



Paper submitted to the Online Information Conference 2009,
TRACK 3: INFORMATION PROFESSIONALS DELIVERING & DEMONSTRATING VALUE
Session: Digitisation, repositories and preservation
Wednesday 2 December, 11.30-13.00

The Institutional Repository in 2010 ... and beyond

Introduction

In february 2003, Clifford A. Lynch identified repositories as essential infrastructure for scholarship in the Digital Age¹. In his paper, he announced the now widely adopted definition of a university based institutional repository 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 organization and access or distribution."*

Although supported by an important technological component, he also underlines that an institutional repository is not simply a fixed set of software and hardware.

This paper aims to provide insights and predictions with respect to the short and long term future of Institutional Repositories, based on the current appraisal and criticism on today's results of more than six years of Institutional Repository reality.

While acknowledging that important challenges still lie ahead, this paper states that the trend towards organizational commitment to the stewardship of digital materials, essentially at the core of institutional repositories, is still very actual today and will be in the future.

The Current Repository Landscape and Trends

Over the past six years, a large number of institutions has launched Institutional Repository related projects, and this number is still rising today.

The directory of Open Access Repositories² service provides a quality-assured listing of open access repositories around the world. OpenDOAR staff harvest and assign metadata to allow categorisation and analysis to assist the wider use and exploitation of repositories. Each of the repositories has been visited by OpenDOAR staff to ensure a high degree of quality and consistency in the information provided: OpenDOAR is maintained by SHERPA at the University of Nottingham.

With currently on average one new listing every day, OpenDOAR lists more than 1500 repositories.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



Aside from growing awareness and increased interest in organizational commitment to stewardship of digital materials, following influences also attributed to these growing numbers:

Cross-Institution efforts

In countries where government bodies, or consortial collaborations of institutions, embarked on institutional repository projects, directly affecting a number of institutions, we see a high uptake of institutional repository platforms today. We focus on two different examples, to illustrate that these kinds of projects can have very different angles and ambitions.

The **DARE** (Digital Academic Repositories) project in the Netherlands was initiated in January 2003 by the SURF foundation, a collaboration between education institutions that drive ICT innovations in the education. This project had the ambition to erect Open Access Repositories at each of the universities in the Netherlands, and attribute the role of a national harvester to the national library. This project concluded in 2006 and succeeded in establishing the institutional repositories and the national portal, currently online under the name Narcis³. DARE was followed by two tenders, resulting in project grants for 7 projects in 2008 and 12 projects in 2009. These are smaller projects focussed on themes including collaboration, enriched publications and dissemination of knowledge in the high school environment.⁴

In the UK, a number of cross-institution repository related initiatives are coordinated and funded by **JISC** (Joint Information Systems Committee). JISC's mission is to provide world-class leadership in the innovative use of Information and Communications Technology to support education and research. JISC is directly funded by the different Higher Education Funding Councils in the UK. With large scale calls and programmes, such as the Digital Repositories Programme 2005-7, the Digital Repositories Programme 2007-8 and more recently, the £14m Repositories and Preservation programme⁵, JISC has driven repository development forward nation wide. Under the umbrella of these programmes, services such as EPrints and OpenDOAR emerged. Aside from the goals to offer online access to digital content, other angles of added value were explored as well, like the potential for Institutional Repositories in the area of Research Assessment⁶

When comparing the DARE approach in the Netherlands, and JISC's programmes around 2005, JISC seemed to pursue a broader scope of research goals, while DARE's approach to get the network of institutional repositories established was more pragmatic. Currently, the umbrella of SURF Share projects shows more diversity in the Netherlands as well.

Current state of the Art in Institutional Repository Platforms

In his "*INFORMAL comparison of some institutional repository solutions*"⁷, Neil Godfrey gives an intro level guideline for institutions interested in "what (else) is out there" into different Institutional Repository platforms. Without going into the specific technical detail about which set of features is offered by which platform, following high level observations characterize the current state of the Art in these Institutional Repository Platforms.

- As demonstrated by the numbers in OpenDOAR, **Open Source Software** systems such as GNU EPrints, DSpace and Fedora Commons together seem to hold the lionshare of the landscape. However, as the registration in OpenDOAR

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



requires an open OAI-PMH interface for harvesting of the metadata, it's possible that other platforms, without this interface, are being used to implement an Institutional Repository vision without being registered there.

- These platforms are all web based applications in which different people can have distinct views on information and tasks, depending on the roles they are attributed.
- Typically, the technical installations of these repository platforms are centralized and managed internally at the institution. Alternatively, commercial service providers, such as BePress and Open Repository and non-commercial service providers such as NITLE and CILEA, offer repository platforms as hosted systems to their client institutions.
- While some systems, such as Eprints and DSpace are easy to install and maintain by a qualified IT Technician, platforms like Fedora require intensive customization effort prior to its use and deployment.
- Although predominantly used to create freely accessible online repositories, most platforms have functionality to restrict access in different ways. For example, the content of the Jim Henson Works repository⁸, implemented with the Fedora Commons platform by the University of Maryland, restricts access to the content to public computers at certain libraries. Likewise, repositories might only offer descriptive information about the content publicly, and restrict access to the actual full text or digital object by embargo's or other mechanisms.
- As illustrated by the DSpace introduction video, repositories are being implemented in a variety of organizations, not only in Academic Research institutions.

The Open Access Movement

The earliest examples of supporting online infrastructures for open access self-archiving ("green" OA) of scientific output, in the form of pre-print archives, were mainly subject specific, such as Arxiv.org. It was when institutions themselves, started to support green OA and think about deposit mandates, that Institutional Repositories gained field as enabling technology for implementing Open Access at the institution level. Lead by the University of Southampton's Open Access Archivist Stevan Harnad, the community of researchers, policy makers and librarians that effectively wanted to realize Green OA at their institution grew steadily. The commitment of a large number of institutions was formalized in the Berlin Declaration on Open Access Knowledge (2003).

Although the implementation of an Open Access vision can be an important argument for the initiation of an institutional repository project, this is very different than the more general idea of organizational commitment to the stewardship of digital material. The core difference here is that the aim of Open Access is providing faster access to (pre-print) journal publications, and furthermore, to make this a free service to the public. While the general definition of an institutional repository, and the commitment behind it, doesn't explicitly challenge certain aspects of traditional publishing, Open Access does.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



Perceived Success

Determining whether existing repository initiatives are currently successful is not a straightforward task. Each individual repository project can have different objectives or milestones, and those might not always be publicly available. Here are some of the common metrics and considerations.

According to the European Repository Landscape study by SURF and DRIVER started in June 2006 and completed in February 2007⁹, based on a large scale survey of academic institutions into the state of their Digital Repositories revealed that on average, 40% of the European academic institutions had a repository for their digital academic output in January 2007.

However, differences between nations are large: in 7 from the 27 surveyed countries, there appear to be no research institutes with a digital repository for research output. On the other side of the spectrum, 7 countries (including the Netherlands and the United Kingdom) were already in an advanced stage where more than 50% of the institutions were running a repository.

So in short, this study categorized certain countries as more advanced comparing to other ones, merely by the fact that more repositories are present, for a higher percentage of institutions. The study itself realizes that the glass is still half empty today, as the existing repositories cover only an estimated 38% of the research output of their institutes from 2005, while an estimated 37% of the academics in their institutes deposited their materials. This illustrates that the fact that institutions or countries are adopting repository platforms, doesn't necessarily means a success in terms of executing the commitment to stewardship of digital materials.

Another measurement for assessing progress or success of a repository project, is the growing or declining number of deposits. Through the OAI-PMH interface, organizations external to the institution are able to harvest the repositories contents and keep track of new deposits. OpenDOAR holds references to the number of deposited items for each of the listed repositories. The Network of European Economists Online (NEEO)¹⁰, an EU funded project measures both the number of deposited full text items as well as individual downloads of these full text items, for the repositories included in the network.

When only regarding the growth of the repository contents, as a measurement for success, the repository's ability to offer access, and serve community of visitors is not being assessed. By keeping into account repository item visits and individual file downloads, the ability and success in offering online access to the stored content can be measured.

Another alternative to measuring this ability, is to assess the repository's exposure in online search engines, i.e. online visibility.

The aim of the Ranking Web of World's Repositories¹¹ is to support Open Access initiatives and therefore the free access to scientific publications in an electronic form and to other academic material. The web indicators are used there to measure the global visibility and impact of the scientific repositories.

According to the Ranking Web of World's Repositories, online visibility of repositories keeps increasing, with more pages and higher average pageranks every six months.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



Likewise, the numbers in OpenDOAR illustrate that more and more content is being deposited in repositories.

Perceived Failure

Before going into detail about today's challenges for repositories, we evaluate recent studies into the causes why a specific repository project hasn't delivered according to what was expected.

In January 2005, in the article "*Understanding Faculty to Improve Content Recruitment for Institutional Repositories*"¹², authors and institutional repository managers Nancy Fried Foster and Susan Gibbons (university of Rochester) admitted that the phrase "If you build it, they will come" didn't work well as a content recruitment strategy for the repository project. She strikes a point still valid today, indicating that Institutional Repository benefits seem to be very persuasive to institutions, IRs fail to appear compelling and useful to the authors and owners of the content. In other words, added value for institutions appears more clear than added-value towards authors, faculty or other content contributors.

The two strategies that were adopted to recruit more content for the Institutional Repository indicate two challenges that are still present today in the context of many Institutional Repository Projects: One was to try a new strategy for recruiting faculty members in which they were approached on their own grounds and were addressed to in their language. The other was to enhance the repository to make it much easier for our faculty members to submit their items to the IR and to showcase themselves and their research. In short, there's project marketing, communication & awareness at one hand, and lowering technical thresholds on the other hand.

More recent, in April 2007, Philip M. Davis and Matthew J. L. Connolly published an introspective study about the reasons for non-use of Cornell University's Installation of the DSpace repository platform.¹³ They state that little has been done to evaluate repository projects, especially with regards to faculty participation. This observation reveals a problem that goes back to the very early phase of the project definition, scope and deliverables. As the metrics for evaluating the project and thus measuring the success should have been ideally defined from the start. Although not explicitly stated, it also leads to believe that the exact objectives of the projects might have been obscure as well.

In their research, they encountered a variety of reasons why faculty were not using the repository: redundancy with other modes of disseminating information, the learning curve, confusion with copyright, fear of plagiarism and having one's work scooped, associating one's work with inconsistent quality, and concerns about whether posting a manuscript constitutes "publishing".

One of the main conclusions is that in the library community, the "crisis in scholarly communication" is often framed as a crisis in access brought about by a combination of high journal price inflation and a stasis in library funding, resulting in journal cancellations¹⁴. In stark contrast, access to the literature appeared to be a non-issue for all faculty interviewed save one. In short, the faculty who were interviewed didn't perceive any added value in the Institutional Repository in the area of retrieval or access to research that might be useful for their own research.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



Likewise, in the area of dissemination of their own work, i.e. using the repository to disseminate their own work, the interviewees were convinced that researchers in their field already had adequate access to their work, and that (apart from one researcher) they believed there was little or no public interest in their work.

Like in the Rochester study from 2005, issues with marketing and communication regarding the initiative were also revealed in this study, stating that *“of the eleven faculty members interviewed, only four knew about the existence of DSpace, and only one member (the historian) had deposited items in it”*.

Future Challenges and Opportunities

From the aforementioned observations of perceived failure and external influences, such as emerging technologies and trends, we isolate a number important challenges today, that will shape the direction of repository projects in the future.

Project Definition: Vision, Strategy and Tactics

For newly starting repository projects, and already established ones, supporting a vision with strategies and relevant tactics are very relevant in the light of aforementioned perceived failures. In the conception of a project vision, there should be a clear distinction between institutional benefits, or benefits for one of the bodies of the institution (library, research coordination office, ...) and benefits for individual researchers. One of the reasons why this might be overlooked in the past, is that the core from C. Lynch's Institutional Repository definition, the *“organizational commitment to the stewardship of these digital materials”*, doesn't automatically translate into immediate benefits. Here are a few examples of different visions that could apply:

An **institution-centered vision** could entail that the institutional repository operates first and foremost as a modern academic bibliography, and that it significantly reduces time and effort for annual research reporting, by eliminating the need for paper delivery of lists. A second institutional benefit from offering an online environment, such as a repository platform, is that all available content can be clearly linked and identified with the institution. In this sense, the repository can carry out prestige and identity from the institute.

A **researcher-centered vision** is aimed at creating added value for individual researchers. The repository offers functionality or services that either solve problems, or introduce new benefits that can save time or help them in achieving tasks in novel ways. From the Cornell study, it became clear that in its current form, researchers haven't been able to identify any added value, resulting in the non-use. Up until today, it hasn't been proven that Institutional Repositories **can't** generate additional online exposure, and increase usage citations for exposed content. However, as became clear from both studies, simply asserting that the repository does this, proves to be insufficient to get significant commitment from faculty. Providing clear and recurrent information on the usage of a researchers material can illustrate the use of having materials in the repository. When the vision requires a strategy that will influence the behavior of a large number of diverse people (cfr different disciplines), the strategy will come at a certain cost and timespan. One possible strategy would be to bootstrap the repository with a batch

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



import or gathering phase, to have a starting mass of content in the repository and follow this up with testimonials resulting from the new exposure of these materials.

A **“general public” centered vision** could mean that the repository generates added value for the outside world, by operating as a portal to relevant content, being frequently visited by national and international visitors, originating from outside of the institution. This vision can link into an institution-centered vision, typically when the institution is an academic research institution. However, when the project originates from a consortium of libraries, who's core task is to serve the general public, this vision applies most. Any strategy implementing benefits for the general public should ensure that (technical) barriers to access are as low as possible, and that the content itself is being presented as attractive as possible to the target audience. Also an appropriate marketing and communication plan will apply to ensure that the initiative effectively reaches the target audience.

Setting SMART Objectives¹⁵

As indicated by the Cornell study, very little has been done in the area of evaluating Institutional Repository projects. As projects will choose different strategies, the ways to evaluate them, metrics that will be chosen, will differ as well. One example that can illustrate this, is the researcher-centric vision of serving the researchers and the internal scientific community of the institution with long term digital preservation (and stewardship) of a well defined set of content types of (academic) output. In this vision, and strategy to implement this individual item page visits, and downloads of full text, will not be a relevant metric to measure success or progress in implementing the vision.

Regarding measuring usage of repositories, different projects by JISC (PIRUS2) and DINI (OA-STATISTIK) are actively working on standardizing methods to measure usage statistics across different repository platforms.

But no matter what the exact goals or objectives will be, measuring and evaluating process towards the goal will be possible when the objectives or more detailed milestones are formulated as SMART: Specific, Measurable, Attainable, Realistic and Time-Bound.

Threats and opportunities from external initiatives

When stepping away from the conceptual definition of an Institutional Repository, and now looking specifically at the current state of the art in Institutional Repository Platforms, several emerging trends and initiatives pose both threats and opportunities for the existing platforms.

Articles and presentations making claims that repositories are “wrong” or “dead”^{16, 17} generally indicate a functional limitation that illustrate that a repository platform has failed or proven inadequate for one type of use case or goal. None of these articles claims that the conceptual idea from Lynch's definition of “organizational commitment to stewardship of created digital materials” itself, has become obsolete.

So the perspective to evaluate emerging technologies and novel approaches is the consideration whether another platform, technology or modus-operandi might be better suited to implement the repository vision, strategy and tactics.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



When considering recent progress in the fields of search engines, with special regard for **Google Scholar** and **Google Books**, the link can be made with “general-public” centered vision behind a repository project. Effectively, if the sole purpose behind the repository project is to make certain collections of Digital or digitized objects available for the general audience, it can be a a very cost effective solution to invite a corporation to perform the scanning and online dissemination for free. However, note that these external initiatives might not serve other visions or strategies, especially the institution-centered vision where branding, and linking between the institution and the content are considered to be a vital aspect of the vision. Although this paper has barely touched the subject of repositories for providing access to data, it’s interesting to note that Google recently abandoned it’s project to offer access to large public data sets.

A movement under the flag of “**Research 2.0**” is evaluating different existing schemes and initiatives in the area of social networking, such as LinkedIn, Facebook, with the objective to create novel, added value services for academic scholars. Some of these initiatives, such as Mendeley¹⁸, profiling themselves as the “iTunes for Academic Research Papers”, are started by (startup) corporations. While others, such as 2collab¹⁹ by Elsevier, and Connotea²⁰ by Nature publishing group, are initiatives by publishers (in the case of 2collab, Elsevier), to contribute to new business models in the publishing industry. All of these ideas are characterized by a very researcher-centered vision, aimed at providing added values for individual researchers. In this sense, they provide a realistic threat to Institutional Repository projects, driven by a researcher-centered vision. However, in important drawback is that often, the development and feature set is completely controlled by the corporation and very little on demand customization is possible.

In various domains, companies and research institutions are creating **layers of value-adding functionality on top of existing repository platforms**. Companies like Symplectic Ltd. (UK), @mire (Belgium) and ATIRA (denmark) are leveraging the potential of repositories in the area of Research Assessment, either by integrating them in the institution’s CRIS (current research information system) structure, or by adding reporting functionality in the repository itself. In the development of their own, freely available repository platform Zentity, Microsoft devotes crucial effort and external support to additional (authoring) tools that allow users to interface directly with the repository, from within their MS applications²¹.

Is there a future for Institutional Repositories in 2010 (and beyond) ?

While existing Institutional Repository technology continues to evolve, new technologies emerge and institutions are (re-)inventing their Institutional Repository strategies.

Depending on the chosen strategy, the implementation of one of today’s Institutional Repository Platforms might make more sense than another strategy.

Throughout different projects, testimonials and studies, it remains clear that organizations are more and more realizing both importance and added value benefits of the central concept behind Institutional Repositories: organizational commitment to stewardship of digital materials.

@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294



@mire nv

Technologielaan 9
B-3001 Heverlee
België

+32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294

- ¹ Lynch, Clifford A. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 (February 2003): 1-7.
< <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml> >
- ² Directory of Open Access Repositories - OpenDOAR
< <http://www.opendoar.org> >
- ³ NARCIS - National Academic Research and Collaborations Information System
< <http://www.narcis.info> >
- ⁴ SURF Share - Open Onderzoek
< <http://www.surffoundation.nl/nl/themas/openonderzoek/pages/default.aspx> >
- ⁵ Repositories and Preservation Programme
< <http://www.jisc.ac.uk/whatwedo/programmes/reppres.aspx> >
- ⁶ IRRA: Institutional Repositories and Research Assessment
< <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2005/irra.aspx> >
- ⁷ Neil Godfrey, INFORMAL comparison of some institutional repository solutions, 19th October 2008
< <http://metallogger.wordpress.com/2008/10/19/comparing-some-institutional-repository-solutions/> >
- ⁸ Jim Henson Works Repository
< <http://www.lib.umd.edu/digital/henson/> >
- ⁹ Maurits Van der Graaf, Kwame van Eijndhoven "The European Repository Landscape : Inventory study into present type and level of OAI compliant Digital Repository activities in the EU", Amsterdam University Press (2007)
< <http://dare.uva.nl/aup/nl/record/260225> >
- ¹⁰ Network of European Economists Online
< <http://www.neeoproject.eu/> >
- ¹¹ Ranking Web of World Repositories
< <http://repositories.webometrics.info> >
- ¹² Foster, N. F. & Gibbons, G. D-Lib Magazine doi:10.1045/january2005-foster (2005).
< <http://www.dlib.org/dlib/january05/foster/01foster.html> >
- ¹³ [Davis, Philip M.](#) , [Connolly, Matthew J.L.](#) Institutional Repositories: Evaluating the reasons for non-use of Cornell University's installation of DSpace, CNRI, D-Lib Magazine vol 13, n 3/4
< <http://ecommons.cornell.edu/handle/1813/5195> >
- ¹⁴ English, R.: Scholarly Communication and the Academy: The Importance of the ACRL Initiative. Portal: Libraries and the Academy 3(2003) 337-340.
- ¹⁵ Drucker Peter F. The practice of management. New York: Harper and Row Publishers, 1954.
- ¹⁶ Marshall, C. C. 2008. From writing and analysis to the repository: taking the scholars' perspective on scholarly archiving. In Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (Pittsburgh PA, PA, USA, June 16 - 20, 2008). JCDL '08. ACM, New York, NY, 251-260. doi: [10.1145/1378889.1378930](https://doi.org/10.1145/1378889.1378930)



¹⁷ Ian Stuart, Repositories are dead, long live repositories
< <http://jiscrepository.ideascale.com/akira/dtd/1412-784> >

¹⁸ Mendeley: Academic Software for Research Papers
< <http://www.mendeley.com/> >

¹⁹ 2Collab by Elsevier
< <http://www.2collab.com> >

²⁰ Connotea: free online reference management for clinicians and scientists
< <http://www.connotea.org/> >

²¹ Microsoft External Research - Tools and Resources for Scholarly Communication
< http://www.microsoft.com/mscorp/tc/scholarly_communication.mspx >

@mire nv

Technologielaan 9
B-3001 Heverlee
België

☎ +32 2 888 29 56

info@atmire.com
www.atmire.com

BTW BE 886.066.294