

Purdue University Purdue e-Pubs

Charleston Library Conference

Moving Technical Reports Forward

David Scherer

Purdue University, dscherer@purdue.edu

Roberto Sarmiento

Northwestern University, r-sarmiento@northwestern.edu

Maliaca Oxnam

University of Arizona, oxnam@u.library.arizona.edu

Charles Watkinson

Purdue University, cwatkinson@purdue.edu

Follow this and additional works at: <http://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at: <http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

David Scherer, Roberto Sarmiento, Maliaca Oxnam, and Charles Watkinson, "Moving Technical Reports Forward" (2012). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/10.5703/1288284315133>

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.

Moving Technical Reports Forward

David Scherer, Scholarly Repository Specialist, Purdue University

Roberto Sarmiento, Head of Transportation Library, Northwestern University

Maliaca Oxnam, Technical Report Archive and Image Library Chair, University of Arizona

Charles Watkinson, Director of Purdue University Press and Head of Scholarly Publishing Services, Purdue University

Abstract

Technical reports have always posed problems for libraries and librarians. They are often bibliographically inconsistent, difficult to source, and published to varying standards of quality. In some science and technical fields, these reports are also large in number and central in importance. Additionally, established workflows for acquiring and preserving technical reports in distributed repositories have been undermined by the transition from print to digital. Overall, the "grey literature" challenges librarians face have increased.

This paper presents three case studies of how academic libraries have found innovative ways to face the problems of technical reports and improve their production, dissemination, and preservation; thus reducing the duplication of research efforts and saving public funds. Transportation is one example of the disciplines where these described changes are taking place, and the opportunities for libraries to improve the technical report workflow in this field will be a particular focus of the session.

Readers can expect to learn about the challenges of handling technical reports in the digital age and the opportunities that exist for improving discoverability and dissemination in the networked environment. A particular focus will be on new roles for libraries and librarians, and how library publishing and data management services can offer new opportunities for partnerships with researchers.

Technical reports have always posed problems for Libraries and Librarians. They are often bibliographically inconsistent, difficult to source, published to varying standards of quality, and are also large in number. Additionally, established workflows for acquiring and preserving technical reports in distributed repositories have been undermined by the transition from print to digital. Overall, the "grey literature" challenges librarians face have increased. The field of transportation is one example of a discipline where such challenges exist, but also where opportunities exist for libraries and librarians to assist researchers in the handling of technical reports, and their overall discoverability and dissemination. Since these publications often constitute the only publication of important, often taxpayer-funded, and consist of primary research, engaging with the information management challenges that technical reports and other forms of "grey literature" pose is an important activity and role for Librarians and Libraries to accept.

This paper will describe through three case studies from Northwestern University, the Technical Report Archive and Image Library (TRAIL), and Purdue University; how libraries and librarians are exploring new techniques and practices for handling technical reports. We illustrate some new roles for libraries and librarians and show how library-based publishing services can create new and fruitful partnerships with researchers and their respective communities.

Transportation Technical Reports Overview: A View from the Eye of the Hurricane

Technical reports are an integral part of research and publication within the transportation community of this country. They constitute one of the principal ways by which federal and state departments of transportation, university transportation centers, public transportation

agencies, and, to a lesser degree, private organizations and present research results.

Historically, the publishing of these reports was done in paper format, but within the last 5 years, the publishing and distribution has moved to CD-ROM and to the web. Reports are typically on rather technical or engineering subjects, spanning all transportation modes, with the bulk of them related to highway construction and safety. Although we have no reliable way to assess the number of technical reports published in any given year, we believe that perhaps over a thousand are published every year.

For the most part, the production of technical reports is due to contractual requirements from the funding agencies to show that the research was conducted. The funding is mostly provided through federal and state grants. The amount of funding varies between tens of thousands to several million dollars per project or multi-year cycle. Thus, there is a lot at stake for the agencies to show returns and provide results. In addition, we need to keep in mind that the funding is mostly generated by our taxes, so we as citizens also have a stake on the production of these reports. Producing agencies see the reports not only as a requirement, but also, as a form of public relations and branding for their agencies due to the money and prestige involved. In addition, research engineers see the reports—and its results—as a way to establish a pecking order within their research community.

Challenges

For the purpose of this paper, we will focus our discussion on technical reports published by the states' department of transportations (DOTs) and by federally funded university transportation centers (UTCs). Technical reports present libraries and librarians with many, varied challenges. However, we can categorize them into three broad classes: lack of consistency, stewardship, and (lack of) access.

Among the most salient challenges we encounter:

- Each state DOT and UTC has its own standard for publication, delivery, writing style, web

retention, etc. This lack of standardization within the 50 DOTs and 22 UTCs does not allow for a single or simple approach to handling the whole body of information.

- Because of this lack of standards, the transportation research community has gone through a slow, painful, and disjointed transition from paper to digital documents. There is simply no consensus among DOTs on how to approach the production of reports in electronic format.
- Although a relatively simple and straight forward form which provides bibliographic information and metadata is required at the front of every report, in many cases it is poorly used or ignored by researchers and/or report writers. In fact, it is not uncommon to hear faculty being surprised by the fact that keywords are required or by the fact that they never had any thoughts regarding the need for the form.

In addition, up to a few years ago, it was not uncommon to talk to faculty/researchers and find out that they have never contacted the university library for any kind of support or never thought of partnering with library staff to receive help with research, production, or archiving of the reports. Fortunately, this is now changing. One unique challenge is faced with the cataloging of state DOT-produced reports. A number of agencies, both with and without libraries, do not catalogue their own reports but rather send it to another agency's library to have it catalogued and stored. This is a burden on libraries willing to perform original cataloguing, and the delay impacts timely access to the reports.

Even though a great number of reports are now produced in electronic format and although regulations require agencies to archive their reports with the National Transportation Library, many do not follow this requirement and then publish the reports on their agencies' webpages from which, after an undetermined amount of time, they are taken down and become unavailable. This misstep in stewardship is also aggravated by the lack of persistent URLs. Another less common, but still often encountered, reason

for the disappearance of reports from DOT websites is the liberal use of claims of “national security” concerns on reports that then leads to agencies “vanishing” them.

Although the transportation research community does have a national distribution list mechanism in place for the dissemination of reports and even though in almost all cases research archiving is a funding requirement, there are still an undetermined number of agencies that for unknown reasons do not comply. For over 13 years now, transportation agencies have been required to submit copies of their reports either on CD-ROM or URLs to the National Transportation Library (NTL) for cataloguing and archiving of digital reports only. Although the NTL holds an incomplete record of digitally produced reports, this is the closest our community has come to having a “national” repository. However, we also need to keep in mind that the NTL only archives digital copies, thus leaving the bulk of the technical reports literature—hard copies—not accounted for and not available in digital format.

Sadly, the archiving of hard copy/historical reports was not, nor is now, comprehensively undertaken by any U.S. public or private institution.

It is important to reiterate that although it is required for agencies to deposit copies of their technical reports with the NTL and the Transportation Research Board (TRB) Research in Progress (RiP) database, some agencies still ignore or fail to follow this mandate and do not file. This lack of systematic, comprehensive archiving affects the nation’s transportation research agenda.

Human factors also contribute to the challenges associated with technical reports. The whole cycle of research, writing and publication of reports is, for the most part, done by engineers. Engineers who, for the most part, write in “technical English,” may pay little attention to style and readability, may not be particularly concerned with providing metadata, and may not pay much attention to publishing formats. Then, when the report is published to the web we encounter management issues related to retention periods and accessibility. All these combine to further

impact the accessibility and stewardship of these reports.

Opportunities for Librarians

Despite all these challenges impacting access, stewardship, and standardization of transportation technical reports, there are very positive signs of improvements. Our nation’s transportation information community is a well-organized group of librarians at the regional, national, and international level who are working to actively address and partner with research engineers to solve or ameliorate these issues.

Currently there are several groups/committees sponsored by TRB (Transportation Research Board), SLA (Special Libraries Association), the NTL (National Transportation Library), and several transportation library consortia, to sponsor research, publication, digitization, bibliographic instruction, cataloguing, etc., to address, solve, create standards, etc., to these challenges. Among the initiatives currently under development we have: standardization of title page forms; maintenance and development of controlled vocabularies for better description; strengthening of archiving and distribution compliance channels; and digitization of historical (hard copy) reports. While not solely concerned with transportation technical reports, TRAIL, described below, is one example of these initiatives. Such efforts are aimed at partnering with transportation librarians, engineers and among these groups. However, it is important to note that most of these groups were established within the last 12 years and that the transportation library community has less than 250 information professionals in the US. Thus, there is still plenty of ground to cover and catching up to do.

Non-transportation librarians have also a role to play in solving these national challenges. The partnering of science and engineering librarians and liaisons with the producers of technical reports and the transportation information community may be very rewarding and exciting work. Besides getting to know/work with committed researchers and with a unique information community, librarians can also become consultants to these agencies/UTCs to

help them manage their information needs and publication of reports. Relatively easy library projects/tasks such as cataloguing, digitizing, website optimization, etc., can be done to disseminate and make the reports available to a wider audience. Partnerships with UTCs and state DOTs may also lead to broader collection development and the provision of (new) services to an underserved population. Finally, these partnerships may provide relatively easy public relations points for libraries and, yes, there is even a possibility of grant money being involved. It can be a true win-win arrangement.

The Format and Intellectual Property Issues of Technical Report Literature: Experiences from the Technical Report Archive and Image Library

What Is TRAIL?

The importance of preserving federal technical report literature was recognized by the Greater Western Library Alliance (GWLA) and the Center for Research Libraries (CRL) in early 2006. Initially formed as a task force and pilot project, the task force was charged to explore the viability of retrospective digitization of the technical report literature that had been distributed by federal agencies. The intent of the project was for preservation, while simultaneously creating better access to the report literature that was otherwise nearly impossible for the average researcher to find and access without mediation from a librarian. A little more than a year later, the Technical Report Archive & Image Library (TRAIL) was established with the goal of digitizing what was referred to as the “legacy” collections of technical reports issued by the federal agencies. “Legacy” was roughly defined as those reports that were issued prior to 1976 and publicly distributed by federal agencies.

So How Many Reports Are There?

One of the most common questions that TRAIL was asked to answer was “So how many reports will you need to digitize?” If anyone were to visit their closest academic library and ask to see their government report collection, aside from the

shelves of paper reports that may be available, most people would be floored to calculate just how many reports are housed in the room full of microfiche cabinets found in many libraries. To that end, doing some rough calculations based on the holdings at several major research universities, it is safe to estimate that there could be hundreds of thousands of reports just waiting to be revitalized in digital form.

Where Does One Start When Faced with Hundreds of Thousands of Reports?

In 2007, TRAIL formed a partnership with the University of Michigan’s Google Books project which enabled the scanning of many physical reports with deposit at the University of Michigan and eventually into HathiTrust. This partnership has been key in enabling the mass digitization of physical federal technical reports. As such, TRAIL identified the agencies and report series that were of highest interest to the sci-tech library community based on reports of patron activities and requests and later followed up this inventory with a study of technical report series that were (1) issued by agencies that are no longer in existence or (2) for which there is not evident stewardship being provided. It was also evident that libraries were interested in knowing that a complete run of a given series or agency’s materials had been digitized. Libraries were less enthused about a random digitization approach where everything was thrown into a big hopper, as there was no way of knowing what was included in this type of collection. As such, TRAIL agreed that it would, to the best of its ability, focus on ensuring that all available reports for a given series were retrospectively digitized, making the “completeness” of a collection a high priority in its approach. These initial environmental scans formed the basis of the collection development focus for TRAIL.

As TRAIL began digitizing materials, it was clear that while TRAIL would send a large number of reports through the UM Google Books digitization stream, arrangements would need to be made for the items that did not meet the scanning requirements set forth. These items mostly contained maps, foldouts, or were of such poor shape/quality that mass digitization would not be

possible. To this end, TRAIL partnered with the University of North Texas Libraries to coordinate the digitization and digital storage for these “special handling” materials.

In addition, TRAIL also came across a number of even more unique and challenging materials that were issued report numbers. An example of one such report included stereographic cards documenting geologic formations and included a fold-up viewer. Definitely not something most people digitize on a regular basis. As TRAIL came across these more challenging types of materials, they were cataloged, noted, and held aside. TRAIL also investigated harvesting reports from federal agencies where the agency had already performed digital conversion of the materials, but for which TRAIL felt it should replicated the holdings for purposes of preservation. As a pilot, TRAIL harvested the National Advisory Committee for Aeronautics (NACA) reports (NACA was the predecessor to NASA.). As of July 2012, TRAIL has digitized and/or harvested a little more than 40,000 reports, approximately 3.5 million pages of report content.

That’s Neat, but Is TRAIL Really Useful to Anybody?

The second most popular question after asking how many reports TRAIL was needing to digitize, is the question of “So is anybody really using this material?” Nearly as soon as TRAIL had posted its initial materials online, it began receiving comments and requests for more materials. Comments included:

“TRAIL is a wonderful digital library.”

“I can’t begin to express the gratitude and research potential your site has provided to us. Thanks to TRAIL, I am able to find and download Bureau of Mines Bulletins pertaining to my area that the Bureau of Mines and Office of Surface Mining in PA don’t even have in their libraries. Your site has made my research efforts a lot easier, saving weeks of time and travel expense.”

“Thank you so much for your help. The 25 year old copy of the monograph that I have

was destroyed by water; finding it online makes me very happy! Please pass on my great appreciation to all those who have put together this very valuable service.”

The comments and requests were global, including interactions with researchers as far away as Chile, Indonesia, Israel, Australia, and many more locations throughout Europe. For several years, usage of the materials was tracked primarily through the interactions TRAIL members had with customers. Statistical usage was difficult for TRAIL to track because of its multi-institutional repository structure, which was even further complicated by a search interface housed at a third institution. Despite these challenges, TRAIL recently enabled statistical tracking for both its main search interface—technicalreports.org and for the “special handling” collections housed at the University of North Texas.

Intellectual Property Issues

In addition to the more obvious complications that can occur with any retrospective digitization regarding format, the concept of intellectual property and copyright as it related to government documents became a very interesting area for TRAIL. There is great misconception that there is no copyright for government documents. And to further complicate matters the term “public domain” is often, unintentionally, confused with the concept of “no copyright.” These are two distinct concepts, with the common characteristic of being “free of copyright restrictions in the US.”

In brief, only those documents authored by an employee of the federal government actually have no copyright restriction, and that is only true within the US, as there can be legal restrictions abroad. In addition, it should go without saying that works authored by an employee of the federal government may contain copyrighted materials as part of the content of the work, which should be taken into consideration. For TRAIL, working closely with the government agencies to ensure the agencies were aware of our digitization and redistribution intent was very important in having confidence in pursuing digitization of the materials. The more

complicated areas involved reports produced by contractors for the government. If one reviews the legislative history pertaining to copyright, it can be documented that Congress decided that contractors and grantees were not employees of the government for purposes of copyright, and as such the resulting works were not considered “works of the US Government.” However, contractors and grantees were often contracted under some form of Federal Acquisitions Regulation, or, FAR. In general, the FAR indicates that the contractor or grantee retains the copyright ownership for the work, but also grants the Government “unlimited rights” that basically allows the Government to share and exercise all the same rights of a copyright owner. The Government sponsor or originating office is responsible for the review and acceptance of

materials on behalf of the Government, including the information quality, classification, and distribution. The sponsor or office is responsible for the primary and secondary distribution of the materials to the appropriate distribution channels. An example of a secondary distribution that impacts libraries would be the distribution of technical reports to the National Technical Information Service or to the Government Printing Office for distribution as part of the Federal Depository Library Program.

Next Steps

While TRAIL has been very conscious to take concerted efforts to notify agencies about TRAIL’s intent to digitize and redistribute the technical report literature, the only official way to ensure that there are no copyright restrictions on the digitized items is to enter into an agreement with each government agency explicating providing to TRAIL the right to digitize and copyright its works. To this end, TRAIL and the HathiTrust have begun conversations with the CENDI Copyright Working Group regarding how to obtain permission for digitization and redistribution from each government agency (CENDI is an interagency cooperative of the sci-tech information managers for each of the major government agencies.). To this end, the Copyright Working Group has put forth a proposal to CENDI to create a template agreement that could be utilized by each agency

to enter into an agreement with trusted mass digitization partners to enable the agency to more easily provide permissions to the partner for the digitization and redistribution of materials produced by the government, under the stewardship of the given agency. If obtained, this type of agreement would allow entities such as TRAIL the ability to further create and enhance access to the technical report literature.

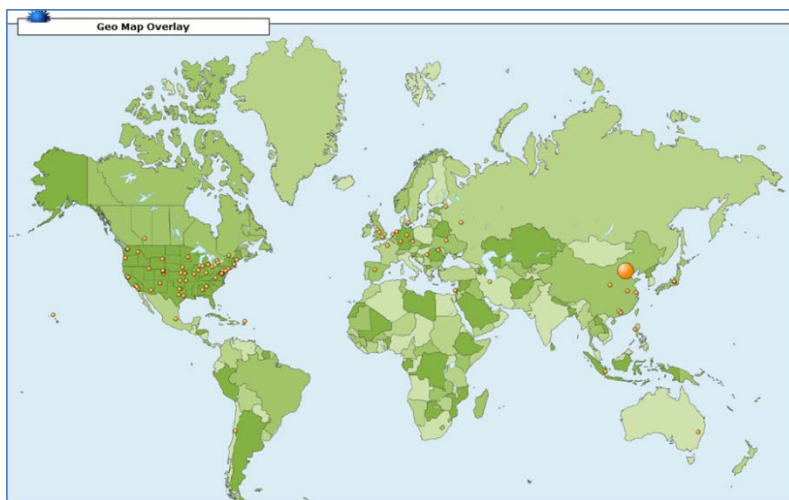


Figure 1. A Map Indicating the Originating Locations for Searches Performed on the Technicalreports.Org Website from January through October 2012

Technical Reports and Library Publishing Services

As described above, librarians have developed innovative strategies to engage with the preservation and access issues involved in managing technical report literature in the digital age. This section of the paper describes an experiment by Purdue University Libraries to improve the practices of technical report production, becoming involved much earlier in the scholarly communication supply chain, and working closely with administrators and authors at Indiana's UTC, the Joint Transportation Research Program (JTRP) at Purdue.¹ JTRP, established in 1936, makes an annual investment of around \$5 million, mostly drawn from the Indiana Department of Transportation (INDOT), in sponsoring around 270 students, faculty, and staff to do transportation research aimed at improving the efficiency of the Indiana transportation system.

The JTRP/Libraries project operates in the context of a broader investment in library publishing services at Purdue Libraries, where a "publishing division" was created in April 2012 to integrate the activities of Purdue University Press, founded in 1960, and a suite of scholarly publishing services developed by the libraries since 2006 under one umbrella. The vision that this project advances is one where Purdue Libraries is better able to serve a continuum or spectrum of publishing needs on campus, ranging from the production of traditionally "informal" publications such as conference proceedings or technical reports, to much more "formal" products such as books and journals aligned with the University's disciplinary strengths. While the publishing division of Purdue Libraries involves an almost unique integration of a university press with digital repository staff, its

¹ A more extensive description of the project at Purdue has been published in the *Transportation Research Record*. Newton, M. P., Bullock, D. M., Watkinson, C., Bracke, P. J., and D. K. Horton. (2012). Engaging new partners in transportation research: integrating the publishing, archiving, and indexing of technical literature into the research process. *Transportation Research Record: Journal of the Transportation Research Board*, 2291, 111–123. <http://dx.doi.org/10.3141/2291-13>

creation can be seen as part of a larger movement to build the capacity of library publishers. A 2010–2011 research project entitled "Library Publishing Services: Strategies for Success" sponsored by IMLS, SPARC, Berkeley Electronic Press, and Microsoft Research, found that over 55% of North American libraries were developing or offering publishing services, and this grew to almost 80% when ARL libraries alone were considered.² The recommendations of this project has recently led to the establishment of an initiative involving over 50 libraries to form a Library Publishing Coalition to increase the capacity of the library publishing field.

The Joint Transportation Research Program at Purdue was a logical partner for a pilot library publishing partnership, not only because of the volume of publications generated by its researchers, but also because of the clear vision of its director, Professor Darcy Bullock, of the importance of effective communication in minimizing duplication of effort and waste of taxpayer funding. Collaboration to digitize technical reports created since 1956 started in 2006, and over 1,500 reports are now available through Purdue e-Pubs (docs.lib.purdue.edu/jtrp), the Libraries' institutional repository and online publishing platform based on the Digital Commons platform. In 2009, a project to streamline the technical report publishing workflow, especially the complex review process, was embarked upon using the backend "EdiKit" system that has enabled Digital Commons to be used so successfully as a journals publishing solution. The Library therefore became an active partner not only in handling the back files, but also in managing the continuing publication of around 20 new reports a year.

As of 2012, the Libraries' publishing division and JTRP jointly sponsor a half-time production editor who is responsible for managing the review process, coordinating a largely automated copyediting and design process handled in XML by an external vendor, Charlesworth, performing quality assurance, and then assigning CrossRef DOIs to reports before they are published through

² The final report of this project can be found online at wp.sparc.arl.org/lps. Further information about the Library Publishing Coalition project can be found at <http://educopia.org/lpc>.

Purdue e-Pubs. Bibliographic standardization has been a major priority of the project, so the initial focus was based on fulfilling the guidelines for deposit set down by the *Transport Research International Documentation (TRID)* index, the largest online bibliographic database of transportation research. However, the opportunities for increased dissemination and impact that enriched metadata offer have driven further expansion of the partnership's indexing initiatives. All new reports are now assigned an ISBN, as well as a DOI, and made available in print-on-demand form through Ingram's Lightning Source system. With the assistance of bepress's DC Publishing Services suite, they are also submitted to library discovery platforms including OCLC WorldCat, Ex Libris Primo, ProQuest Summon, and EBSCO Discovery. The opportunity for scholars to obtain a copy of a new report in print through Amazon or Barnes & Noble, for example, as easily as they can retrieve a free PDF through Google Scholar has not only excited JTRP's funders, the Indiana Department of Transportation, but has fulfilled a real need that transportation scholars and practitioners have shown for their reports to be available in both print and electronic form. The range of publications has also expanded, with Purdue Libraries now collaborating with JTRP to publish the Proceedings of the Road School conference which annually attracts over 2,000 government officials, contractors, and engineers from all over the region.

The further investment that both Purdue Libraries and JTRP have made in their partnership to improve technical report publication has been repaid by impressive COUNTER compliant usage statistics that show over 450,000 downloads of the 1,500 reports since the project started in 2006. In 2012 alone, the reports were downloaded over 150,000 times, and there is a trend of ca. 10% increase per month. Moreover, at a time when universities are particularly concerned about

"global impact," Google Analytics shows that a substantial number of users were international, with particularly strong representation from India, China, and South Korea. The Analytics reports also provide encouraging information for transmittal to funders, showing INDOT a strong spread of users across Indiana. This has allowed University press releases to feature JTRP as a prime example of how Purdue delivers a good return on investment to Indiana's taxpayers by communicating applied research to the transportation engineers who can use it to improve local transportation infrastructure.

Conclusion

The production, dissemination, and preservation of technical reports, and other forms of "grey literature," have traditionally been seen as more of a problem than an opportunity for libraries and librarians. As the experiences at Northwestern University show, digital technologies have exacerbated, rather than lessened, the challenges of management and preservation in transportation, an important discipline that relies heavily on technical report production in North America where approximately \$180 million is spent annually by the US Department of Transportation in supporting State planning, research, and University Transportation Centers. However, the response from users and the technical report creators to library-based initiatives such as TRAIL and the Purdue University-JTRP partnership show that librarians and libraries are facing and overcoming these obstacles. These responses, measured both qualitatively and quantitatively, show that the investment by libraries and librarians can generate substantial good will, new and productive partnerships, and improve the scholarly impact of technical reports; thereby proving the benefits and outcomes outweigh the challenges when facing the vast array of materials in the technical literature ecosystem.