaining, if necessary, in its full functionality.

29) Discussions about establishing such a network are already well advanced between the AHDS and the King's College Centre for Computing in the Humanities, the Sheffield Humanities Research Institute, the Glasgow Humanities Advanced Technology and Information Institute, and The Centre for Data Digitization and Analysis at Queen's University Belfast. All these centres have shown strong support for the scheme. Others could also be included: it is important that the network should be as inclusive as is consistent with its main aims. The centres in this network would bring together research practitioners to work alongside research technicians, systems engineers, software developers, as well as specialists in preservation and what I have termed above the methodologies of use. The centres would also be encouraged to act as hosts for other content owners, including cultural heritage and memory organizations, and institutional library and archival collections, to ensure the development of a critical mass of interoperable content.

30) Centres would need to agree to support and work with a common set of tools, technologies, and technological infrastructure, and to work to agreed technical and metadata standards. New Centres would be able to join the network but would be subject to some form of certification. The Network would be coordinated by a central Executive who would ensure compliance with the technical infrastructure and standards. The Executive would also be responsible for the preservation of the content through a shared preservation infrastructure. Centres would be required to share any tools they developed to enhance searching, accessing, and using the digital information. Through the Executive there could be layered onto this model enhanced cross-search and discovery services, and, in the longer term, aggregation and harmonization of distributed content using web services and grid technologies. These services would provide:

a) aggregation of related resources by, for example, subject, theme, content type (including digitized information and objects, publications, catalogue records and bibliographies) that adds value through harmonization, clustering and adding additional information using grid technologies and advanced data and text mining techniques;

b) the application of ontologies and thesauri for discovery and use at a deep, rich level of granularity across multiple collections;

c) visual searching and browsing;

d) in general, the means of ensuring that data creation projects are adequately informed by what I have termed above the methodologies of use as well as of creation and preservation;

e) a service layer interacting with Google and other such Services to ensure full exposure of content.

The specifications of the infrastructure need not be particularly restrictive. It could include a variety of solutions, including proprietary systems, so long as the AHDS can develop, support and migrate them. The cost of the underlying technical infrastructure would be included as part of the core grant for the AHDS, whilst the cost of creating and customizing the front-end, and developing the functionality specified by the grant holder, could be built into the grant application. Different levels of support could be offered within this model, ranging from basic programming to create and customize the front end through to a full partnership. The development could be based either at the project's institution with help from AHDS staff, or at the AHDS.

31) When completed, the resource would either be kept at the AHDS, or mounted on the grant-holder's website and connected to the AHDS technical infrastructure to ensure cross-

searching and full interoperability with other AHDS collections. The AHDS would also monitor the need for updates and migrations across systems and hardware. These could either be done in collaboration with the grant-holder's institution or the resource could transfer to the AHDS in the event of the host institution being unable or unwilling to continue to maintain and upgrade it. Resources could be personalized for the grant-holder or his/her institution, and, even if kept only at the AHDS, could look to the world as if they resided on the grant-holder's site. The grant-holder could retain ownership of the resource, with the AHDS having the right to disseminate and update the technical infrastructure in which it sits. The key consideration is that the technical sustainability of the resource would be assured by the AHDS, with the minimum updating required to ensure continued accessibility and functionality covered by the AHDS's core funding. The AHRC could consider making funds available against further grant applications after the end of the project period for more extensive technical upgrading and academic updating.

32) It is obvious that having a network as a required development partner, rather than the AHDS alone, greatly increases the flexibility of solutions available to data projects. Equally importantly, the network organization can ensure that a common body of expertise, tools and resources is built up over time, so that each project can benefit from those that have gone before it. In particular the network would

a) be cost-effective, since technical development costs would be kept low through the use of standard tools and a network experienced staff, avoiding reinventions of the wheel;

b) help close the circle between creators and users of data, since AHDS could channel expertise on uses of data (e.g. from the ICT Methods Network) into the data creation process;

c) be able to liaise and share expertise with other agencies in the data development field, notably JISC and the Data Curation Centre;

d) take care of technical quality assurance--though a parallel system of academic quality assurance would need to be implemented;

e) help to ensure the resource's wide dissemination, through the AHDS catalogue system, and by appropriate exchange of information with organizations such as Intute Arts and Humanities;

f) increase the scope for harmonization or interoperability between different resources through the use of a consistent technical envelope;

g) give the AHRC added value from a service which it already funds.

Recommendations

33) The AHRC should encourage the AHDS to develop further the proposed network of expert centres coordinated by it. The network should be available to act as technical development partners in future data resource creation or enhancement projects. Its aim will be to ensure that project development is informed by the appropriate expertise and adopts appropriate technical solutions, so as to optimize technical quality and allow the project output to be maintained after completion in its full form, if necessary, by the AHDS. The network should be as inclusive as is consistent with this aim.

34) A new Technical and Sustainability Appendix should be devised which will not only ask for more detailed and robust guarantees of technical sustainability than at present, but will also seek to optimize the quality of outputs through formal quality assurance procedures, user testing, documentation, reusability, and, to the extent that it is feasible, harmonization

and interoperability. Without wishing to impose the specific quality assurance procedure developed by the related *ICT Strategy Project* and described above, I propose that projects should be expected to use procedures of at least equal rigour. The Appendix should also require detailed information as to how the project will ensure the academic sustainability of its output, in terms of keeping the content up to date.

35) This Appendix should be completed to the appropriate AHRC panel's satisfaction by all future Strategic RE projects, and all Research Grant projects whose output includes substantial digital resources intended for wider use. Panels should be rigorous in requiring the maximum reasonable guarantee of technical sustainability, as defined earlier; how far they should seek guarantees of academic sustainability could no doubt be decided case by case.

36) It need not be a requirement for projects to maintain on-line resources at the AHDS, but all should be required to demonstrate a guarantee of sustainability at least equal to AHDS maintenance. The requirement to deposit underlying data at the AHDS should remain,

37) Associating the network of expert centres as collaborators in an RE or similar bid would not be a requisite for funding. But the requirement to demonstrate a guarantee of sustainability at least equal to AHDS maintenance would be an incentive to work with the network, while not precluding other collaborators or host sites.

38) If the Research Committee is minded to consider the matter further, we will administer a questionnaire to existing RE projects in order to ascertain their need for additional resourcing and support to ensure sustainability in both the technical and the academic sense. While I would not expect any commitment at this stage from the AHRC to provide a related stream of funding, we would try to make some estimate of order of funding needed. There would be little point in our doing this, however, unless the Research Committee is prepared to give thought to the matter when the relevant information is available.

39) If funding is made available to update existing projects for purposes of sustainability, the same broad considerations should apply as for new projects: i.e. they should not be required to work with the network of centres, but the updating must either involve maintenance of the full resource by the AHDS, or an equivalent guarantee of sustainability.

DAVID ROBEY

February 25, 2007

Appendix 1. RE Awards completed by end 2005

Archer	Ian	University of Oxford	On-line version of Royal Historical Society Bibliographies on British and Irish History	http://www.rhs.ac.uk/bibl/bibwel.asp
Bailey	Geoffrey	University of Newcastle upon Tyne	Web Access to Rock Art: the Beckensall Archive of Northumberland Rock Art	http://rockart.ncl.ac.uk/
Bell	Maureen	University of Birmingham	The British Book Trade Index on the Web	http://www.bbti.bham.ac.uk/
Bishop	Julia	University of Sheffield	A Descriptive Catalogue of the James M. Carpenter Collection of Traditional Song and Drama	http://www.hrionline.ac.uk/carpenter/index.html
Breward	Christopher	The London Institute	Conservation, cataloguing and indexing of journals held as part of the EMap archive	http://voyager.arts.ac.uk/
Brown	Cedric	University of Reading	The Beckett International Foundation Catalogue On-line	http://www.unicorn.rdg.ac.uk/uhtbin/webcat/0/1/0
Bruce	Margaret	University of Manchester Institute of Technology	Digital resource For Textile Arts, Crafts and Design	currently offline; should be back by end nov/dec 2006
Burnett	Charles Stuart Freeman	University of London	A Catalogue of Medieval British Manuscripts Containing Commentaries on Aristotle	http://www.britac.ac.uk/arp/index.html
Charlton	David	University of London	The British Contribution to Series A/ii of Repertoire International des Sources Musicales (RISM)	http://www.rism.org.uk/ukandireland.html
Clarke	Elizabeth	University of Warwick	Web publication of the metadata from the Perdita project	http://human.ntu.ac.uk/research/perdita/index.html
Corrigan	Karen	University of Newcastle upon Tyne	A linguistic Time-Capsule: The Newcastle Electronic Corpus of Tyneside English	http://www.ncl.ac.uk/necte/index.htm
Cramp	Rosemary Jean	University of Durham	Enhancement of CASSS Digital Archive	http://www.dur.ac.uk/corpus/index.php3
Crawford	Michael Hewson	University of London	The Italic Epigraphy Project: Text and Monument	Output Not Digital
Creaton	Heather	University of London	London's Past Online: a bibliography of Greater London's history	http://www.history.ac.uk/cmh/lpol/
Crossley	Paul	Courtauld Institute of Art	Corpus Vitrearum Medii Aevi	http://www.cvma.ac.uk/
Cunliffe	Barrington Windsor	University of Oxford	The development of the Celtic Coin Index	http://www.writer2001.com/cciwriter2001/index.htm
Cunningham	Hamish	University of Sheffield	Generic tools for linguistic annotation and web-based analysis of literary Sumerian	http://www-etcsl.orient.ox.ac.uk/
Dack	James	University of London	Repertoire International de Litterature Musicale (UK operations)	http://pages.britishlibrary.net/rilm-uk/
Divall	Colin	University of York	Joining Tracks: enhancing academic access to the National Railway Museum Library	http://libcat0.york.ac.uk
Duffy	Eamon	University of Cambridge	Calendar of Papal Registers relating to Great Britain and Ireland. The Lateran Registers of Pope Leo X (1513-1521). Calendar with indices to be edited by Mrs AP Fuller	not found: AHDS has contacted
Edwards	Warwick	University of Glasgow	Early 18th Century Scottish Manuscripts	Output Not Digital
Eley	Penelope Anne	University of Sheffield	Partonopeus de Blois: an Electronic Resource	http://www.hrionline.ac.uk/partonopeus/
Ellis	John	University of London	TV Times Digitisation Project	http://tvtip.bufvc.ac.uk/index.php
Ferdinand	Christine	University of Oxford	The Works of William Congreve: Completion of D F McKenzie's edition (Oxford University Press, 2002)	Output Not Digital
Fraser	Valerie	University of Essex	Latin American Art: an on-line research resource	http://www.ueclaa.org/ueclaaOnline/index.jsp
Geddes	Jane	University of Aberdeen	The St Alban's Psalter: on the Web	http://www.abdn.ac.uk/stalbanspsalter/index.shtml

Goldman	Lawrence	University of Oxford	Victorian Social Reform: A Bibliography of the Published Papers of the Social Science Association 1857-86	http://napss.modhist.ox.ac.uk/index.html
Gray	Peter	University of Southampton	EPPI: Enhanced British Parliamentary Papers on Ireland, 1801-1922	http://www.eppi.ac.uk/
Harrison	Colin	University of Oxford	John Ruskin's Teaching Collections	http://ruskin.oucs.ox.ac.uk/
Higson	Andrew	University of East Anglia	The British Cinema History Research Project	http://www.uea.ac.uk/eas/britcin/
Hitchcock	Tim	University of Hertfordshire	The electronic Old Bailey Sessions proceedings, c.1670-1778	http://www.oldbaileyonline.org/
Hockey	Susan	University of London	From Archive to Researcher: a Generic Tool Set	http://leaders.sourceforge.net/
Horton	Mark Chatwin	University of Bristol	Papers of Isambard Kingdom Brunel (1806 - 1859)	http://www.bris.ac.uk/is/library/collections/specialcollections/archives/brunel/ikbrunel.html
Howe	Anthony C	University of East Anglia	The Letters of Richard Cobden (1804-1865) [FIRST GRANT]	Output Not Digital
Jefferies	Janis Kay	University of London	The Constance Howard Resource and Research Centre in Textiles	http://www.vads.ahds.ac.uk/collections/CHM.html
Jefferys	Kevin	University of Plymouth	Historical database of twentieth century local elections in Great Britain	http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=5319
Keay	Simon James	University of Southampton	Roman amphorae: a digital resource	http://ads.ahds.ac.uk/catalogue/archive/amphora_ahrb_2005/
Matras	Yaron	University of Manchester	Morphosyntactic typology of Romani: database enhancement	http://www.llc.manchester.ac.uk/Research/Projects/romani/
McDermott	Anne	University of Birmingham	The Johnson Dictionary project	http://www.fab24.net/jd100203/index_.htm
McRae	Andrew	University of Exeter	Early Stuart Libels: an electronic edition of political poems from manuscript sources	http://www.earlystuartlibels.net/htdocs/index.html
Mercier	Pascal	University of Sheffield	The Andre Gide on-line press archive	http://www.gidiana.net/GA.htm
Miers	David	Cardiff University	Law making in Wales: an on-line analysis	http://www.wales-legislation.org.uk/
Millett	Bella	University of Southampton	A trial electronic edition for the Early English Text Society	www.tei-c.org.uk/Projects/EETS/
Mulryne	Ronnie	University of Warwick	Digitisation of Renaissance Festival Books in the Collections of the British Library	http://www.bl.uk/treasures/festivalbooks/homepage.html
Munro	Julia	University of Reading	The Rural History Centre Library Cataloguing Project: assimilating the MAFF Library	http://www.rhc.rdg.ac.uk/webview?webviewinterface=21
Myles	Florence	University of Southampton	French interlanguage oral corpora	http://www.flloc.soton.ac.uk/
Nesi	Hilary	University of Warwick	British Academic Spoken (BASE) corpus	http://www2.warwick.ac.uk/fac/soc/celte/research/base/
Ovenden	Richard	University of Oxford	Catalogue of the papers of Sir Isaiah Berlin	http://www.bodley.ox.ac.uk/dept/scwmss/wmss/online/modern/berlin/berlin.html
Owen	Hywel Wyn	University of Wales, Bangor	Creating a place-name database in Wales: computerization of the Melville Richards Place-name Archive	http://www.e-gymraeg.co.uk/enwaulleoedd/amr/
Pagel	Ulrich	University of London	Digitisation and Access Enhancement of the Tibetan Dunhuang Manuscripts at the British Library	http://idp.bl.uk/idp.a4d
Parker	Robert	University of Oxford	Lexicon of Greek Personal Names	http://www.lgpn.ox.ac.uk/online/index.html
Peacock	David	University of Southampton	Stone in Archaeology: towards a digital resource.	http://ads.ahds.ac.uk/catalogue/archive/stones_ahrb_2005/
Pearn	Alison	University of Cambridge	Online calendar of the correspondence of Charles Darwin	http://darwin.lib.cam.ac.uk/
Poesio	Giannandrea	University of Surrey	Dance Data On-Line	http://libweb.surrey.ac.uk/DServe/DServe.exe?dsqApp=Archive&dsqCmd=Index.tcl
Reid	Norman	University of St Andrews	Cataloguing the Papers of Wilfred Ward (1856-1916)	http://specialcollections.st-and.ac.uk/projmss.htm
Richardson	David	University of Hull	The Trans-Atlantic Slave Trade: a revised and enlarged database	not found: AHDS has contacted
Rodger	Nicholas	University of Exeter	Guide to the Naval Records in the Public Record Office	not found: AHDS has contacted

Schlicke	Paul	University of Aberdeen	The Clarendon Edition of Charles Dicken's Sketches by Boz and other early sketches	Output Not Digital
Skretkowicz	Victor	University of Dundee	Dictionary of the Scots Language	http://www.dsl.ac.uk/dsl/index.html
Street	Sean	Bournemouth University	Project to Digitise the Archive of the Independent Local Radio (ILR) Programme Sharing Scheme	http://www.bournemouth.ac.uk/library/resources/special_independent.html
Taylor	Roger	De Montfort University	Photographs exhibited in Britain 1839 - 1865	http://www.peib.org.uk/
Trotter	David Andrew	University of Wales, Aberystwyth	The Anglo-Norman On-line Hub (Phase 1)	http://www.anglo-norman.net/
Turville-Petre	Thorlac	University of Nottingham	A Key to English Place-Names	http://www.nottingham.ac.uk/english/ins/epntest/keytoepn.html
Tyler	Colin	University of Hull	Unpublished manuscripts of the British Idealists; lectures by Cook Wilson; and one volume of reviews	Output Not Digital
Upton	Clive Stanley	University of Leeds	The Leeds Archive of Vernacular Culture	http://www.leeds.ac.uk/english/activities/lavc/
Walker	David	University of St Andrews	Dictionary of Scottish Architects	http://www.scottisharchitects.org.uk/
Warner	Anthony	University of York	The parsed corpus of early English correspondence	http://www-users.york.ac.uk/~lang22/PCEEC-manual/index.htm
Wathey	Andrew	University of London	The Digital Image Archive of Medieval Music (DIAMM)	www.diamm.ac.uk
Watry	Maureen	University of Liverpool	The Science Fiction Hub: a subject portal for science fiction studies	http://www.sfhub.ac.uk/
Welch	David Albert	University of Kent at Canterbury	The Digitisation of the Modern Cuttings Collection, Centre for the Study of Cartoons and Caricature	http://library.kent.ac.uk/cartoons/
Whyte	Iain	University of Edinburgh	Redesigning the City: The Percy Johnson-Marshall Collection	http://www.johnson-marshall.lib.ed.ac.uk/index.html
Willis	David	University of Cambridge	A historical corpus of the Welsh language	http://people.pwf.cam.ac.uk/dwew2/hcwl/menu.htm
Woodland	Christine	University of Warwick	Cataloguing the archives of the Trades Union Congress, 1970-90	http://www.warwick.ac.uk/services/library/mrc/ead/292d0700.htm
Worth-Stylianou	Valerie	Oxford Brookes University	CESAR a comprehensive online repository of French Theatre resources in the 17th and 18th centuries	http://www.cesar.org.uk/cesar2/