

Jun 2nd, 12:00 AM

Assessing Data Linking in Transportation Technical Reports

Lisa Zilinski

Purdue University, ldz@andrew.cmu.edu

Paul Bracke

Purdue University - Main Campus, pbracke@purdue.edu

David Scherer

Purdue University - Main Campus, daschere@andrew.cmu.edu

Lisa Zilinski, Paul Bracke, and David Scherer, "Assessing Data Linking in Transportation Technical Reports." *Proceedings of the IATUL Conferences*. Paper 2.
<https://docs.lib.purdue.edu/iatul/2014/dlspace/2>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries.
Please contact epubs@purdue.edu for additional information.

ASSESSING DATA LINKING IN TRANSPORTATION TECHNICAL REPORTS

Lisa D. Zilinski

lzilins@purdue.edu

Paul J. Bracke

pbracke@purdue.edu

David A. Scherer

dscherer@purdue.edu

Purdue University Libraries
West Lafayette, Indiana, United States

Abstract

In 2013, the Purdue Libraries partnered with the Joint Transportation Research Program (JTRP), a collaboration between Purdue University and the Indiana Department of Transportation (INDOT), to incorporate the publication of research data sets into the technical report publication process. This project extended an existing process, developed in partnership between the Purdue Libraries, Purdue University Press, and JTRP, which leveraged Purdue's institutional repository to replace artisanal production processes with a more professional approach (Zilinski, Scherer, Bullock, Horton, & Matthews, forthcoming; Newton, Bullock, Watkinson, Bracke, & Horton, 2012). The existing workflow focused on production of textual documents, but did not address opportunities related to research data or geospatial information within the reports. The new workflow was intended to improve access to, and usability of, research data embedded in the reports by combining functions of Purdue e-Pubs with functions of the Purdue University Research Repository (PURR). This paper will present a framework for repository assessment including performance indicators and instruments measuring the impact of leveraging data and institutional repositories for linking publications to data. This framework will allow us to address research questions such as: What is the strategic value of linking data and publication for researchers and research centers? How can repositories contribute to key performance outcomes for researchers and research center administrators? What is the value of this initiative to library administration? Why is this relationship and initiative important to the libraries in developing repository and data services?

Keywords

library assessment, scholarly communication, institutional repositories, data publication

Introduction

Assessment is an important component of offering information services, including repository services.

Currently, assessment is generally focused on value to individual researchers, rather than examining the overall value and benefits of the repository service to organizational stakeholders and consumers. At the dawn of the era of institutional repositories (IRs), Crow (2002) posited that the value proposition of institutional repositories was that they are key in developing a new paradigm of scholarly publishing and that they could enhance institutional visibility and prestige. While these factors remain a significant part of the value of IRs, there are other ways in which IRs may provide value to stakeholders.

In the past decade, there has been much research conducted on the evaluation and assessment of repositories. There are a few different tools available that can assess and audit the reliability and integrity of repositories, such as the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) audit (Digital Curation Center, 2009) and the ISO 16363 Trusted Digital Repository (TDR) Checklist (International Organization for Standardization, 2012). These tools allow for the assessment of the repositories but do not assess the impact and value the repositories have on the different stakeholders.

Background

Value and Impact of Repositories

Burns, Lana, and Budd (2013) outlined both a set of propositions about the value of repositories and cost-metrics to demonstrate this value. According to them, the value of repositories is derived from: (a) establishing capacity for libraries to participate in system of scholarly communication; (b) acceptance and use of repositories, both from the perspective of authors and consumers; (c) discoverability of content; and (d) impact on reputation of larger institutions. Cost per visit, search, and download were then proposed as value metrics. Cullen and Chawner (2011) described that the value of repositories lay with the evolution of repository-based services that could meet the four key functions of the scholarly communication system of registration, certification, awareness, and archiving as prescribed by Roosendall and Geurts (1997).

Cullen and Chawner (2010) described the inherent conflicts between various repository stakeholders, especially in regards to faculty researchers. They noted that although there may be conflict between the needs of researchers and their institution, the value of the repository for both parties laid in the capabilities of the repository and repository services that lessened or eliminated the barriers of usage and deposit by providing enhanced scholarly communication services that could fill the gaps in the more historically established scholarly communication systems. In their study, St. Jean, Reih, Yakel, and Markley (2011) further mentioned that respondents highlighted the importance of awareness from all vested parties, including but not limited to content creators, administrators, and end-users as a means to measure value and impact of repositories and repository services.

Although it is noted that repositories carry a level of value for authors and stakeholders, it has been more problematic to assess their impact. Scholarly communication librarians and repository managers alike have seen that the “build it they will come” mantra has not carried over to institutional repositories. This has led to the perspective of repositories not making the impact that had been expected (Cullen & Chawner, 2011). A major reason for this perspective is also derived from the fact that the measurement of impact is based more upon local opinion than on standards that could be applied across a repository or repository system. In their study of repositories being formed in academic institutions of New Zealand, Cullen and Chawner noted that the institutions viewed their repositories as a success even though there were no formal standards. In his 2013 study focusing on investigating webometric measures to evaluate repositories, Smith found that conventional measurements of research impact did not correlate with the measurement of impact for repositories.

A noted and novel usage of repositories has been through library publishing services and activities, which has furthered their impact and value. These efforts are also noted for their involvement and partnership of university presses. Crow (2009) noted that there are numerous projects and potential opportunities in which the libraries and university press could create partnerships between these two distinct, yet parallel campus organizations. Clement (2011) further explained, “the university is a publishing dynamo that, ironically, hardly knows in any central way what is going on from unit to unit.” By working alongside one another, university presses and library publishers could not only identify local opportunities, but also understand what gaps have not been filled in the more traditional scholarly publishing models. Walters (2012) found that the value of university presses and library publishing services could be in providing the level of satisfaction that may be met in lieu of dissatisfaction faculty have in regards to the publishing services from existing publishers. By collaborating, university presses and libraries have focused on incorporating their current and newly formed skill sets to the solution itself instead of focusing on the source of the solution. Huwe (2013) described these new skill sets as the quest for creating solutions which would span commercial and professional boundaries.

Bankier and Perciali (2008) noted that library publishing activities that work with their repositories are very well suited for campus-grown scholarship, including but not limited to journals, technical reports, and student scholarship. By collaborating in library publishing ventures, the press becomes more exposed to campus-based publishing activities while the libraries, in-turn, can address the knowledge they have within the campus disciplines through their subject specialists and departmental liaisons. Hannsson and Johannesson (2013) described in their study on academic library support for scholarly publishing the

knowledge of publication habits within different subject domains. This knowledge is something librarians more or less explicitly relate to when discussing publication strategies with faculty.

Organizational Value & Impact

It is important to situate assessment of IRs within a broader institutional context, however. Administratively, it is important for any unit within an institution of higher education to demonstrate institutional centrality and alignment with mission in order to improve long-term sustainability and financial support, especially during times of budgetary challenges (Casella, 2010). Although there are a variety of strategies for doing so (Hackman, 1985; Ashar & Shapiro 1988), participation in IR activities may contribute to demonstration of centrality for a variety of stakeholders. Additionally, there are a multitude of changing contexts within higher education that should be taken into account when considering the value of IRs, or other library services for that matter, for stakeholders. Research centers and departments, for example, may be attempting to respond to institutional incentives related to interdisciplinary research (Etzkowitz & Leydesdorff, 2000), market-oriented behaviors (Slaughter & Rhoades, 2004; Bercovitz & Feldman, 2008), and benchmarking (DiMaggio & Powell 1983). Jain (2011) found the organizational benefits of IRs could be classified into three categories based upon the primary groups of stakeholders and creators: the institution, authors and content creators, and the library.

While measures of scholarly value are critical, as indicated by the usage of content and returns to the researcher and institution's academic prestige, it is also important to understand the ways in which IR stakeholders must demonstrate value within their institutions and the role IR assessment might play. An internal assessment could be utilized to measure the value of the repository. One suggested way to determine the internal value of the repository is to look to other established measurements that have been utilized to measure the impact of library-based services and activities. Casella (2010) suggests that the internal value proposition of the balanced scorecard (BSC) perspective translated from traditional library performance could be utilized to determine the alignment of the repository the institutional mission and goals. Casella also elaborated that several performance indicators could also be further utilized to assess internal and external values from the perspectives of users, additional internal perspectives, and financial perspectives.

Strategically, it is also important for those in charge of implementing the IR and repository services to understand what they wish to gain from the repository and the services they offer to determine what focuses should be further developed (Jain, 2011). By becoming a central unit of campus scholarship for faculty and students, the repository also has the potential to become a tool to manage the university's image (Giesecke, 2011).

Preliminary Framework for Assessment

This project will develop a framework for the assessment of institutional repositories. The primary outputs of this project will be:

1. A framework for assessing value and impact of repositories, including organizational and administrative impact, and performance indicators; and
2. Instruments that can measure the impact to stakeholders, including the library and partner organizations of leveraging institutional repositories.

The framework will be evaluated and refined by assessing the repository collaborations between the Purdue Libraries and the Purdue Joint Transportation Research Program (JTRP), and later confirmed in a more generalizable manner through assessments in other contexts.

This project will build from a preliminary framework, as depicted in Table 1, taking a holistic view of the value of IRs to a range of institutional stakeholders. Performance indicators and assessment instruments will be developed to enable the evaluation of the IR from the viewpoints of each group. The preliminary stakeholder groups are: researchers, research centers, departments and colleges, the library, and university administration. For each of these groups, both scholarly and administrative value will be assessed. Although, in the preliminary matrix, most of the types of value cut across all stakeholder groups, it is expected that differentiation in interests among stakeholder groups will emerge and the assessment framework will become substantially more nuanced.

	Researcher	Research Center	Departments and Colleges	Library	University Administration
Scholarly Value					
Academic Impact	Bibliometric Indicators (Citation Counts, h-index, Altmetrics)				
Visibility and Dissemination	Coverage in Search Engines, Ranking in Search Engines, Inclusion in A&I Databases, Inclusion in Linking Services				
Impacts on System of Scholarly Communication	Changes in Acceptance of OA by publishers, faculty, administrators; Acceptance of Data as Publication Form				
Social Impacts	Press Coverage	Outreach Impacts	Outreach Impacts	Outreach Impacts	Outreach Impacts
Administrative Value of Repositories					
Cost	Dissemination Costs	Dissemination and Administrative Costs	Dissemination Costs	Dissemination, Preservation, and Administrative Costs	Dissemination, Compliance, and Preservation Costs
Compliance	Compliance with Funder Mandates				
Trustworthiness	DRAMBORA, ISO 16363 (TDR)				
Centrality	Alignment with Institutional Mission and Priorities				

Table 1: Preliminary Assessment Matrix

Although this paper proposes a view of IR assessment that takes into account both scholarly and administrative values of IR programs, the research context of IRs dictates that scholarly value must remain primary. In a college or university context, impacts on academic prestige will be core to the ways in which value will be measured, and one could view the potential contributions of IRs to this end in several ways. First, existing models for assessing research quality will continue to be important, including both bibliometric indicators such as impact factors and h-index, as well as emerging approaches such as altmetrics. Second, quality research is more impactful when it is visible. Accordingly, assessment of the contributions of IR to research visibility will be important to all stakeholders in the university. The discoverability of IR resources within web-scale and institutional infrastructures for discovery and access, therefore must be made measurable. Third, a primary academic value of IRs is their role in effecting change in the system of scholarly communication. Although it can be difficult to directly measure the impact of an individual repository or repository collaboration upon the larger system, it is possible to document innovative efforts, such as the linking of textual and data objects within IRs, that have an impact on aspects of the broader scholarly communication environment. These impacts could manifest themselves in a variety of ways, ranging from impact on institutional prestige from developing new approaches to more fundamental impacts on the acceptance of new forms and models of scholarly communication. Finally, academic research is conducted to expand human knowledge and impacts of research on the general public can accrue prestige to researchers and institutions, but also have an impact on the communities in which universities are situated - a particularly important consideration for public institutions.

Models for assessing IRs that focus on scholarly value, however, do not fully demonstrate the impacts of IRs on stakeholders. IRs can have a range of administrative or organizational impacts upon the institutions in which they are situated. In an era of constrained budgets for higher education and libraries on a global level, budgetary impacts are critical. It is important for libraries, for example, to be able to demonstrate that they are offering repository services in a cost-effective and trustworthy manner, and that the returns to investments in repositories are accrued at a level acceptable both to library and university administrations. Looking beyond the library, however, it is also important to develop better measures for cost avoidance in other areas of the library. External funding for research is important for most institutions and demonstrating the ways in which IRs allow stakeholders to meet funder mandates is another manner of demonstrating administrative impact. Finally, there are ways in which individuals or units within an institution of higher education can demonstrate organizational centrality and alignment with institutional mission through documentation of IR activities. For example, IRs present a way for libraries to demonstrate stronger connections to the campus research enterprise. They may provide other opportunities for academic units to demonstrate centrality - through maximization of unit prestige, or alignment with emerging academic concerns such as interdisciplinarity.

Putting the Framework into Practice

The collaboration between the Purdue University Libraries and JTRP was a logical choice for evaluating and refining the assessment framework. There are a number of stakeholders involved in the collaboration, including Libraries faculty and staff, research center administration and researchers, and civil engineering faculty. In reviewing the literature, we did not find any current assessment method that would allow us to measure the impact or value of the relationship between the library and the research center. Our goal is to develop a model for assessing stakeholder value through developing criteria and processes for assessing the work done in constructing the repository systems for the Joint Transportation Research Program (JTRP).

Proposed Case Study

JTRP is a partnership between the Indiana Department of Transportation (INDOT) and Purdue University, established in 1936. This partnership supports collaboration between higher education institutions, agencies and industry to further Indiana transportation infrastructure research including continuous improvement in planning, design, construction, operation, management and economic efficiency (JTRP website). Since 2006, the Purdue University Libraries has been working closely with JTRP authors and administrators to improve production, dissemination, and stewardship of technical reports produced over

the last 60 years, and resulted in the development of the JTRP Technical Reports collection in Purdue's institutional repository, e-Pubs (Newton et al., 2011).

The impact of this collaboration was demonstrated using COUNTER-compliant usage statistics that show the 1,500 technical reports published in e-Pubs were downloaded over 668,000 times between 2006 and 2013. In 2012 alone, the reports were downloaded over 190,000 times, increasing about 10% per month.

It is projected that downloads will approach 250,000 in 2013, which is approximately a 30% increase over 2012 (Zilinski et al., forthcoming). These metrics are only a partial indicator of the value of the repository and focus purely on quantitative measures and highlights the gap between current assessment activities and what is needed to indicate stakeholder value.

In 2013, we expanded this partnership to include the publication of data sets in the Purdue University Research Repository (PURR), Purdue's data repository. The Libraries led the collaboration to incorporate the publication and linking of research data sets to the corresponding technical reports by developing a comprehensive workflow. This workflow combined the JTRP technical report publication workflow with the PURR data set publication workflow. We were able to create points of cross-repository communication and interaction, which assisted the authors, editor, and repository managers in coordinating the final publications. The DOIs of the datasets were linked and published in the metadata and on the cover page of the technical report while the technical report DOI was published in the metadata and landing page each of the corresponding data set publications. Currently, we are measuring the impact through the use of Altmetrics, downloads, access, and citations. Through this work, we identified four practices that we identified as best practices: link workflows, interact early with the data repository, use traditional publication attributes such as bibliographic metadata and DOIs, and monitor and evaluate usage and access statistics (Zilinski et al., forthcoming).

Applying the Matrix

The JTRP assessment pilot consist of two phases, which will represent first stages in fully developing the proposed assessment framework. The first phase will employ a grounded theory approach to further developing a model of stakeholder value in IRs. The researchers will conduct interviews with stakeholders, which will provide a number of insights into the question of IR assessment. For example, semi-structured interviews based on the research literature on IR assessment and organizational behaviors will allow us to verify findings from the literature on IR assessment and organizational behaviors, and also determine the performance metrics and evidence that is most valued by stakeholders in demonstrating IR value. The interviews will also allow the researchers to identify gaps in the preliminary framework and develop nuanced approaches to developing assessment instruments for individual stakeholder groups. Finally, the interviews will help the researchers prioritize the development of assessment instruments representing various aspects of the preliminary assessment framework.

Following the interview process, a prioritized list of assessment instruments will be developed and piloted with JTRP stakeholders. On a practical level, this will ultimately result in an assessment report for the use of JTRP and Libraries staff. On a theoretical level, this will provide a suite of assessment instruments validated in a single case that can be tested and generalized through further applications at Purdue and other institutions.

References

- Antelman, K. (2004). Do Open-Access Articles Have a Greater Research Impact? *College and Research Libraries*, 65(5), 372-382.
- Ashar, H. & Shapiro, J.Z. (1988) Measuring Centrality: A Note on Hackman's Resource-Allocation Theory. *Administrative Science Quarterly*, 33(2), 275-83.
- Bankier and Perciali. (2008). The Institutional Repository Rediscovered: What Can a University Do for Open Access Publishing? *Serials Review*, 34(1), 21 – 26. doi: 10.1016/j.serrev.2007.12.003
- Burns, C.S., Lana, A., & Budd, J.M. (2013). Institutional Repositories: Exploration of Costs and Value. *D-Lib Magazine*, 19(1/2), doi:10.1045/january2013-burns
- Casella (2010). Institutional Repositories: An Internal and External Perspective on the Value of IRs for Researchers' Communities. *LIBER Quarterly*, 20(2), 210-225.

- Clement, R. (2011). Library and University Press Integration: A New Vision for University Publishing. *Journal of Library Administration*, 51(5-6), 507-528. doi: 10.1080/01930826.2011.589330
- Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper. *ARL Bimonthly Report* 223.
- Crow, R. (2009). Campus-based Publishing Partnerships: A Guide to Critical Issues. Washington, D.C.: SPARC
- Cullen, R. & Chawner, B. (2010). Institutional Repositories: Assessing Their Value to the Academic Community. *Performance Measurement and Metrics*, 11(2), 131-147. doi:10.1108/14678041011064052
- Cullen, R. & Chawner, B. (2011). Institutional Repositories, Open Access, and Scholarly Communication: A Study of Conflicting Paradigms. *The Journal of Academic Librarianship*, 37(6), 460-470. doi: 10.1016/j.acalib.2011.07.002
- Digital Curation Center & Digital Preservation Europe. (2009). Digital Repository Audit Method Based on Risk Assessment (DRAMBORA). Glasgow, Scotland.
- Giesecke, J. (2011) Institutional Repositories: Keys to Success, *Journal of Library Administration*, 51(5-6), 529-542, DOI: 10.1080/01930826.2011.589340
- Hackman, J.D. (1985) Power and Centrality in the Allocation of Resources in Colleges and Universities. *Administrative Science Quarterly*, 30(1), 61-77.
- Hannsson, J. and Johannesson, K. (2013). Librarians' Views of Academic Library Support for Scholarly Publishing: An Every-day Perspective. *The Journal of Academic Librarianship*, 39, 232-240. doi: 10.1016/j.acalib.2013.02.002
- Huwe, T. (Sept/October 2013). Digital Publishing: The Next Library Skill. *Online Researcher*. 51-55.
- International Organization for Standardization. (2012). ISO 16363:2012: Space data and information transfer systems – Audit and certification of trustworthy digital repositories. Geneva, Switzerland.
- Jain, P. (2011). New Trends and Future Applications/Directions of IRs in Academic Institutions. *Library Review*, 60(2), 125-141. doi: 10.1108/00242531111113078
- Joint Transportation Research Program. (n.d.) Purdue University Retrieved January 27, 2014 from <https://engineering.purdue.edu/JTRP>
- Lynch, C. A. (2003). Institutional repositories: Essential Infrastructure for Scholarship in the Digital Age. *Portal: Libraries and the Academy*, 3(2), 327-336. doi: 10.1353/pla.2003.0039
- Newton, M.P., Bullock, D.M., Watkinson, C., Bracke, P.J., & Horton, D.K. (2012). Engaging New Partners in Transportation Research: Integrating the Publishing, Archiving, and Indexing of Technical Literature into the Research Process. *Transportation Research Record*, 2291, 111-123. doi: 10.3141/2291-13
- Roosendall, H.E. & Geurts, P.A.T.M. (1997). Forces and Functions in Scientific Communication: An Analysis of their Interplay. Presented at the CRISP 1997 Cooperative Research Information Systems in Physics. University of Oldenburg, Germany, 1997.
- St. Jean, B., Rieh, S.Y., Yakel, E., & Markey, K. (2011). Unheard Voices: Institutional Repository End-Users. *College and Research Libraries*, 72(1), 21-42.
- Smith, A.G. (2013). Web Based Impact Measures for Institutional Repositories. Proceedings of the ISSI 2013 Conference. Retrieved January 24, 2014 from <http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/2881/paper.pdf?sequence=2>
- Walters, Tyler. (2012). The Future Role of Publishing Services in University Libraries. *Libraries and the Academy*, 12(4), 425-454. doi: 10.1353/pla.2012.0041
- Zilinski, L.D., Scherer, D.A., Bullock, D.M., Horton, D., & Matthews, C.E. (forthcoming). Evolution of Data Creation, Management, Publication, and Curation in the Research Process. *Transportation Research Record: Journal of the Transportation Research Board*.