Institutional Repositories

On institutional repositories, how they came to be, and how they are fitting into the digital library

by Richard Jones

Introduction by way of a brief history

The first seeds of the institutional repository can be traced back as far as the seminal articles by William Gardner and Stevan Harnad in 1990, when networked electronic communication was starting to become a viable tool for the dissemination of scholarly literature. In his article "Scholarly Skywriting and the prepublications continuum of scientific enqury", Harnad states that:

> "The whole process of scholarly communication is currently undergoing a revolution comparable to the one occasioned by the invention of printing." (Harnad, 1990)

These early moves towards digital repositories which really took off some ten years later were primarily suggestive of disciplinary archives, borne out by the establishment in 1991 of the Los Alamos arXiv[1] for high-energy physics (now based at Cornell). The first scholarly recorded proposal for an institutional archive came later, in 1994, in response to Harnad's "Subversive Proposal for Electronic Publishing" (Okerson and O'Donnell, 1995), in which he states:

> "For centuries, it was only out of reluctant necessity that authors of esoteric publications entered into the Faustian bargain of allowing a price-tag to be erected as a barrier between their work and its (tiny) intended readership, for that was the only way they could make their work public at all during the age when paper publication ... was their only option."

The subsequent discussion which took place by email on various lists (particularly Virginia Polytechnic Institute's VPIEJ-L list and the University of Vermont's SERIALST list) among interested scholars, including Paul Ginsparg, the originator of arXiv, brought to light many considerations. These included the technical requirements at the time, centralised versus decentralised storage models, and the formulation of the copyright issues that would play a dominant role in the self-archiving debate. Here, Nobel Prize winner Joshua Lederberg introduced into the discussion the idea of institutional rather than disciplinary archives:

---

[1] www.arxiv.org/

> "...instead let each institution set up its own ftp-able archives for all of its scholars. That way, each place can also set up its own ground rules."[2]

It is notable that this discussion focused around setting up FTP archives for scholars to use with immediate effect. The institutional repository in its current form is a centralised service, often run by the institution's library and the applications themselves require internet technologies to be well developed; therefore, the concept could not become as it is today without the time for the underlying technology to mature. The discussion and movement was also primarily driven by the scholars themselves rather than librarians; this is not surprising, since the practice of circulating preprints of articles between academics has long been commonplace. Again, though, the skills required to operate the modern institutional repository have long resided in the library, and their involvement is virtually necessary.

It is perhaps for these reasons that a large corpus of literature on the subject did not start to develop until around the turn of the millennium. Between 2001 and 2003 there was an explosion of articles covering the groundwork for institutional repositories, most notably "The Case for Institutional Repositories" (Crow, 2002) which came from the Scholarly Publishing and Academic Resources Coalition (SPARC), set up by the Association of Research Libraries (ARL) in 1998 to address the issue of high scientific journal prices.

In addition to this, software to support the creation of e-print archives really started to become available in 2001 with the release of EPrints.org[3], and later in 2002 the release of DSpace[4]; between them the most dominant open source repository packages. It is worth noting that by this point the idea of archiving e-theses institutionally was already well under development, and ETD-db[5], the major software package in this field, was released in 1999 by Virginia Tech and the Networked Digital Library of Theses and Dissertations (NDLTD). The role of the e-theses efforts should not be underestimated in the development of the institutional repository, since they provide the infrastructure to gather some institutional research under a degree of control by the organisation which is not so straightforward in practicing academic research.

We can suggest, therefore, that the origins of the institutional repository are fairly complex insofar as the modern interpretation of the term. Dominant factors would include:
- Pre-existing or under-development e-theses archives;
- Pre-existing departmental e-print archives

---

[2] The full discussion is presented abridged in Okerson and O'Donnell, 1995
[3] www.eprints.org/
[4] www.dspace.org/
[5] scholar.lib.vt.edu/ETD-db/

- Grass-roots faculty practices of making e-prints available on personal web pages (Johnson, 2002);
- Subject repositories such as arXiv;
- Institutional desire to preserve for both posterity and portfolio;
- Support for the philosophy of Open Access;
- Pre-existing distributed document servers;
- A reaction to the "Journals Crisis".

Nearly simultaneous to these factors coming together to produce the first fully featured institutional repositories, the development of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[6] made the desire for interoperability and cross-searching disparate repositories a much more viable goal. First convened in 1999, the OAI group produced the first stable version of the protocol in 2001. By utilising simple Dublin Core[7] at the most basic level, and by providing a simple URL based query schema, this protocol dramatically lowered the implementation barriers for interoperable archives.

A critical factor, though, underpinning much of the development of all forms of Open Access repository has been the so-called "Journals Crisis". ARL statistics show that serials costs between 1986 and 2004 increased in price by approximately 273%, while the Consumer Price Index (CPI) increased, approximately in line with library budgets, by around 73%[8]. This major disparity in the rate of increase of cost has caused numerous budgetary crises in libraries struggling to balance the need to maintain journal access with the rapidly diminishing budget available for other essentials, such as books. This problem has been further complicated by the switch to e-journals by publishers and the introduction of the "Big Deal". This works by offering electronic journals on a multi-title, multi-annual basis, and has a relationship to the institution's print cost such that the lower the print spend, the higher (relatively) the electronic spend. The outcome is that switching to electronic delivery only is not necessarily cheaper.

This problem is particularly relevant for institutional repositories, since the cost bearers are the institutions themselves. Therefore, while disciplinary repositories came first, institutional repositories can be seen as a way that an institution can directly address the issues themselves, with the long-term goal of reducing the e-journal costs, or even altering the publishing model entirely. The effort required, though, is not to be underestimated, and the stakes are high for the main participants (faculty, librarians and publishers), while overcoming the traditional publishing paradigm is a major challenge (Johnson, 2002).

[6] www.openarchives.org/
[7] A basic 15 element metadata set, often thought of as the 'lowest common denominaor' for metadata. See hwww.dublincore.org/
[8] www.arl.org/stats/arlstat/graphs/2004/aexp04_pub.pdf

The remainder of this chapter discusses how institutional repositories have been defined, how they are compared and contrasted with other similar kinds of repositories, and what kinds of uses they have found within their host institutions. Since the institutional repository, although old enough in concept, is still young in implementation, we will look at some examples of how they are being configured and presented, and the place that they have found themselves within the Digital Library. It is also of interest for us to delve briefly into the more technical details and concepts that underpin the repository, including common and desirable features, digital preservation activities (and especially the Open Archival Information System reference model), and information management challenges that are presented. We will also look at the social features of the repository such as how it is viewed in relation to intellectual property rights, and the challenges that libraries face in producing advocacy strategies to populate their repositories.

The many faces of the institutional repository

Institutional repositories tend to have a very wide remit. They mean many different things to many different people, and are used in a variety of ways. The sorts of content types, for example, that we see include e-prints (both pre- and post-prints), gray literature (especially e-theses), working papers, technical reports, books and book chapters, conference papers and posters, and even some administrative records. Some broad working definitions have been drafted which try to encompass the functions of this repository type. For example, Clifford Lynch, the director of CNI[9] defines it as:

> "...a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organizational and access or distribution" (Lynch, 2003)

While Lynch defines the institutional repository as primarily service orientated, and by commitment from the institution, Raym Crow of SPARC defines it from a slightly different, but compatible, angle in terms of the origins of the material as:

> "any collection of digital material hosted, owned or controlled, or disseminated by a college or university, irrespective of purpose or provenance" (Crow, 2002)

---

[9] www.cni.org/

Crow goes on to note that the college or university boundary is not absolute, and that organisations finding use and benefit from an institutional repository could include government departments, non-governmental organisations (NGOs), museums, independent research organisations, federations of societies and even commercial entities.

Further to looking for broad definitions of the repository, there have also been efforts to define some general properties that either define the institutional repository, or are the natural outcome of maintaining one. From the literature available it is possible to define six characteristics without too much difficulty (Crow, 2002; Johnson, 2002; Lynch 2003; Genoni, 2004) :

- Institutionally defined;
- Scholarly;
- Cumulative and perpetual;
- Open and interoperable;
- Capturing and preserving events of campus life;
- Searchable within constraints.

We will not discuss the justifications for these characteristics in depth, but some notes on the second and third points are warranted. That a repository be *Cumulative and Perpetual* suggests that it continues to gather materials, and continues to store those materials (under some selective preservation strategy) in perpetuity. That is, the archive is not static and it is not short-term. Meanwhile, being *Open and Interoperable* is the only way that an institutional repository can reasonably interact with other repositories, and placing access barriers of any kind will no doubt lead to a lack of use globally, and ultimately obsolescence (which would make being *perpetual* considerably more challenging). There are further discussions on these characteristics in "The Case for Institutional Repositories" (Crow, 2002) and "The Institutional Repository" (Jones et al, 2006:53-54).

We could try to contrast these properties with other library information systems such as subject repositories, learning object repositories, institutional record archives, library catalogues and metadata aggregators. The boundaries between all these are extremely blurred, and we find, for example, institutional repositories containing learning objects or institutional records. In addition, the term "learning object" itself has a fuzzy definition, and could encompass materials such as journal articles. Areas which make the institutional repository what it is, then, appear to be its institutional branding for the end-user, its further focus on materials that reflect the nature of institution members irrespective of their use in the outside world, and perhaps the idea that it is related to preservation more so than other repository types (in that it has a highly interested curating entity).

What we generally find is that most modern institutional repositories are primarily e-print and e-thesis archives, most likely because of the development history and origins of the repository. Other content types such as multimedia, course materials and datasets are emerging, but they are yet to be considered commonplace.

When choosing how to organise institutional repository holdings, which we expect to reflect the fact that it is institutionally defined, we find a number of things. The obvious pattern to follow is that of the institution's internal organisation, and this structure is common. We often find that at the lower levels of the categorisation that a content type (e.g. e-theses, e-prints) structure is used to further refine the collections. Some repositories are actually finding themselves cross-institutional, either because of common research goals in small organisations, convenience in terms of technical support or geographical proximity and branding.

Other collection structures that are found include those which rely purely on categorisation by content use type (e.g. technical documentation, learning objects), while some use a very pure subject heading structure. We also see that many repositories exhibit inconsistent structures, which can be ascribed to a number of causes: first, that devolved administration produces structures in one area of the repository that are not informed by those in other areas; second, that some structures may be more appropriate to different disciplines or organisational units; and third that due to the youth of many of the systems their place is uncertain in their institution and information environment.

The institutional repository in the digital library

If the institutional repository does not yet inhabit a defined place in the information environment, then they are not sufficiently well established to even be considered essential elements. It may be, for example, that Open Access Journals and disciplinary repositories will prove the most effective and popular in the long term. Nonetheless, they are creating for themselves a place in the Digital Library, itself a relatively new entity. Aside from the traditional library Online Public Access Catalogue (OPAC), other components of this environment include the increasing prevalence of portals, both library, institutional and even disciplinary. Electronic journals are a staple part of the digital library environment, as well as e-books and internal services such as departmental web pages containing course and reading lists, for example. Learning object repositories are also making appearances, and there is always the wealth of general information literature that has been provided by libraries since before the electronic age containing help and tutorials for users. The rest of this book introduces many of these facets, and we can see how rich the information environment becomes.

These content types exhibit large amounts of cross-over. For example, materials held in institutional repositories could be catalogued in the OPAC, while it is often necessary to surface resources in multiple portals, each of which may have differing ideas as to implementation; repositories will also often contain material that is available in some form in an e-journal, and traditional library resources could be seen as relevant material for storage in the repository. The outcome of this is that the repository must be able to interoperate: exporting records for OPACs, providing web service interfaces for portals, and being flexible enough in metadata capture to hold a variety of unusual or unpredictable objects. Many regard the institutional repository as one kind of storage mechanism among a sea of other systems appropriate for their use type (for example, the advanced management and presentation of image collections may be better done elsewhere). That is, the institutional repository is not the complete solution for institutional digital asset management, but it may rapidly become an important part. Creating well organised networks of information will the the ultimate goal, and the repository will be one of the participants.

In addition to this, the global information network of different types of repositories is increasingly gathering more nodes. Repositories should, as Crow notes, be open and interoperable, and using technology such as OAI-PMH we are seeing the rise of repository information networks as theoretically depicted
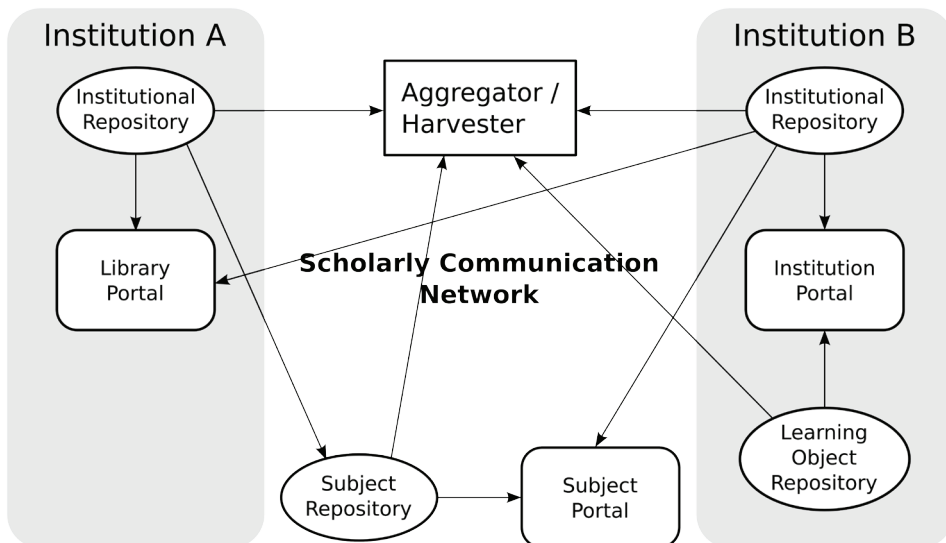

Figure 1. A theoretical repository, portal and harvester network for scholarly communication

in *Figure 1*. Here, the different information sources, ranging from institutional and disciplinary repositories through to databases of aggregated metadata from all types of repository are interacting. The goals of this interaction include making available information in as many places as possible and providing users with the opportunity to access materials in a number of different ways, and ideally in a full-text form. Cross searching, and tools such as OpenURL make tracking down the desired information that much easier, and if one route is not necessarily open to a user, another may be.

Technicalities and details

It is useful, in order that we understand what the implementation of an institutional repository really brings with it, for us to examine in more detail the sorts of features and requirements that might be appropriate, as well as the complicating social issues that arise. We can start with an idealised approach to the general features that would be appropriate for a repository of this nature. Many of these systems will exhibit some of these features, although it is a list that may ultimately be attained rather than that already exists. It should also be borne in mind that institutional requirements for the repository will vary, since we have noted that it is not a well-defined system. The following, nontheless, outlines are fairly fully functional instititutional repository (Jones et al, 2006:55-62):

- *A strong development community*; particularly important in open source software, a community behind the package driving the repository is always an asset. This is an aid to technical support, further development and so forth.
- *Easily integratable*; since the repository will be one part in an already established electronic landscape it is useful to have a system which can easily embed itself both technologically and culturally.
- *Security*; some particular uses of the repository will require direct interaction with particular users, and the facilities to authenticate these users and authorise their activities is necessary. In line with integration, a system which can talk to established institutional security mechanisms is ideal.
- *Archival integrity*; to maintain holdings with some certainty over time it is necessary to ensure that content is not tampered with, and that the audit trail and provenance of the item are kept in good order.
- *Administrative tools*; repositories typically require tending by trained administrators. Good tools to aid this administration become very important as the popularity of repositories increases.
- *Licensing and Restrictions*; while we aim to capture as much material as

possible, some may still have legal barriers to open access. The repository, in these cases, should try to hold the material under a relevant licence and restrict public access until such time as it can be released as opposed to rejecting the content.

- *Web service interfaces*; in particular, support for the OAI-PMH is very important, but other protocols such as SRW/U and even z39.50 could have a place in the repository.
- *Metadata management*; it is necessary for both exposure to end users and to web services that metadata is in good shape and appropriate. Systems should gather adequate metadata both for description and access as well as ideally more technical metadata to aid preservation.
- *Federation, devolution and scalability*; as repositories grow both in holding size and usage it will be ultimately necessary to devolve features such as administration and storage in order to achieve the necessary scalability.
- *Ingest and egress routes*; it is of paramount importance to ensure that materials can both get in and out of the repository with relative ease. All barriers to ingest and egress must be kept as low as possible to make adoption more attractive, and interoperability easier.
- *Preservation activities*; the opportunity is available for institutions who wish to preserve their intellectual outputs to do so within the remit of the institutional repository. It can provide both the software tools and the cultural change which will gather output from all locations and make them available as targets for digital preservation (Wheatley, 2004).

We also see repositories being imbued with other characteristics, whose value-added services could be the gateway to really embedding them within institutional working practices. These include tools such as researcher home pages, which can act as a full-text CV for academics. Full-text searching is common, and we are finding increasing use of true subject classifications, making the institutional repository more and more a core service. End user features such as annotation and discussion forums for items have appeared, and the option to have the item printed by the institution is being offered in some cases. The CDSware[10] team at CERN in Geneva are offering on-the-fly file format conversions to aid in digital preservation, as well as giving the option to web-cast streaming content. The repository then becomes just one facet of an increasingly sophisticated information and research tool.

The major technical challenges for the repository are really shared by many forms of modern archival systems. With the advent of grid technologies, using federated storage is becoming important both to store the digital content

---

[10] cdsware.cern.ch/

as it grows in size and to aid in preservation. Meanwhile, end user tools are constantly in development to aid cross searching of multiple databases from various institutional and subject portals, which has major challenges in terms of user interface development and interoperability standards. OAI-PMH use has become widespread, but this brings with it many information management challenges which have yet to be satisfactorily addressed; these include record deduplication (OAI, 2005) and metadata enhancement (for an example, see McClelland, 2002).

With digital artifacts as young as ten years we are already seeing a high degree of obsolescence both of format and storage media, so digital preservation is one of the big topics being addressed in the library, information and computing communities. This area has no guaranteed solutions (and may never have any), although there is a lot of development and recommendations in areas such as content migration, emulation of software environments, or even development of generic platforms upon which digital preservation activities can take place (see Wheatley, 2004 for an excellent introduction to digital preservation in the context of institutional repositories). In 2002 the OAIS Reference Model ISO standard (CCSDS, 2002) emerged as a recommendation for how an archival system can aim to support preservation activities. The basic workflow of the system is presented in *Figure 2*.
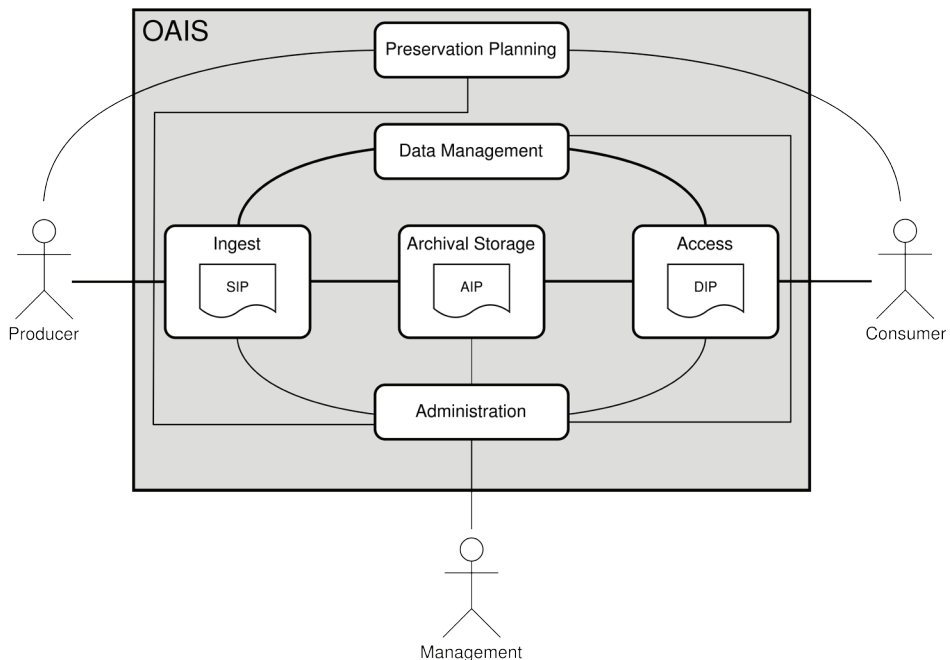

Figure 2. OAIS reference model overview (Jones et al, 2006:78)

Figure 2 shows three main participants in the archival procedure: the producer, the consumer, and the manager. The producer generates a Submission Information Packet (SIP) containing all the relevant metadata and content for the item. This is then converted by the system into an Archival Information Packet (AIP), which is the preservation copy. The consumer accesses a Dissemination Information Packet (DIP) through the access area of the system, which determines which AIP to convert and deliver, based on the working data in the Data Management section. The OAIS system knows how to transform data between these forms when necessary. Across all of this, then, there is administration as handled by the manager, and preservation planning spanning all three participants. Detailed discussions on this OAIS reference model can be found in CCSDS, 2002 and Lavoie, 2004.

Some of the most critical details and technicalities of the institutional repository, particularly those that act as barriers to adoption, are social, though, not technical. Problems that were proposed early on, before many academics had an opportunity to come across the institutional repository or open access were that there may be a quality control element somehow missing from the open access repository model, and that the entrenchment of the paper publishing model, due to its monopoly, would be difficult to break.

Already Open Access Journals are addressing the issue of quality control by applying equally valid peer review. They are also trying an experiment in the publishing business model that will not necessarily mean that traditional publishers will have to fight the open access movement, but instead modify their working practices to take advantage of it. The role the institutional repository plays in this process is threefold: as part of a global effort to improve open access, they act as a catalyst for change that is directly accessible to the researcher; they can hold and disseminate materials that become available either through open access publishing or from publishers who are amenable to the self-archiving process; they smooth the path to open access by giving a forum for academics who are without an open access journal to appropriately publish their work.

Concerns that repository implementers must be prepared to deal with that stem from this slowly changing model are primarily centered around copyright and intellectual property rights (IPR). This is a clear hang-over from the entrenched publishing model, and it takes time to demonstrate that handing over copyright to a publisher is not an absolute requirement, and that the availability of materials in an open archive does not preclude immediate copyright or IPR infringement by unknown parties. The real crux of the problem is in the unclear position that repositories of all types hold within the publishing environment; contrary to traditional services, libraries are finding themselves increasingly in the role of a publisher which brings with it many new responsibilities (Jones et al, 2006:145-

147). Much material that academics would like to deposit is already copyright by the original publisher, while other material (especially theses) could be safely published online now, but damage the possibility of later publication of derived works because of prior publication. This paralysis can be worked at by chipping away at the copyright problems by asking publisher permissions to place materials online, addition by the author of friendly clauses to publisher contracts, and by encouraging authors to publish in open access journals wherever possible.

There has been much work in this area since repositories started to become reality. The Securing a Hybrid Environment for Research Preservation and Access (SHERPA)[11] project in the UK now maintains a list of publisher policies[12] with regard to self-archiving practices, and a similar service is provided by the EPrints. org community[13]. This allows academics to rapidly ascertain the default copyright status of their work for deposit in an institutional repository.

Repositories and their managers must mitigate their risks carefully in this field. There are large administrative overheads in ensuring that all holdings are copyright safe. While checks must be made by any responsible organisation, sometimes it is impossible to know whether an item is truly clear for deposit. The only person really in a position to know this is the author, so repositories should be sure to licence works appropriately as they are deposited. The main stakeholders in the licensing process are the author, the institution and the end-user; a licence for any deposited material should include a statement by the author that the material is theirs to deposit, preservation rights for the institution (transform and migrate and so forth), as well as reuse/distribution rights that the author can agree to in order to allow reasonable open access to their work.

One study by the Zwolle group[14] has attempted to identify all the major stakeholders in the creation and dissemination of scholarly materials, and investigate and promote balanced approaches to rights management. They hope that this is ...

> "...a crucial step toward the development of policies or agreements that seek to assure to the stakeholders the ability to use and manage the works in fulfilment of their most important interests". (Zwolle Group, 2001)

Further considerations in this regard by repository managers will include local factors such as Freedom of Information (FOI) legislation, as well as the necessity in some cases to implement access restrictions on holdings.

When attempting to sell the repository to faculty it is necessary to find

---

[11] www.sherpa.ac.uk/
[12] www.sherpa.ac.uk/romeo.php
[13] romeo.eprints.org/
[14] www.surf.nl/copyright/

arguments to satisfy concerns about these legal situations. For example, a common misconception is that publishing materials in an institutional repository may increase the risk of plagiarism, and while in reality the chances appear small, and the likelihood of a plagiarist being detected are greatly increased. Repositories could also provide automated plagiarism checks as a value-added feature. For detailed information on the legal considerations for electronic resources see Oppenheim, 1999, and for further discussion of advocacy strategies see Jones et al 2006 (pp 111-138).

Case Study: Bergen Open Research Archive (BORA)
The Bergen Open Research Archive (BORA) is the institutional repository of the University of Bergen. It was first launched in late 2004, as one of the first institutional repositories based on DSpace in Scandinavia. The administrative and technical base of this repository lies in the library's acquisitions department, where the skills concerning management of other e-resources such as e-journals, e-books and the library catalogue reside.

BORA is intended to work alongside other repositories managed by the library (ones which are not necessarily open to the public), such that as a whole they meet Lynch's criteria of offering services for management and dissemination of institutional digital assets. BORA then provides the public face of the university's archived research, containing research articles, working papers, books (such as this) and some masters and doctoral theses. It also broadly meets the six general criteria of an institutional repository:

- Institutionally defined: it is branded by the university and the collections are structured according to the institution's organisation.
- Scholarly: there is a wide range of purely scholarly material held therein.
- Cumulative and perpetual: it is constantly and increasingly gathering new materials, and although copyright issues have caused items to be removed in the past, most items remain and are persistently identifiable.
- Open and interoperable: it is compliant with the OAI-PMH 2.0 protocol, and provides data to the Norwegian Open Research Archives (NORA)[15] aggregator.
- Capturing and preserving events of campus life: it will accept most documents that are of relevance to the university, although less obvious uses have yet to be thoroughly investigated.
- Searchable within constraints: it offers a full-text search interface as well as browsing by some metadata elements including a controlled vocabulary of terms.

[15] www.ub.uio.no/nora/

As time goes on BORA finds itself involved in more interactions with the existing digital library environment, including providing and being a target of SFX services, being cross-searchable from the library portal, being a data-provider for several OAI harvesters, and interacting with research reporting systems. The BORA brand has also grown in the Bergen area such that it has become an umbrella term for several institutional repositories sharing a common interface, of which the University of Bergen is now only one facet.

Conclusions

Current adoption levels of institutional repositories are pleasingly high. There are two major registries of open access archives: OpenDOAR[16] (Directory of Open Access Repositories), and ROAR[17] (Registry of Open Access Repositories). These registries show that adoption in developed countries is already very high, with some developing countries also starting to take the advantages that open access confers. The area is in rapid development, and ROAR statistics show a steep incline from around 2001 to the present in the availability of both open archives and open records in those archives available via OAI[18]. By simultaneous reference to the increase in literature which follows a similar timeline, we can surmise that we will continue to see further growth in this area in the short term at least.

A successful institutional repository requires institutional commitment; startup resources are relatively high, and there is a necessity for both technical and advocacy skills to be employed. To make the repository part of every day working practice of the academic is no small task, and tireless commitment to marketing and improving the service, as well as reacting to feedback from the users is necessary.

The true maturity of the institutional repository is some way off, and it will not reach it until the many copyright and working practice obstacles have been successfully tackled. Technologically, the existance of the institutional repository is straightforward, although we will continue to see new technologies, especially in the field of federated storage and devolvable interfaces, being incorporated into repository software.

This chapter has attempted to expose the institutional repository in the place that it has started to find for itself within the digital library. This landscape is constantly shifting as new technologies come and go, and the repository, like many of its information system peers, will have to adapt to changes rapidly, whilst

---

[16] www.opendoar.org/
[17] archives.eprints.org/
[18] These statistics only cover open archives registered with ROAR and which support and make available records using OAI-PMH. See http://archives.eprints.org/index.php?action=analysis

always keeping in focus its original goals and intentions. With the pervasiveness of the internet, and the large number of high quality public search services, repositories like this are going to need to form the cornerstone of the institution's commitment to bringing quality information into the public domain; a challenge which they should take to with great determination.

Bibliography

CCSDS. 2002. CCSDS 650.0-B-1: *Reference model for an open archival information system (OAIS)*. Blue Book 1. Available from: http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf [accessed: 9 Feb 2006]

Crow, R. 2002. *The Case for Institutional Repositories: A SPARC Position Paper* [online]. SPARC. Available from: http://www.arl.org/sparc/IR/ir.html [accessed 9 Feb 2006]

Gardner, W. 1990. *The Electronic Archive: Scientific Publishing for the 90s.* Psychological Science 1 (6): 333-341.

Genoni, P. 2004. *Content in Institutional Repositories*. Library Management 25 (6-7), pp300-6.

Harnad, S. 1990. *Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry*. Psychological Science 1 (6): 342-344.

Johnson, R. 2002. *Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication* [online]. D-Lib Magazine, 8(11). Available from: http://www.dlib.org/dlib/november02/johnson/11johnson.html [accessed 9 Feb 2006]

Jones, R., Andrew, T., MacColl, J. 2006. *The Institutional Repository*. Chandos Publishing, Oxford.

Lavoie, B. 2004. *The Open Archival Information System Reference Model: Introductory Guide* [online]. Digital Preservation Coalition. Available from: http://www.dpconline.org/docs/lavoie_OAIS.pdf [accessed 9 Feb 2006]

Lynch, C. 2003. *Institutional repositories: essential infrastructure for scholarship in the digital age*. ARL. Available from: http://www.arl.org/newsltr/226/ir.html [accessed 9 Feb 2006]

McClelland, M., McArthur, D., Giersch, S., Geisler, G. 2002. *Challenges for Service Providers When Importing Metadata in Digital Libraries* [online]. D-Lib Magazine, 8(4). Available from: http://www.dlib.org/dlib/april02/mcclelland/04mcclelland.html [accessed 9 Feb 2006]

OAI. 2005. *Implementation Guidelines for the Open Archives Initiative Protocol for Metadata Harvesting* [online]. Open Archives Initiative. Available from: http://www.openarchives.org/OAI/2.0/guidelines.htm [accessed 9 Feb 2006]

Okerson, A., O'Donnel, J. 1995. *Scholarly Journals at the Crossroads: A subversive Proposal for Electronic Publishing* [online]. ARL. Available from: http://www.arl.org/scomm/subversive/ [accessed 9 Feb 2006]

Oppenheim, C. 1999. *The Legal and Regulatory Environment for Electronic Information* 3rd Edition. Tetbury: infonortics

Wheatley, P. 2004. I*nstitutional Repositories in the context of Digital Preservation* [online]. Digital Preservation Coalition. Available from: http://www.dpconline.org/docs/DPCTWf4word.pdf [accessed 9 Feb 2006]

Zwolle Group. 2001. *Copyright Management for Scholarship* [online]. SURF. Available from: http://www.surf.nl/copyright/ [accessed 17 Feb 2006]