Institutional Repositories: an Internal and External Perspective on the Value of IRs for Researchers' Communities

Maria Cassella

University of Turin, Via Po, 17 - 10124 Torino, Italy,
maria.cassella@unito.it

Abstract

Institutional repositories represent extremely innovative technology, but repository managers still struggle to bring together a critical mass of content and to demonstrate their overall impact on research. In this paper I propose a set of Performance Indicators (PIs) to assess institutional repositories' success. Fourteen internal indicators are selected and inserted in the quadruple 'balanced scorecard' perspective. Three more indicators from an external perspective are then proposed and discussed by the author.

I hope that this study will foster the development of standard Performance Indicators for IRs in the very near future, in order to help IR managers to demonstrate their repositories' cost-effectiveness and success.

Key Words: institutional repositories; institutional repositories assessment; performance indicators; measurement; impact; balanced scorecard

Introduction

Institutional repositories (IRs) are some of the most innovative and creative components of digital libraries. They are intended to be a showcase for the research output of an academic or research institution. Nevertheless, repository managers still struggle to bring together a critical mass of relevant content from researchers and to attract attention and funding from decision makers. Despite librarians' efforts to promote their IRs 'the research policy

and decision-makers remain unconvinced that the IR is a strategic research tool' (Thomas, 2007).

This lack of awareness of the strategic importance of IRs becomes a critical issue in the current period of economic crisis and threatens IRs' long-term sustainability. Despite the fact that IRs are mainly implemented by means of open source software and 'costs in establishing a repository are modest when compared with other library initiatives' (Westell, 2006, p. 212), repositories do incur operating costs due to the technological infrastructure necessary to support their implementation and maintenance, the procedures involved in running value-added services and in supporting preservation strategies (e.g. system backup) , the costs of the staff involved in their management, and the costs of the vendor fees (Markey *et al.*, 2007).

In a time of financial crisis the need to assess IRs' success by adopting valid, appropriate and relevant Performance Indicators (PIs)[1] has become stringent.

This article proposes that the evaluation of repositories should be undertaken in a broad holistic framework which takes into account both internal and external measures. I will discuss a core set of fourteen Performance Indicators according to the quadruple structure of the balanced scorecard (see section 3): the user perspective, the internal perspective, the financial perspective, and the learning and growth perspective.

I will also discuss a few external indicators as measures of successful repositories. I hope that this study will foster the development of standard Performance Indicators for IRs in the very near future in order to help IR managers to demonstrate their repositories' cost-effectiveness and success.

Literature Review

In the plethora of articles on Institutional Repositories we can identify a well-defined subset of papers covering the subject of the assessment of IRs. Soon after the emergence and subsequent rapid growth of IRs in the United States and in Europe, IR managers were concerned about how to attract a critical mass of relevant content in their repositories, and therefore an early set of studies explored strategies and methods for performing qualitative

assessment, as authors' attitudes towards self-archiving were regarded as key factors of IR success. One of the early qualitative assessment studies on authors' perceptions towards self-archiving was a survey sponsored by JISC in 2005 1,296 researchers were investigated by Alma Swan and Sheridan Brown, consultants of Key Perspectives (Swan & Brown, 2005). They found out that 81% of respondent authors would have been willing to self-archive their research output in the case of an institutional mandate. To better understand faculty needs, Foster and Gibbons (Foster & Gibbons, 2005) performed a work-practice study based on anthropological observation of how faculty members of the University of Rochester do their research and writing in the same year.[2] Since 2005 manifold case studies have predominated in the literature on IRs. Almost every case study stresses the necessity of a pre-implementation qualitative assessment among researchers[3] in order to understand the different communities' needs and align the repository strategies and policies to the researchers' and the stakeholders' agenda.[4]

Later on, studies on IRs began to focus more narrowly on metrics and quantitative evaluation analysis. Westell (Westell, 2006) suggests eight factors of success selected for the evaluation of Canadian IRs. Out of these eight, six measures are internal factors: mandate, integration with planning, funding model, measurement, promotion, and preservation strategy, while two measures are external: relationship with digitisation centres and interoperability. Thibodeau (Thibodeau, 2007) proposes a more general framework for IR evaluation articulated along five dimensions: service, orientation, coverage, collaboration and state. Within each dimension Thibodeau poses questions to help IR managers to assess success. Swan (Swan, 2007) recommends a quality framework based on four domains: content recruitment, user awareness and involvement, workflow practices, and financial discipline.

More metrics-focused is the analysis of Xia and Sun (Xia & Sun, 2007b) who suggest the following IR assessment measures: number of deposits broken down in subcategories, i.e. by class, by subclass, by department/faculty, by version, by type, numbers by date, by depositor, by availability, by location, numbers by any other categories classified by a particular repository, availability of full text, cost per deposit, usage assessment. Xia and Sun also stress the importance for IR evaluation of factors such as authors' attitude, information on depositor, usage assessment and interoperability. In a subsequent article Xia and Sun (Xia & Sun, 2007a) develop an evaluation method

of repositories based mainly on two assessment factors: depositorship, i.e. depositor identity (which conflates author and editorial processes) and availability of full text. They selectively apply these criteria to a small number of repositories.

Carr and Brody (Carr & Brody, 2007) also suggest metrics to measure daily deposit activity. They criticise occasional mass deposit and recommend the use of the ROAR registry of institutional repositories to trace a more regular deposit activity. They propose a three-tiered breakdown of deposit repositories: repositories with a daily count of 1–9 deposits (weak); repositories with a count of 10–99 (healthy) and repositories with more than 100 deposits daily (batch imports). Last but not least Gwenda Thomas (Thomas, 2007) designs an IR performance indicator framework. The author bases her analysis on her experience with the IR of the Rhodes University and proposes a two-tiered structure: significant PIs to evaluate IR impact and secondary-level PIs to demonstrate IR cost-efficiency. The focus of the paper is on qualitative indicators which examine the fit between purpose of the service and research goals of the University, that is, the effect of the service on the end user or institution.

Another strand in the literature focuses more narrowly on external measures, defining them as the alignment of IRs with institutional goals and individual needs (Blythe & Chachra, 2005) and the interaction and impact of the library on scholarly life on campus (Markey *et al.*, 2009).

The Repositories' Value Proposition: Internal Measures in the Structure of the BSC

Performance measurement is a very complex task. It has different perspectives and can have different points of view. Reliance on one perspective and on one single point of view can be fatal. In 1992 Robert S. Kaplan and David P. Norton presented performance measurement in the quadruple 'balanced scorecard' (BSC) perspective. Their aim was to give top business managers a comprehensive evaluation framework by providing answers to four questions:

1. How do customers see us? (customer perspective);
2. What must we excel at? (internal perspective);

3. How do we look for shareholders? (financial perspective);
4. Can we continue to improve and create value? (innovation and learning perspective).

In the context of library performance measurement the BSC structure can be translated as:

1. user perspective;
2. internal process perspective;
3. financial perspective;
4. learning and growth perspective.

The BSC has already been adopted to revise the 45 PIs of the second edition of ISO 11620 : 2008, Information and documentation — Library Performance indicators, and the second edition of the IFLA handbook on 'Measuring quality'.

The internal value proposition for repositories can also be drawn in the BSC structure. The BSC model will help IR managers to align their repository strategies to the institutional mission and goals and to identify priorities in performance measurement by focusing on a core set of meaningful PIs.[5]

Core PIs from the User Perspective

Very differently from the commercial sector, IR assessment is mainly driven by the user perspective, which helps gauge the real level of commitment of the affiliated researchers'[6] communities to the IR.

The PIs proposed from the user perspective are all metrics representing an interaction between the researchers, in the twofold role of depositors and retrievers, and the repository:

1. Percentage of researchers depositing in the IR.

This measure should take into account only authors affiliated with the institution. It really measures researchers' commitment to the IR activity.

2. Average number of items deposited per researcher.

Year after year the average number of items deposited per researcher[7] shows the overall growth and the uptake of the self-archiving practice among

faculty members. This PI is a simple measure to calculate and quite meaningful, because within the same institution researchers' deposit activity may be very different from one researcher to another, from one community to another, and depending on whether deposit has been mandated or not. Due to different scholarly communication patterns some research groups are more naturally proactive in self-archiving. The age of depositing authors could also be a factor that influences deposit practice, as young researchers may be more aware of the advantages of open access (OA) and more keen to self-archive their scholarly output.

3. Number of communities depositing in the IR.

This PI gauges the level of diffusion and penetration of the repository among the different research groups active in an institution. As stated above it should always be taken into account that self-archiving is a practice differently adopted by researchers. Therefore a widespread deposit practice covering all affiliated research communities is a good indicator of the repository's success.

4. Number of items downloaded annually/monthly/daily.

This PI gives repository managers the measure of the real use of the repository among researchers. Downloads are a measure widely adopted by repository managers as download data are logged by repository software as a by-product of the Web request they receive. So far no suitable standard supports the production of sets of statistics for repositories, but usage reporting is a fast developing repository feature. To date two projects have explored the feasibility of defining a set of agreed standards for measuring and reporting usage of materials deposited in IRs: the Interoperable Repository Statistics (IRS), funded by JISC, and the DINI German Project 'Open Access Statistics' (OA-S).

Due to the lack of standards for the production of repository statistical reports, the German DINI Certificate 2006 also requires that access statistics should be accompanied by documentation explaining which criteria were used to collect the basic data and how they were processed (Dobratz & Scholze, 2006, p. 8).

Core PIs from the Internal Perspective

PIs from the internal perspective are all the indicators related to the assessment of the collections and of the value-added services developed by a repository.

As a matter of fact the variety, the richness and the completeness of the collections deposited in the repository and the number of value-added services developed for authors are strategic components of a successful repository.

5. Total number of items deposited annually.

As stated above this measure is universally adopted to assess the success of a repository. In our view this PI is not particularly meaningful for gauging how successful a repository is and it should be always applied in combination with other measures.

The total number of items deposited is strictly related to the repository deposit policy. Due to strong dissimilarities in institutions' goals and missions, deposit policies may be very different from one repository to another. They cover a number of strategic issues, mainly who is entitled to deposit, what type of materials are accepted and what formats the repository is going to accept. To date, the great majority of active repositories allow for the deposit of both full-text documents and metadata records, the latter often representing the critical content mass of a repository.

Another practice heavily influencing the number of items deposited is automatic record ingestion. As stated by Carr and Brody (Carr & Brody, 2007), very large repositories often harvest their content automatically. However, 'automatic ingestion of data may be not consistent with capturing the community's scientific and scholarly output.

6. Number of items deposited daily.

This PI is a relevant measure of the role gained by the repository in the institution among the researcher groups that recognise the advantages of depositing their scholarly output in the IR: visibility, broad dissemination, a tool to maintain updated CVs, a tool for preserving their digital scholarly production and so on.

7. Full-text availability of documents.

Ideally, a repository should contain the entire scholarly output of the institution. However it should be recognised that due to embargo and copyright issues only a percentage of the scholarly production can be self-archived as full text by researchers, even where document deposit has been mandated.

The calculation of this measure includes every full-text resource type that a repository can hold: journal articles, dissertations, unpublished reports and

working papers, book chapters, multimedia, learning objects and so on. To become a useful and successful scholarly communication tool a repository should contain predominantly 'qualified content',[8] i.e. full-text documents.

8.   Full-text availability of articles.

This is one of the most meaningful measures of a successful repository. In the digital environment researchers use and value a great variety of scholarly resources (Maron & Smith, 2008);[9] however, scientific articles still represent the target scholarly output, notably in STM disciplines. The calculation of this measure should include all article versions deposited in the IR:[10] pre-prints, post-prints, and publishers' pdf versions.[11] To identify article versions in repositories we recommend the adoption of the Version Identification Framework (VIF) which provides practical advices to authors and content creators, repository managers and those involved with repository software on how to identify versions better.[12]

9.   Number of collections active in the repository.

The number, the variety, the richness and the quality of the collections of a repository are good indicators of IR success and attract researchers' attention both as depositors and as retrievers.[13]

10. Number of value-added services.

According to Clifford Lynch (Lynch, 2003) a repository is above all a 'set of services'. As we previously stated for the collections, a rich array of value-added services offered is certainly one of the best ways to align the repository to researchers' needs and is a sign of the maturity and ongoing development of the repository (Swan, 2007). Value-added repository services may include: production and export of CVs, delivery of statistics on deposited items, reference citations services, author identification services, guides and FAQ on copyright for authors, controlled vocabulary add-on, commenting add-on, customized TOC, current awareness services, alerts and RSS/Atom feeds, and so on.

Core PIs from the Financial Perspective

The Open Access paradigm is neither free nor costless. Both self-archiving (the so-called Green Road) and publication in OA journals or the author-pays model (the so-called Gold Road) are subject to costs.[14]

The total costs of a repository include both initial start-up and annual maintenance costs (Swan, 2007). Depending on a set of variables and on how ambitious the repository is, costs can vary substantially.

According to the Association of Research Libraries' SPEC Kit 292, repositories' start-up costs range from $8,000 to $1,800,000, while operating costs range from $8,600 to $500,000, with a mean of $113,543 (Bailey *et al*., 2006).[15] As many repository projects are only funded in their start-up phase,[16] long-term sustainability is becoming a great concern among IR managers who are increasingly asked to demonstrate the cost-effectiveness of their repositories. Financial indicators are a strategic tool to assess IRs' immediate plain costs, but repositories' cost-effectiveness can really be measured only in the long term. 'Digital repository as a revolutionary yet experimental way of scholarly publishing has the potential to alter the tradition of information acquisitions and dissemination.' (Xia & Sun, 2007b, p. 77).

Therefore all possible outcomes, benefits, and consequences deriving from the broad dissemination of scholarly institutional output should be taken into account and promoted to stakeholders when assessing the repository's costs.

11. Cost per deposit.

Xia and Sun (Xia & Sun, 2007b) propose to calculate the cost per deposit by dividing the number of an IR's content documents by the total amount of the IR's expenses. All costs necessary for running an IR should be taken into account to calculate the cost per deposit: installation costs, hardware costs, costs for implementing value-added and customised services, quality control procedures costs, preservation procedures costs, costs for outsourcing services, and staff and promotion expenditure. Houghton *et al*. (Houghton *et al*., 2009) also estimate the costs of the time academics spend on depositing. To date repositories do not perform a quality certification function and no costs for peer review should be included in the calculation of this measure. However, in the very near future the certification function may become an automatic overlay service also offered by repositories. In this case certification should be calculated among other IR costs.

The cost per deposit can obviously be broken down per annum by dividing the number of items deposited annually by the annual maintenance costs.

12. Cost per download.

This is a very meaningful measure to evaluate the scholarly efficacy of repository collections. As stated earlier, the problem with statistical reports on repository usage is that to date no suitable standard supports the production of sets of statistics for repositories.

Some repository software system statistics offer the interesting functionality of breaking down statistics per research community. As a matter of fact the 1.6.1 DSpace release produces usage statistics broken down to the level of both communities and collections so that the cost per download per community and per collection can be assessed.

The cost per download should preferably be calculated per annum, but as many IRs' software packages allow to break down access statistics per month, figures on monthly costs per download can also be derived.

Core PIs from the Learning and Growth Perspective

While a lasting funding practice is essential to guarantee the long-term sustainability of repositories, allocation of human resources to IR activities is a strategic factor for assessing the institutional investment in the repository's future development.

Therefore we selected two staff-related PIs to be inserted in the learning and growth perspective:

13. Number of FTE repository staff.

According to the CLIR report on institutional repositories in United States (Markey *et al.*, 2007), in 2007 the overall average of people involved in their institutions' IR efforts was 7.2. This is of course only an average figure. In fact the number of people working for the IR may vary substantially during the IR implementation process, from one year to another, and from one facility to another.

As costs for staff and vendor's fees represent about 75% of the total IR costs (Markey *et al.*, 2007), a growing FTE staff curve over the years is a good indicator of increasing funds being allocated to the repository and indirectly of its increasing reputation among stakeholders.

Staff should always be calculated in FTE. No volunteer or occasional staff should be included in calculating this measure.

14. Expenditure on staff education and training activities.

Successful IRs require a combination of different skills and expertise: repository software as well as digital curation expertise, scholarly communication as well as IPR expertise, expertise on national and international Open Access projects (e.g. EU projects), expertise on knowledge management, expertise on social marketing and promotion strategies (Gierveld, 2006), and so on.

Therefore ongoing staff training is an essential prerequisite to building a successful repository and to assessing the repository's potential for growth and development.

The Repositories' Value Proposition: External Measures

External IR indicators should help managers to gauge the impact of the repository both at national and international levels and to assess its value as a research tool for end-users (i.e. for academic communities working for other institutions).

To measure the external value of a repository we propose the adoption of three main indicators:

1. Interoperability

Institutional repositories are not local phenomena. Technologically, organisationally and politically they are founded on the idea of interoperability. Interoperability indicates 'an ability and openness on behalf of the institution or library to contribute to national and international scholarship' (Westell, 2006, p. 216).

Interoperability also allows the repository to dispel the notion of digital archives as a silo of knowledge; it helps the repository to interact with a network of repositories (national repository networks such as DAREnet in the Netherlands, OA-Network in Germany, Arrow in Australia), to create new repository models (inter-institutional repositories such as the White Rose shared repository of the Universities of Leeds, Sheffield and York, the Kultur Consortium in UK), to guarantee a sustainable repository model and to ensure a long-term preservation strategy.

2. Total amount of external funding secured for IR projects (per annum)

A successful interoperable repository is also a repository that attracts funding.

The capacity of the IR to attract external funding either from policy makers, foundations, institutions or from private companies at local, national and international level is necessarily an indicator of the visibility and of the reputation attained by the repository as a hub of knowledge. Depending on the mission of the institution they serve and on the profile of their collections some repositories are more likely to attract local and national funding while others may become more active at the international level.

3. Participation in national and international projects

The Open Access paradigm is a global paradigm. It has no boundaries and contributes to the world-wide dissemination of the science. After ten years of repository development, projects supporting the self-archiving OA strategy are flourishing both at the national and international level.

The degree of repositories' participation in these projects assesses the level of internationalisation of an institutional repository and indirectly is a tangible indicator of the IR quality. In Europe the Digital Repositories Infrastructure Vision for European Research (DRIVER), now DRIVER II, provides a pan-European repository framework to research content across academic disciplines. Over 2,500,000 scientific publications, found in journal articles, dissertations, books, lectures, reports, etc., are harvested regularly from more than 249 repositories from 33 countries. In order to take part in DRIVER II repositories have to meet technological standards and must incorporate advanced end-users functionalities.

Conclusions

In this article we have recommended a set of performance indicators to measure repositories' success. The next step should be to adopt and test these indicators in order to assess the effective usefulness and relevance of the proposed PIs.

Acknowledgments

References

Bailey, C.W. *et al*. (2006): *Institutional repositories*, Washington, DC: Association of Research Libraries (SPEC Kit 292); available at: http://www.arl.org/bm~doc/ spec292web.pdf.

Björk, Bo-C. *et al*. (2010): 'Open Access to scientific journal literature: situation 2009', *PLoSONE*, 5(6), http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal. pone.0011273.

Blythe, E. and V. Chachra (2005): 'The value proposition in institutional repositories', *Educause review*, 40(5), p. 76–77; available at: http://www.educause.edu/ EDUCAUSE+Review/ERVolume402005/EDUCAUSEReviewMagazineVolume40/ 158012.

Carr, L. and T. Brody (2007): 'Size isn't everything: sustainable repositories as evidenced by sustainable deposit profiles', *D-Lib Magazine*, 13(7/8), http://www. dlib.org/dlib/july07/carr/07carr.html.

Dobratz, S. and F. Scholze (2006): 'DINI institutional repository certification and beyond', *Library Hi Tech*, 24(4), p. 583–594; available at: http://conference.ub.uni-bielefeld.de/2006/proceedings/dobratz_scholze_final_web.pdf.

Foster, N.F. and S. Gibbons (2005): 'Understanding faculty to improve content recruitment for institutional repositories', *D-Lib Magazine*, 11(1), http://www.dlib. org/dlib/january05/foster/01foster.html.

Gierveld, H. (2006): 'Considering a marketing and communications approach for an institutional repository', *Ariadne*, 49 (October), http://www.ariadne. ac.uk/issue49/gierveld/.

Houghton *et al*. (2009): *Economic implications of alternative scholarly publishing models: exploring the costs and benefits: a report to the Joint Information Systems Committee*, January, http://is.gd/3Q5KF.

Kaplan Robert, S. and David P. Norton (1992): 'The Balanced Scorecard: measures that drive performance', *Harvard Business Review*, January/February, p. 71–79.

Lynch, C. (2003): 'Institutional Repositories: essential infrastructure for scholarship in the digital age', *ARL Bimonthly Report* 226, http://www.arl.org/newsltr/226/ ir.htm.

Markey, K. *et al*. (2007): *Census of institutional repositories in the United States: MIRACLE project research findings: a CLIR report*; available at: http://www.clir.org/pubs/ reports/pub140/contents.html.

Markey, K. *et al*. (2009): 'Secrets of success: identifying success factors in institutional repositories', Open Repositories 2009 Conference Proceedings; available at: http:// hdl.handle.net/1853/28419.

Maron, Nancy L. and K. Kirby Smith (2008): ' Current models of digital scholarly communication: results of an investigation conducted by Ithaka strategic services for the association of research libraries', *Journal of electronic publishing*, 12(1), http://quod.lib.umich.edu/cgi/t/text/text-idx?c=jep;view=text;rgn=main;idno=333 6451.0012.105.

Poll, R. (2009): 'Benchmarking in the form of performance indicators and balanced scorecard', in *Library statistics for the Twenty-First Century world*, edited by Micheal Heaney, Saur: Muenchen, pp. 61–70.

Poll, R. and P. te Boekhorst (2007): *Measuring quality: performance measurement in libraries*, Saur: Muenchen (IFLA publications, 127).

Rieger, O. (2007): 'Select for success: key principles in assessing repository models', *D-Lib Magazine*, 13(7/8), http://www.dlib.org/dlib/july07/rieger/07rieger.html.

Swan, A. (2007): 'The business of digital repositories', in *A DRIVER'S Guide to European Repositories* (edited by Weenink, K., L. Waaijers and K. van Godtsenhoven). Amsterdam: Amsterdam University Press, http://www.keyperspectives.co.uk/ openaccessarchive/reports/The_business_of_digital_repositories.pdf

Swan, A. and S. Brown (2005): *Open Access self-archiving: an author study*, Technical report, http://eprints.ecs.soton.ac.uk/10999/.

Thibodeau, K. (2007): 'If you build it will it fly? Criteria for success in a digital repository', *Journal of digital information*, 8(2), http://journals.tdl.org/jodi/article/ view/197/174.

Thomas, G. (2007): 'Evaluating the impact of the institutional repository, or positioning innovation between a rock and a hard place', *New Review of Information Networking*, 12(2), pp. 133–146.

Xia, J. (2008): 'A comparison of subject and institutional repositories in self-archiving practice', *The Journal of academic librarianship,* 34(6), pp. 489–495, doi: 10.1016/ j.acalib.2008.09.016

Xia, J. and L. Sun (2007a): 'Assessment of self-archiving in Institutional Repositories: depositorship and full-text availability', *Serials Review*, 33(1), pp. 14–21, doi:10.1016/j.serrev.2006.12.003.

Xia, J. and L. Sun (2007b): 'Factors to assess self-archiving in institutional repositories', *Serials Review,* 33(2), pp. 73–80, doi: 10.1016/j.serrev.2006.09.002.

Westell, M. (2006): 'Institutional repositories: proposed indicators of success', *Library Hi Tech*, 24(2), p. 211–226.

Websites Referred to in the Text

ARROW, http://www.arrow.edu.au/

DAREnet, http://www.narcis.info/

DRIVER, http://www.driver-repository.eu/

DSpace manual release 1.6.1, http://www.dspace.org/1_6_1Documentation/

Kultur Consortium, http://www.jisc.ac.uk/whatwedo/programmes/reppres/sue/kultur.aspx

Interoperable repository statistics, http://irs.eprints.org/

NISO/ALPSP Journal Article Versions (JAV): recommendations of the NISO/ALPSP JAV Technical Working Group (NISO RP- 8- 2008), http://www.niso.org/publications/rp/RP-8-2008.pdf

Open Access statistics, http://www.dini.de/projekte/oa-statistik/english/

Ranking Web of World Repositories, http://repositories.webometrics.info/about.html

Repositories Support Project, http://www.rsp.ac.uk/

Version Identification Framework, http://www2.lse.ac.uk/library/vif/

The White Rose repository, http://eprints.whiterose.ac.uk/

Notes

---

[1] For the scope of this article we will adopt and adapt the definition of 'performance indicator' given in the ISO 11620: 2008 (2nd edition) Information and documentation — Library Performance indicators: 'A performance indicator is a numeric, symbolic or verbal expression, derived from library/repository statistics and data used to characterize the performance of a library/repository'.

[2] The core team that conducted the study at the University of Rochester included two librarians, a computer scientist, an anthropologist, a programmer, and a graphic designer.

[3] See in particular the Section 10 titled 'Institutional Repositories research studies' of the second edition of the Institutional Repository Bibliography by Charles W. Bailey, http://www.digital-scholarship.org/irb/irb.html.

[4] The benefits of conducting a stakeholders' analysis are highlighted by Rieger, O., 2007.

[5] According to Poll (2009) 'the BSC will function best with only a few chosen indicators, generally not more than 20'.

[6] Among IRs' users we include: researchers, students, notably PhD students, administrators, institutional leaders, service staff, end-users. Researchers are by far the most numerous and relevant user category.

[7] The calculation of this PI should also include items deposited by third parties on behalf of the researchers. Students archiving their thesis and dissertations should be excluded.

[8] According to Xia 'any repository with unqualified deposits such as a large number of non-full-texts will keep users away from using it' (Xia, 2008, p. 494).

[9] More precisely, Maron and Smith (2008) list eight types of digital scholarly resources: e-only journals, reviews, preprints and working papers, encyclopedias, data, blogs, discussion forums, professional and scholarly hubs.

[10] A very recent study by Björk et al. (Björk *et al.*, 2010) has assessed that in 2009 only 24% of all green open access articles were deposited in institutional repositories. Authors rather prefer self-archiving in subject-based repositories (43%) and on personal websites (33%).

[11] A list of publishers allowing the deposit of their pdf article version is available on the Sherpa/RoMEO website at: http://www.sherpa.ac.uk/romeo/PDFandIR.html

[12] In alternative the NISO/ALPSP Journal Article Versions (JAV): recommendations of the NISO/ALPSP JAV Technical Working Group (NISO RP-8-2008) can be adopted to identify article versions.

[13] An external measure of success is the repositories ranking available through the Ranking Web of World Repositories, performed by the Cybermetrics Lab of the Spanish Research Council.

[14] Costs of the OA publishing model have been assessed in the study of Houghton *et al.*, 2009.

[15] Swan (2007) also gives some useful data on repositories' costs by comparing a sample of four repositories: set-up costs range from 3,900 pounds (SHERPA Nottingham repository) to 2.4 million dollars (MIT DSpace) and running costs range from 30.000 euros (National University of Ireland repository) to 285.000 dollars (MIT DSpace).

[16] This statement is notably valid for the repository's 'institutional business model' which is one typology of a set of business models for repositories. Swan (2007) identifies five operational models for repositories: institutional model, community model, public sponsor model, subscription model, and commercial model.