

AHRC ICT Methods Network Expert Seminar

SUSTAINABILITY OF DIGITAL RESOURCES IN THE ARTS AND HUMANITIES

Franklin-Wilkins Building, King's College, London, Wednesday 29 November 2006

On 29 November 2006, the Methods Network convened an Expert Seminar to address a range of issues concerning the long-term sustainability of digital resources, approaching these from a variety of perspectives. Representatives from the AHRC ICT Programme; the Research Information Network; the Andrew W Mellon Foundation; the Humanities Research Institute, University of Sheffield; an AHRC ICT Strategy Project; and PIs from major academic research projects presented papers which were followed by a series of extensive discussions.

The two principle concerns in considering the long term preservation and accessibility of scholarly digital resources may be described as academic sustainability – keeping a resource current in terms of its content, and technical sustainability - maintaining a platform for a resource to run on.

It is vital to collect *evidence of value* of scholarly digital resources to academic and other communities in order to justify the current and future funding of such resources. Qualitative evidence, demonstrating what such projects have achieved in terms of advancing scholarship, is crucial. More complex issues, such as whether all scholarly digital resources should be considered for perpetual preservation and maintenance, were just touched upon in this discussion.

This report outlines a series of responses made by participants to key themes in an attempt to identify the mechanisms necessary to effectively sustain resources including: establishing technical and other development models; developing funding models; creating a formal network of centres of expertise; establishing and maintaining trustworthy repositories; and (as part of the evidence of value) effectively promoting the value of scholarly digital resources.

The key point recognized at this seminar was that the consideration of sustainability issues is crucial to all stages of the life-cycle of digital projects and that there was a need for structure and policy frameworks which could be applied to the ongoing curation and maintenance of digital scholarly resources.

Addressing the sustainability issue will also mean addressing appropriate issues of training and awareness, including postgraduate training. However, it was not within the scope of this seminar to explore such issues. (The Methods Network is in conjunction with JISC developing the ICT Guides which will feed into Postgraduate Training. A two-year funded post has just been established to develop these Guides.)

Presentations

David Robey, AHRC ICT Programme, University of Reading.

This paper focused on sustainability issues that arose with regard to the digital outputs of research projects funded by the AHRC Resource Enhancement Scheme. Drawing upon ongoing discussions with the AHDS, the AHRC ICT Programme Steering Committee, RIN and other expert bodies and individuals, the paper focussed on ways in which applicants to the AHRC could work with the AHDS and other centres of expertise in order to achieve sustainable digital research resources. The output of projects funded by the AHRC Resource Enhancement Scheme since 2004 is primarily digital and this mass of materials constitutes a significant body of evidence of current practice in developing and managing digital resources.

Suzanne Lodato, The Andrew W Mellon Foundation, New York, USA.

The approach to sustainability issues by the Mellon Foundation, referring particularly to the Scholarly Communications funding programme (the scope and funding of which was expanded 1999) was the subject of this presentation. Sustainability is key to all aspects of the Andrew Mellon Foundation funding programme, particularly where the outcome is a digital resource. The Andrew Mellon Foundation places great emphasis on an iterative process when developing funding applications, retaining control of key issues such as sustainability planning throughout the grant

period. Since 2001 the Andrew Mellon Foundation have experienced a 57% increase in Scholarly Communications funding and a 560% increase in funding for digital humanities resources.

Discussion session led by Christie Carson and John Rink, Royal Holloway, University of London.

Christie Carson is PI for the AHRC funded project 'Designing Shakespeare' a resource comprising four databases: text; images; 3D models; audio and video interviews. Previously she was PI for the Cambridge King Lear CDROM Text and Performance Archive' (2000), a digital resource currently in danger of becoming obsolete; a process made additionally complicated by licencing issues. John Rink is PI for the AHRC funded project 'Chopin's First Editions Online' and the Mellon-funded 'Online Chopin Variorum Edition'.

Michael Jubb, Research Information Network, London.

This paper and the accompanying consultation document circulated at the seminar is based on a series of RIN internal discussions, and discussions with bodies such as JISC, the E-Science Centre and other key stakeholders. The document highlights issues around the stewardship of digital data for research purposes. One of the key concerns of the consultation document is to determine how to appraise the value of different kinds of digital research data and to manage the increasing mass of digital research output. A proposed framework for consultation comprises: Standards and Quality Assurance; Access, Usage and Credit; Benefits and Cost Effectiveness; Preservation and Sustainability; Formal and Legal Responsibilities. The RIN has a remit to coordinate and facilitate the development of both policy and practice in the provision of information services for the research community.

David Bates and Jane Winters, Institute of Historical Research, University of London.

This paper discussed the findings of the AHRC ICT Strategy Project, 'Peer Review and Evaluation of Digital Resources for the Arts and Humanities'. This project proposes to establish a framework for evaluating the quality, sustainability and impact over time of digital resources for the arts and humanities. The issues addressed are financial (funding long term maintenance), technical (with the responsibility on creators to take account of technical sustainability) and academic (funding and time needed from the academic community).

David Shepherd, Humanities Research Institute, University of Sheffield.

The Humanities Research Institute (HRI) at Sheffield University is one of the leading 'centres of expertise' in the UK and HRI aims to be a national and international leader in the creation of digital resources. The importance for sustainability of integrating the creation of digital resources with a wide range of scholarly activities was emphasized, as were the numerous partnerships, actual and potential, with users, libraries, publishers as well as within and between academic institutions; the principles and practicalities of sustaining digital resources are, or should be, inseparable from those of sustaining academic research in the arts and humanities as a whole.

Key Summary Points

Establishing Technical Development Models

It was recognized that the way in which individuals and organizations involved in the production of scholarly digital resources work together needs to be more clearly articulated. Crucial to this is developing full awareness and understanding of the work required at the intersection of the technical and the scholarly. A number of organizations and disciplinary groups in the UK and internationally have developed their own protocols to deal with some of these issues, but there is no overall articulation of what is being attempted and what the key principles are.

As far as AHRC-funded projects were concerned, the co-development of such projects between the AHDS, the Principal Investigator (PI) and the institution concerned was recommended, as was helping to establish effective partnerships between PIs and technical developers - incorporating the expertise of regional centres such as CCH, HRI, HATII etc. Such practice already exists to a certain extent.

Building sustainability discussions into the early stages of a project's technical development encourages thought about potential use and re-use of individual digital resources and increases the awareness that sustainability issues such as preservation, usability and reusability must be considered at all stages of the life-cycle of digital projects – from selection of materials to curation, maintenance and management of these resources. It was proposed that a more rigorous quality control process should be established for the creation and publication of digital resources, including ensuring that all appropriate steps are taken by projects to incorporate adequate documentation; incorporate user testing; ensure that web logs are kept etc.

To ensure adherence to technical requirements it was recommended that agreements should be firmly established at the outset of the funding process, or that effective monitoring during the project could be established, as a funder has

limited control once a project is under way. It was also acknowledged that applicants would be more likely to give such issues serious consideration if funding was conditional upon their doing so.

The AHRC ICT Strategy project on peer review and evaluation noted in their survey results that in determining the value of a digital resource, technical elements were not given as much serious consideration as might be expected. An important point, emphasizing that researchers do not always recognize or articulate the transformative effect of digital resources on their research practice, relying usually on such resources to simply provide increased or enhanced access to what is already available.

Generally (survey results indicated) there was also a low awareness of the importance of sustainability of digital resources.

It was suggested that a process should be established to inform projects of the crucial importance of technical development issues and it was assumed that this would be planned as part of the development of the AHRC Strategic Resource Enhancement scheme.

The issue of tools development was also raised and it was acknowledged that there are currently insufficient resources for developing tools in the UK. Such issues are being addressed by the AHRC/EPSRC/JISC Arts and Humanities e-Science Initiative, but were acknowledged to be problematic within the AHRC. It was noted that the Andrew Mellon Foundation supports the development of software tools and advocates sharing tools and developing tools in collaboration with other projects – a useful model to study. The Andrew Mellon Foundation also encourages development of open-source technology (usually Scholarly Communications staff require that open source technology be maintained for at least five years after the grant period).

Procedural and Rights Models

It was agreed that the Andrew Mellon Foundation's assessment criteria employed as part of their Scholarly Communications funding programme was a valuable model. Active evidence of 'intellectual leadership' is expected throughout the process of any application. Technical standards are agreed at the outset, or institutions receive a planning grant to establish technical standards before project implementation. The host institution must have the means to ensure long-term preservation of and access to any digital resource, whether in-house or through an external service. Internal collaboration, for example with librarians and digital technologists is expected, as well as cross-institutional collaboration. Understanding of user needs and rights must be clear. In addition a viable 'business' model developed with appropriate consultants, and financial support must be detailed. The Andrew Mellon Foundation has non-exclusive rights to distribute resources (that is, software and digital files) that it funds to the academic and cultural communities for non-commercial purposes. Award holders are prompted to think carefully about issues of rights and sustainability and seek professional consultation where necessary.

It was agreed that exploration of such models would be a key part of the agenda for taking the discussions forward into developing structural and policy frameworks.

IP Issues were discussed briefly, as this issue underpinned some of the difficulties in putting effective sustainability measures into place. It was acknowledged that it was not within the scope of this seminar to assess this issue in detail. One of the potential solutions is Creative Commons licencing and the suggestion of whether the AHRC might be able to help with the investigation of potential implementation of this was broached.

Funding Models

The concept of an iterative approach to funding, as practiced by the Andrew W Mellon Foundation, was discussed. The Andrew Mellon Foundation model is highly effective and UK funders might well consider adopting elements of this model in their own funding procedures. However, the fact that the Andrew Mellon Foundation solicits applications rather than issuing a call for proposals to the community was seen as a negative aspect of their practice.

An iterative process, particularly in terms of the technical aspects of a project would, by drawing on the expertise of the proposed network of expert centres, help ensure that all possible positive steps toward ensuring sustainability were taken.

One proposal made at the Expert Seminar was for a two-stage application process: an initial summary submission assessed for scholarly value; and a second, more detailed submission, incorporating a technical appendix; Peer reviewers would be chosen primarily for their subject expertise, but their ability to assess the technical elements of a proposal must also be taken into account. There was some concern however that an iterative approach and increased preliminary consultations would increase the amount of time needed to complete each proposal. It was suggested that FEC would make this more practical. Involving expert centres in the application process would also help smooth

the path of consultation. This is of course already partially in place as all AHRC applicants are expected to contact the relevant AHDS centre.

It was considered that funding to cover the initial costs of support from expert centres could become part of the AHRC funding package. Other potential funding streams include those currently available within some universities, who offer 'seedcorn' funding specifically for this process and those offered by the British Academy, which provides funding for pilot projects as a preliminary to submitting an AHRC application. It was acknowledged that there was resistance within the AHRC to a 'small grants' funding model, so the British Academy's model is invaluable. The British Academy also offers small grants to research projects in order to assist with the academic maintenance of projects, and is currently considering extending this scheme to cover database sustainability.

The issue is to find a balance between responsive and directed approaches to funding. Following the Mellon example levels of competition could be reduced by focusing more on Strategic mode funding, but this might not always be to the perceived benefit of the scholarly community

A Network of Centres of Expertise

Formal recognition of a UK network of centres of expertise was proposed and emphasis was placed on the importance of ensuring that the proposed network continue to support individual centres of expertise such as the HRI, CCH, HATII (to name but a few) to function as they do at present. It was felt that overwhelming centralization or standardization should be avoided and that it would not be appropriate for centres of expertise to be in any way a subset of bodies such as the AHDS.

It was suggested that the role of the AHDS in coordinating systematic and effective channelling of expertise from the network to support future resource development/enhancement projects be confirmed and further defined and that agreements be brokered with academics confirming that they would work closely with the network in keeping with agreed strategies and technical solutions. Either the AHDS or the host institution would then agree to undertake the maintenance of the resource for an appropriate period. The length of the 'appropriate period', needs to be defined and agreed.

Another aspect of channelling expertise, that is, the issue of disseminating and sharing project expertise more widely, was also raised. This would ensure that standards established and experience gained in one project (for example a project such as DIAMM, which as part of its output creates digital images of exemplary quality and rigorous metadata standards) could be made available to projects with similar aims and methods. Managing such 'channelling' effectively would require the expert knowledge and familiarity with appropriate projects which the UK's centres of expertise would have.

Potential models for integrating resources, such as the collaborative practices already in evidence within (for example) CCH and other centres of expertise could also be investigated.

Funding the support work of the network was an issue of some concern and it was suggested that this aspect could be coordinated as part of the AHDS mission, and included as part of the Full Economic Costing of research projects. (i.e. funding bodies such as the AHRC and JISC could cover the costs of supporting network partners as part of FEC).

A question was raised as to whether the individual centres of expertise would be expected to be funded under FEC or on a consultancy basis.

Other stakeholders such as librarians, archivists and information technologists should be incorporated into the network of centres and brought into any ongoing discussions. The work of digital libraries and librarians over the last ten years should also be explored.

Perpetual Repositories?

There was some concern that academic institutions would not be able to guarantee that they could maintain resources in perpetuity although in general terms it was thought to be more effective to trust curation of resources to institutions rather than individuals.

It was suggested that the curatorial role of the AHDS could be enhanced i.e: their role extended from depository to maintainer of online resources. At present many resources do not deposit fully with the AHDS, and even supportive institutions (e.g. CCH, HRI) cannot guarantee perpetual support to projects, particularly if suitably qualified staff (both technical and academic) are no longer available. The AHDS is in a singularly appropriate position to be able to offer support, at least on the technical maintenance side. The current development of a collections management database

by the AHDS and the increased emphasis the AHRC will place on depositing data with the AHDS should also help in developing the AHDS' role.

The importance of educating PIs was emphasized. Although it is self-evident that scholars would benefit ultimately from the creation of sustainable materials, academics are not always concerned with sustaining their projects over a long period. Those who are not interested in maintaining a resource over a long period should be made aware that they must give serious consideration to sustainability issues, perhaps even more so than those who are dedicated to maintaining a resource that they have created.

Five university repositories exist although only two have any significant holdings, and even in these institutions it is clear that policies and procedures need to be defined and implemented. It was however agreed that national institutions such as the British Library would also be likely to be trustworthy repositories.

In a recent survey the RIN found that universities (with some notable exceptions) tended to shy away from issues to do with curation and preservation of data as something that was beyond their expertise. However now that the UK is in the development stages of an institutional repository movement – a movement driven by institutional librarians (and JISC are currently investigating the creation a national infrastructure for digital repositories) this situation may change.

Maintaining centres over the long term is an issue which must continue to be highlighted and it was suggested that income generation – selective charging (not for profit) for the use of a resource - should also be considered as one possible support mechanism. However it was hoped that the current sustainability issue could be resolved within the framework of current funding models.

Promoting Value of Digital Resources

It is still necessary to contend with a lack of awareness, a lack of understanding, and a lack of acknowledgement of scholarly digital resources, although of course, resources hosted by recognized centres of expertise do have an implicit guarantee of value. Is there a need to certify centres of expertise in some way?

Peer review is well established for analogue resources, but not for resources in the digital domain. Even where peer review does take place a division between the technical and academic aspects persists. The use of and placing value on digital resources is becoming more common but it is not embedded in the research culture. This process would be facilitated by establishing a recognized mechanism of assessing the value of scholarly digital resources.

This leads on to the issue of citation - and in particular the problem of citing a regularly updated digital resource. Is it possible to recover citations from an earlier version of a digital resource? Creators of digital resources should be encouraged to provide clear citation guidance. Following on from this, researchers may well be using digital resources to inform analogue outputs, but there is no means of assessing this if it is not acknowledged, and formalizing citation guidelines might help, by raising awareness that the scholarly significance of an academic digital resource is equal to an analogue resource.

It was recommended that more investment in training researchers both to use and to evaluate digital resources should be made. For example scholarly journals could commission reviews of digital resources and encourage authors to cite digital material. There are portals in place such as Intute which do review resources, but these are not intended to be of peer-review standing.

Research councils and other funding bodies should be encouraged to conduct post-completion assessments of projects and to publish the results. Although it would be the academic content that was primarily under scrutiny, reviewers should also have the ability to assess the technical aspects of projects and a framework for peer review and evaluation would help support this process.

Hazel Gardiner
Senior Project Officer
AHRC ICT Methods Network
February 2007