Digital Sustainability and Digital Repositories

Kevin Bradley
Sustainability Officer
Australian Partnership for Sustainable Repositories
National Library of Australia
kbradley@nla.gov.au

Abstract:
The tasks associated with managing and backing up digital data are well known to IT managers, but the mere presence of the data stream is not the only criterion for preserving and maintaining digital content. Digital sustainability recognises that the continuity of digital information goes well beyond basic storing and managing of data and is integrated into the lifecycle of the information object. It includes technical, social and economic considerations. In 2004 a DEST-funded project, the Australian Partnership for Sustainable Repositories (APSR) initiated an investigative process to establish a centre of excellence for the management of digital collections. The APSR has an overall focus on the critical issues of the access continuity and the sustainability of digital collections, and this paper draws on the investigation of these issues in the University sector.

Introduction

Routine IT management tasks such as backing up digital data are so fundamental to the management of data that even relatively unsophisticated home computer users are aware of the imperative. Though this routine task is a necessity to any ongoing retention of data, and may possibly prove to be comparatively routine in the long term, it will not of itself ensure access to the content. There are more complex issues associated with the preservation of digital information which are concerned with access to meaningful and comprehensible information rather than just bit streams.

The apparent fact that digital information technologies have superseded, or at least extensively supplemented, the paper-based methods of documentation and timely communication in universities, as elsewhere, and the manner in which they have been incorporated into the processes of peer review, has led some commentators to describe this as "a new model" of scholarship (Smith 2003). If this new model of scholarship is to survive, or if digital resources are to be incorporated into the old model, there is a necessity for the apparently fragile digital research and publication resources to endure. To this end, the Australian Partnership for Sustainable Repositories (APSR) was established in 2004 to investigate sustainability of digital research outcomes in the university sector. Digital sustainability recognises that the continuity of digital information goes well beyond basic storing and managing of data and is integrated into the lifecycle of the information object. This includes technical, social and economic considerations, and these concerns now underpin the design and implementation of our repositories.

APSR

The Australian Partnership for Sustainable Repositories was funded by the Department of Education, Science and Training (DEST) as part of the Systemic Infrastructure Initiative, to promote excellence in sustainable digital repositories. Its work commenced in May 2004 and the work under the current round of funding will be completed by December 2006. The partnership complements and collaborates with the work done in the Meta Access Management Project (MAMS), Australian Research Repositories Online to the World (ARROW), and Australian Digital Thesis Program Expansion and Redevelopment Project (ADT). APSR also benefits from and collaborates with other digital sustainability initiatives worldwide. Wherever possible, APSR documents and adopts existing international standards or is actively involved in developing and implementing emerging standards.

Terminology

Repositories are not just the technology that manages data; in fact, the word repository is used here to describe the broad socio-technical make-up of the repository. Repositories are "composed of people, activity, artefacts and technology"(Bishop, House et al. 2003), which characterisation necessarily includes social relations, user expectations, funding and resources, organisational stability and even the distinction between hardware and software. A digital object storage system, or a digital mass storage system refers to the hardware (and included software) systems that store the data, and repository software describes the software technology that is developed for the purpose of centralising, preserving, and making accessible an institution's intellectual capital (Johnson 2002). Taking into account the full

gamut of socio-technical infrastructure that constitutes an institutional repository will allow us to ensure the sustainability of the content.

Sustainability and sustainable has invaded the English language. The Victorian Government has a Department of Sustainability and the Environment (http://www.dse.vic.gov.au/dse/) and the ACT Government has an Office of Sustainability which is "committed to creating a sustainable Canberra" and "developing, facilitating and coordinating the implementation of guidelines, policies and procedures related to sustainability" (http://www.sustainability.act.gov.au/). In both cases the actual object of sustainability is unclear, but seems to imply something of economic and environmental sustainability. In the face of such obfuscation, or at least haziness, it might be best to be clear about the usage of the word sustainable in the context of digital preservation.

In the digital repository sense, we use the word to mean building an infrastructure, both social and technical, which is economically viable for maintaining valuable data without significant loss or degradation. This includes, as has been discussed above, the whole socio-technical composition of the repository. Clearly, it is not possible to preserve digital information without a sustainable organisational, economic, social, structural and technical infrastructure.

Digital preservation, on the other hand, is specific in its aims: to preserve the viability, renderability, understandability, authenticity, and identity of a digital item. The focus of digital preservation is on the item; its concerns are with the unique and individual characteristics of each digital object. Sustainability and preservation are two different aspects of the same problem, they are not, however, coterminous terms, they each refer to different issues.

Partners

The APSR is structured around test bed projects, which test and analyse issues associated with local implementations undertaken by the partners, as well as central projects concerned with sustainability and the communication of issues associated with sustaining digital information.

The partners in APSR are the Australian National University, the University of Sydney, University of Queensland, the Australian Partnership for Advanced Computing (APAC) and the National Library of Australia. APSR partners are committed players in the field of sustainable digital repositories. The partners participate in APSR to enhance and extend their existing work and make it useful and relevant to the sector at large. The university partners run test beds, where they use their own repository environments to trial APSR initiatives. Specialist expertise is provided by the National Library (digital preservation) and APAC (e-data sets). All the partners contribute to the central sustainability and outreach programs.
The test bed partner projects are complex and comprehensive, based on university-wide initiatives. They incorporate consideration of community-specific repositories, existing repositories, as well as the investigation and development of more broadly applicable repository software systems. Consequently, all partners are involved in, and drawing on, much more than the development of the specific repositories. Nonetheless, these repository software systems are critical to the development of sustainable digital data and reflect on, and lead to, important developmental work, and each of the university partners have produced instances of software or repository tools.

The Australian National University (ANU) has concentrated on developing the open source repository software "DSpace". ANU programmer Scott Yeadon is a committer to the DSpace project, has been responsible for coordinating at least one major release of DSpace in the last 12 months, and has been instrumental in all the releases that have occurred in that time. The development has resulted in a number of outcomes: an easily installable version of DSpace, or "DSpace in a box", the DSpace-Cocoon trial for encoding complex digital objects, a critical issue for the preservation and re-use of digital items whose complexity make sustainable access processes problematic, user-driven batch upload, a statistical analysis tool and an image thumbnail (derivative) generator. There has also been collaboration with the federated search and access systems, PictureAustralia and MusicAustralia and the development of a draft ANU Institutional Repository Policy. (Raftos 2005) The close involvement in the development of code has allowed APSR to indirectly influence issues that impact on sustainability and preservation. Two of the sustainability projects (see below), the Automatic Obsolescence Notification System (AONS) and the PREMIS (PREservation Metadata: Implementation Strategies) metadata functional specification and development will build on the relationship.

The University of Queensland (UQ) has concentrated similar resources on the development of a repository based on the open source FEDORA repository software. The development has resulted in a flexible, digital repository and workflow management system known as Fez, and which is available under open source license from links on the APSR website http://www.apsr.edu.au/fez.htm. UQ's programmer Christiaan Kortekaas led the development of the Fez FEDORA implementation which has received critical acclaim. The system incorporates a search engine, registration process, level III authentication, and is designed to help university repositories comply with the Research Quality Framework (RQF) requirements. This latter point recognises the economic issues associated with sustainability (Weaver 2005). The UQ's Fez will also be a test site for the AONS project and the PREMIS implementation.

The University of Sydney has identified a critical need in drawing together a number of disparate, pre-existing repositories. To address this issue it is developing an interoperable open source, OAI-compliant middleware solution to enable digital objects (text, image, sound etc) in diverse discipline-based object repositories to be searched, discovered, retrieved and, if required, transformed or integrated. The middleware is currently in testing in image banks in Medicine, Veterinary Science, Archaeology, Architecture and soon Botany, as well as tests on the PARADISEC archive. The concept and initial design were developed by Ian Johnson, and further development work is now being coordinated by Daniel Burn. More information about the middleware can be found at http://www.ispheres.org/

The sustainability stream of APSR is led by the National Library of Australia. This part of the project has been successful in defining the issues that are critical in managing a sustainable repository and some of these are discussed below: a detailed analysis of the issues is available at the APSR website http://www.apsr.edu.au/publications/WPapers.htm. The earlier part of the sustainability project included a survey of data collections in the university environment, and the results of this have been used to inform the design of various parts of the APSR projects. These include the Automatic Obsolescence Notification System (AONS), the functional specification of the PREMIS metadata set and a project to work with research communities on their data management needs.

Automatic Obsolescence Notification System (AONS)

Ensuring long term access to digital objects within institutional repositories is a massive problem for libraries, archives and universities, and its solution is one of the key objectives of the APSR. DSTC's PANIC (Preservation Architecture for New Media and Interactive Collections) project has been working on an integrated, extensible architecture based on preservation metadata, automatic notification services, software and format registries and semantic grid services - that offers a sustainable, dynamic approach to the long term preservation of large collections of heterogeneous scientific data. The first step in implementing an automated digital preservation system is obsolescence detection and notification.

The objective of this project is to evaluate the obsolescence detection and notification component of PANIC through a collaborative project between DSTC, APSR and the UK Digital Curation Centre (DCC) and to deliver a working prototype for installation on APSR Repositories. The project leverages off the experience and development expertise of the DSTC and the work being carried out at the DCC.

For the purposes of this project, we will investigate building the obsolescence detection and notification module on top of the UK Digital Curation Centre's Representation Information Registry (or a local mirror of this):
http://dev.dcc.rl.ac.uk/twiki/bin/view/Main/DCCApproachToCuration).

When there is an incompatibility between an object's current preservation/formatting metadata and the latest information recorded in the registry, then a message is sent to the nominated person(s) or software agent, notifying them of a potential risk.

Requirement Specification of Preservation Metadata Capability in Open Source Repositories

As has been identified (Wheatley 2004), none of the of the open source repositories has the capability of managing the metadata to ensure appropriate sustainability, and few have been identified as considering this a priority. The objective of this project is to develop a requirements specification for preservation metadata with the intent of coding the requirements into appropriate repositories as a subsequent stage in this process. The process will be informed by the Preservation Metadata Project being presently undertaken.

The primary cost in the process of making a coded change to a software repository is generally in the specification and modelling of the requirements, certainly until that is clarified, the process cannot be costed. The development of the requirements specification will be undertaken in consultation with the lead programmers in the partner institutions so that the implementation can be undertaken as soon as is practicable. Though the specification is implementation independent, it is very likely to find a practical application in DSpace and FEDORA, and it is therefore proposed that the specification be developed with the support of the DSpace committer and the developer of Fez.

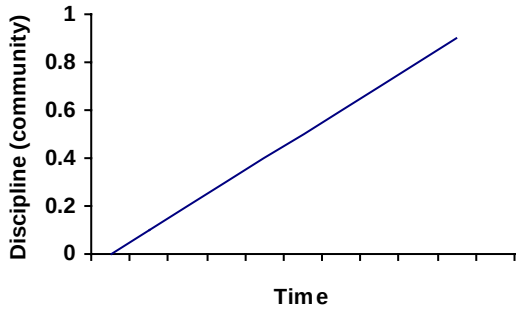
A project to work with research communities on their data management needs

A clear distinction in approach has been identified between specialist data intensive repositories and general or institutional repositories. That there are different approaches to data management resulting in different requirements, design and workflow approaches is clear. The actual distinction however, is not so easily determined, as the technologies, issues and processes are all, more or less, shared. Nonetheless, knowing the degree to which a data manager has to be a subject specialist to manage data, or a library and information management specialist, is a critical to APSR. It goes to the initial stages of data repository funding and responsibility.

Specialist technical requirements

Critically, a number of scientific data sets claimed that maintaining responsibility for their data was essential to the sustainability process. Many of the scientific data managers could not envisage a role for a general repository for the specialist data. The need for specialist knowledge led most of these data set managers to employ subject specialists rather than information management specialists.

The data specialists argue that decisions about the management of the data can alter the meaning of the content, and that only a subject specialist is adequately equipped to make those decisions without compromising the content of the repository. This may be the case, but alternatively, it may be that the explicit metadata necessary to make a dataset transportable has not yet been developed or made, and that local expertise compensates for the lack of documentation. It can be argued that the need for metadata increases with the distance from either the community, or the creation date (see graph below).



The need for metadata: the greater the separation in time or discipline, the greater the need. (DCC UK)

As an indicator for APSR deliberations, it might be that if the data cannot be managed away from the designated community, then it will consequently not be possible to sustain the data when there is a large temporal distance between its use and its creation.

Institutional or General Repositories

A design tenet of specialist repositories is that system design should divide on discipline issues rather than on a technical basis, which requires domain expertise and knowledge. However, an institutional repository is predicated on the principle that consolidating technical services will deliver economically sustainable efficiencies, and a concentration of expertise around the issues that inhere in digital preservation and sustainability.

As a consequence, the data management and research communities project will investigate these and other issues, while supporting the development of sustainable practices in those areas.

Issues

There are a number of issues which have been raised in the APSR project so far, more than can be comfortably considered in this paper but which are available at http://www.apsr.edu.au/publications/WPapers.htm. A selective summary of the issues is considered here, as something of an appendix to the projects, and a justification for the work undertaken.

Data sustainability

The sustainability of the raw data is the retention of the byte-stream in its proper and logical order so that it can be delivered as required identically to how it was first deposited. This is simply the issue of managing and duplicating the data, a task which is fundamental to all data sustainability. Though fundamental to all digital archiving and retention practices, the process still lacks official standardisation.

In spite of the lack of internationally agreed standards for data back up and safety most IT professionals are in general agreement about the processes necessary to produce safe, redundant copies of data. Most IT professionals would be able to measure whether they meet a notion of best practice, and are even likely to call the process "archiving". Many archives hope that by this approach the retention of the byte-stream in its proper and logical order will be achieved, so that it can be delivered as required identically to how it was first deposited. Nearly all digital preservation experts simply cite the Cedars' Guide to Digital Preservation Strategies which states that "preserving a byte-stream is a relatively routine IT operation." (Cedars 2002).

Data backup has certain risks associated with it, especially when the refreshment of the data incorporates a change in physical technology. The data format may remain the same, but the act of restoring the data to its new carriers exposes it to the risk of undetected, or at least unnoticed, changes to the data or metadata. Though automated systems will detect errors, there is a need in all digital archiving communities to identify an owner of the data. The data owner authorises non-routine actions which may become necessary. The identified owner determines whether the action has been successful. The data owner determines the quality status of any changes to the data.

Clear responsibility for the preservation of the content is a mandatory requirement of an archive. The UNESCO Guidelines for the Preservation of Digital Heritage (Webb 2003) state that "Responsibility is a crucial issue in the preservation of digital heritage. The starting point for action is a decision about responsibility." This issue is one at both the individual and the institutional level. The socio-technical nature of digital sustainability applies at even the most technical level.

It should be emphasised that the technical data management issues are not the main problem in digital preservation or sustainability. The technical problem of ensuring that the data is retained in its logical order is one that can be demonstrated to be achievable. However, the

lack of standardisation and the absence of any testing of possible data refreshment and system migration scenarios means that there is some unnecessary risk and uncertainty associated with long term data management even before the major preservation issues are addressed.

The sustainability of meaningful access to the content

The major issue associated with sustainable access to data is accessing the content, so that not only the data is retained, but also it can be rendered in future technical environments and can be comprehended in a manner commensurate with the creators' intentions and user expectations.

The ability to sustainably and accurately render an item regardless of how the technological environment has changed is the fundamental aim of digital preservation. There is no better measure of the success of the preservation of a digital object than the ability to sustain access to a digital item. APSR has an "overall focus on the critical issues of the access continuity and the sustainability of digital collections". Access by the designated community is a critical measure of sustainability; access to an authentic rendering of the informational content in both the short and long term must be the primary aim of any digital preservation program. This is stated succinctly in the UNESCO Guidelines for the Preservation of Digital Heritage: "Digital materials cannot be said to be preserved if access is lost"(Webb 2003).

Digital preservation practitioners describe various strategies for maintaining access to the content in a digital repository, all which can be categorised as a form of migration, emulation or encapsulation. When the operating systems, the software environment, the access tools or the formats change, access to the content is no longer routine, and some alternate strategy will have to be applied to ensure access.

The most useful tool in this process is the use of appropriate preservation metadata. Much of the digital archiving community considers that the preservation metadata elements are adequately described in the document "OCLC/RLG Metadata Framework to Support the Preservation of Digital Objects"(PREMIS 2003) and many are involved in implementing the recommendations in the Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group (PREMIS 2005). APSR has identified the specification of a preservation metadata schema for the various repositories as being a critical outcome of the project.

Preservation metadata is not a technical solution, but rather a mechanism by which future solutions and procedures can be implemented. Any action intended to sustain digital objects must be informed by appropriate and relevant data about the objects. There will undoubtedly be new technical solutions to sustainability issues and preservation metadata is designed to take advantage of such solutions. In essence, preservation metadata is necessary to sustain access to digital information.

Preservation metadata invariably becomes a responsibility of the repository. This is because the information is mostly harvested at ingest, or is generated as a result of particular preservation action through time. Though a small amount of data may need to be entered, mostly data can be managed in an automated way. This requires that the repositories and tools reflect this automated practice. DSpace, however, suggests that responsibility for access belongs with the user community (DSpace, Michael J. Bass et al. 2002). However, as

the technical environment changes, so will the expertise of the user community, leading to a greater responsibility on the part of the repository managers for maintaining access.

The OAIS model suggests that there is a responsibility on the archive to deliver an archival information package that contains all that is "needed to make the content data object understandable to the designated community." (OAIS 2002, page 2-5). This requires that the repository managers "must understand the Knowledge Base of its Designated Community to understand the minimum Representation Information that must be maintained." (OAIS 2002, page 2-4). It may be quite reasonable to assume that a designated community will have particular expertise and technology in providing access to digital objects encoded in a specific way, especially for specialist users; however, more general users of digital materials may not have particular format expertise nor the need to maintain it. It is also very likely that the knowledge base of any designated community will change with time, so that the subject expertise may remain but the technical expertise associated with particular digital delivery technologies may change. It will be necessary to determine in which cases the ongoing support of access is the responsibility of the repository or that of the designated community. An agreed level of sustainable access by the repositories is required as guidance for the designated community of users (Bradley 2005).

The critical issue of responsibility will be played out in the digital arena as time progresses. Certain types of data with community based formats and standards, such as seismic data, or high energy physics data, has a technological community who are equipped to manage the data and formats of their particular area of expertise. Any actions they take should, of course, be informed by the digital archiving community, which has, and is, determining appropriate action and procedures. Other, possibly humanities based, communities (though many have developed expertise in the use of particular systems), do not necessarily have the ability to manage the issues associated with ongoing sustainability and preservation of digital objects, nor with managing what will have become, historic formats and technologies. Clearly, this sort of scenario creates a dilemma with regard to who should sustain access and who should undertake preservation action at the appropriate time. Some of these dilemmas can be explained in the models that might explain the relationship between users, rights owners and repositories.

The organisational structure of digital sustainability

The university environment contains many different organisation structures and relationships with the management of digital repositories. The different ways that this issue is managed may, in some circumstances, lead to a situation where no individual or institution is motivated to take responsibility for undertaking the preservation actions necessary for maintaining the digital content. A simple example of this may be where a general repository may not wish to expend resources on maintaining access to some particular content in its care, and the designated community no longer exists or no longer has the technical capability to maintain access, but the data retains some sort of historical or other research value.

Lavoie (Lavoie 2003) makes the point that the formation of repositories and associated sustainability activities will produce limited types of organisational structures to support the process. The type of structure is dependent, according to Lavoie's model, on three key roles; the rights holder, who holds the intellectual property rights to the digital materials; the archive, which is the area responsible for providing services to ensure the long-term preservation and accessibility of digital materials; and the beneficiary or beneficiaries, those

that benefit from the long-term retention of digital materials. The relationship of the entities, or roles, within that structure tends to lead towards particular types of natural incentives to undertake sustainability or preservation action, and precludes those incentives, or introduces disincentives in others. Lavoie argues that by understanding the economic models we are in a position to encourage preservation action by applying an incentive where such incentives do not naturally occur.

The obvious assumption behind Lavoie's work is that the motivation for the creation and use of a repository transpires due to a particular business need, and the continued sustainability of that repository is dependent, at least in part, on there being a continuing business need. The important conclusion is that, though there may be a business need, and even a cultural or intellectual imperative to sustain a digital repository, there may not be sufficient incentive for any individual entity to undertake the tasks that ensure sustainability due to the nature of the relationship between roles and the consequent organisational model. In such cases it is economically necessary to apply corrective economic policy measures and provide appropriate incentives (Bradley 2005).

A practical comparison might be made between a repository which holds community specific scientific data, and one which supports a much broader collection of general materials.

The Research School of Earth Sciences at the ANU stores seismic data generated by remote sensors. The data has been collected for many decades and is used by researchers at the school and by associated members of the community. The file formats are defined by the community, and are in fact be coded by them. They are virtually meaningless to anyone outside of the community. The Research School of Earth Sciences maintains its data systems, and has successfully managed data migration over a number of format and technology changes.

The right to preserve the data, the benefits from preserving that data, and the means of preserving the data are all located in the same entity. The core business for that department depends on sustainable access to the data and so it is clear that the decision to take any particular archival action will be informed and motivated by the best interest of the collections. Also, as the formats are so specific to the community, there are no other bodies to undertake the tasks, and so the community has direct responsibility for its own technology, and is the source of the most knowledge about its technology.

The converse of this is where all the roles and responsibilities are distributed in separate entities. An example might be a centralised, library-based repository, which holds published materials for an individual academic in a humanities research department. The data is in a standard proprietary format, such as Microsoft Word. The rights in the publication are probably with the individual, but maybe also with the University, and maybe the journal which published the article. The beneficiary may be a student of the Faculty of Arts, for example. The roles all belong to separate entities, and so there is no clear line of responsibility between those who might benefit from the preservation of the resource, those who have the right to initiate it and those who have the means to do it. Further, the fact that the file format is a broadly used one encourages "economic freeloading"; that is, no one takes responsibility in the hope that someone else will solve the problem.

In fact, recognising that sustainable access to Microsoft Word may be a preservation issue, the APSR has retained a programmer to investigate rendering all such documents into XML

during ingest into the digital repository. This technical solution makes it possible to re-present the original document in alternate formats on other, and future, operating platforms. However, this effectively makes responsibility for access the responsibility of the repository, rather than the user community, which is the reverse of the expected policy.

The APSR is, as part of its core projects, attempting to quantify and recognise the issues associated with organisational structure.

The economics of sustainability

One of the maxims that accompany the formation of digital collections is that they will not survive the type of benign neglect that is appropriate for the preservation of paper-based materials. In other words, digital materials require persistent attention, and no one-off action will sustain them for extended periods of time. There are some who claim that through such approaches as digital commons, data will survive with minimal back up strategies. However, the cost of providing meaningful access to the content through the use of developing digital archaeology skills will be cripplingly high, and not necessarily successful. Though Henry Gladney (Gladney 2004) consequently defines digital preservation action as an economic issue, it is one where he sees immediate action as necessary and accordingly advocates investment in the present to ensure access in the future. Lavoie notes that with the advent of digital authoring and distribution technologies, our developing capability to manage and sustain such information is being outstripped by our ability to produce it. He argues that, along with the necessary technological infrastructure for sustainability "must come the development of the associated *economic infrastructure*" (Lavoie 2004).

Understood in these terms, digital preservation is as much an economic issue as a technical one. The requirements of ongoing sustainability demand at their base a source of reliable funding, necessary to ensure that the constant, albeit potentially low level, support for the sustainability of the digital content and its supporting repositories, technologies and systems can be maintained for as long as it is required. Such constant funding is not at all typical of the university based communities that build these digital collections, many of which tend to be grant funded on an episodic basis. APSR recognises the need to develop costing models for sustainability of digital materials according to the specific requirements of the various classes of content, access and sustainability.

Conclusion

Sustainable digital repositories and preservation plans are about content. Access to valuable information is what drives all the plans and strategies associated with digital preservation. The only real measure of success of a sustainable digital repository, and its concomitant digital preservation plan, is access by the user community to valid and appropriate content. As discussed above, the relationship between data, the user community and a sustainable repository does not always result in clear and unambiguous roles for each; nonetheless, the partnership is a necessity. The value of the data, and the decisions made to sustain and preserve those items must be made not only with technical knowledge and expertise, but also with content, value and usage knowledge. As any sustainable digital access approach will require ongoing support from the institutions that house them, the collection they house must be demonstrably useful. This latter statement does not necessarily mean discarding unused material; libraries and archives have long recognised the need for extensive collections to

meet the broad needs of users, and the fact that the perceived relevance of collections changes in the face of research needs.

Repository software systems are not in themselves preservation solutions; rather they are tools in the socio-technical infrastructure that will make access to digital information sustainable. The socio-technical imperative of digital repositories and preservation applies in all aspects of its implementation. Repositories are the product of the interaction of people and technology, and as such, their sustainability is dependent on the continuation of that interaction.

References

Bishop, A. P., N. A. V. House and B. P. Buttenfield (2003). Introduction: digital libraries as sociotechnical systems. Digital library use: social practice in design and evaluation. A. P. Bishop, N. A. V. House and B. P. Buttenfield, MIT Press**:** 1-21.

Bradley, K. (2005). APSR Sustainability Issues Discussion Paper. http://www.apsr.edu.au/documents/APSR_Sustainability_Issues_Paper.pdf. Accessed September 2005

Cedars (2002). Cedars Guide to: Digital Preservation Strategies. http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html Accessed January 2005

DSpace, Michael J. Bass, David Stuve, R. Tansley, Margret Branschofsky, Peter Breton 1, P. Carmichael, Bill Cattey, Dan Chudnov and J. Ng (2002). DSpace-A Sustainable Solution for Institutional Digital Asset Services Spanning the Information Asset Value Chain: Ingest, Manage, Preserve, Disseminate. Internal Reference Specification, Technology & Architecture Version 2002-03-01. Masssachusetts Institute of Technology, Hewlett-Packard Company. http://libraries.mit.edu/dspace-mit/technology/architecture.pdf Accessed January 2005

Gladney, H. M. (2004). "Persepctives on Trustworthy Information." Digital Document Quarterly 3(4 (4Q2004)).http://home.pacbell.net/hgladney/ddq_3_4.htm Accessed 5 Jan 2005

Johnson, R. (2002). "Institutional repositories: partnering with faculty to enhance scholarly communication." D-Lib Magazine 8(11).http://www.dlib.org/dlib/november02/johnson/11johnson.html Accessed

Lavoie, B. F. Role in relation to data or project Name of data or project Interviewed by at place of interview January 2005.

OAIS (2002). Reference Model for an Open Archival Information System (OAIS). City, Management Council of the Consultative Committee for Space Data Systems (CCSDS). http://www.ccsds.org/documents/650x0b1.pdf Accessed January 2005

PREMIS (2003). Preservation Metadata for Digital Materials. PREservation Metadata Implementation Strategies, OCLC. http://www.oclc.org/research/projects/pmwg/background.htm Accessed November 2004

PREMIS (2005). Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group. PREservation Metadata Implementation Strategies, OCLC. http://www.oclc.org/research/projects/pmwg/premis-final.pdf Accessed September 2005

Raftos, P. (2005). The Australian National University: DSpace Testbed Program. http://www.apsr.edu.au/documents/DSpace_testbed_program.pdf Accessed September 2005

Smith, A. (2003). New Model Scholarship: How Will it Survive? Washington D.C, Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub114/contents.html Accessed February 2005

Weaver, B. (2005). eSpace development at the University of Queensland Library. http://www.apsr.edu.au/documents/eSpace%20development%20at%20the%20University%20of%20Queensland.pdf Accessed September 2005

Webb, C. (2003). UNESCO Guidelines for the Preservation of Digital Heritage. City, UNESCO. http://unesdoc.unesco.org/images/0013/001300/130071e.pdf Accessed 29 November 2004

Wheatley, P. (2004). Institutional Repositories in the context of Digital Preservation. Digital Preservation Coalition. Accessed 19 January 2005