

Providing public access to grey literature at the National Transportation Library

Mary E. Moulton^{a*}

^a *National Transportation Library, U.S. Dept. of Transportation, Washington D.C.,
U.S.A*

*1200 New Jersey Avenue SE, Washington DC 20590 mary.moulton@dot.gov

Mary Moulton serves as Digital Librarian at the National Transportation Library. She leads the development and enhancement of digital repository services, information organization, digital curation, and discovery and use of transportation resources.

This manuscript was first submitted on July 7, 2020 for publication to the journal, *The Serials Librarian*, published by Taylor & Francis, for their "Special Issue: Grey Literature."

Providing public access to grey literature at the National Transportation Library

The National Transportation Library (NTL) at the United States Department of Transportation (USDOT) provides national and international access to the crucial transportation information that falls within the scope of grey literature, including the results of U.S. government funded research. Founded as an all-digital library in 1998, NTL's collections include full-text-born digital and digitized publications, data products, and other resources. All items are in the public domain and available for reuse without restriction. Since 2016, NTL has led the implementation of the USDOT's Official Public Access Plan issued in response to the February 22, 2013 Office of Science and Technology Policy (OSTP) Memorandum for the Heads of Executive Departments and Agencies entitled Increasing Access to the Results of Federally Funded Scientific Research. This paper discusses the effect this plan has had on a grey literature library and the efforts to create and maintain a public access repository, as well as exploring relationships between repository platform, contents, and people.

Keywords: public access, digital repositories, national libraries, transportation

Introduction

The National Transportation Library (NTL) at the U.S. Department of Transportation (USDOT) was established in 1998 by the Transportation Equity Act for the 21st Century (TEA- 21; P.L. 105-178) and is administered by the USDOT Bureau of Transportation Statistics (BTS). NTL's fundamental purpose is to preserve and serve transportation grey literature in the form of statistical publications, data products, and all USDOT funded transportation research reports. Additionally, NTL offers data services, reference and research services, and support for the Nation's transportation knowledge networks. NTL's authorized role was expanded in the 2012 Moving Ahead for Progress in the 21st Century (MAP-21; P.L. 112- 141). Specific mandates for NTL in section 6304 include the following:

- Serve USDOT, other Federal agencies, public and private

organizations, and individuals within the United States and internationally

- Acquire, preserve, and manage transportation information and information products and services for use by USDOT, other Federal agencies, and the public
- Provide reference and research services
- Serve as the central repository for USDOT research results and technical publications
- Act as a central clearinghouse for transportation data and information of the Federal Government
- Serve as coordinator and policy lead for transportation information access
- Coordinate among and cooperate with multiple external parties to develop a “comprehensive transportation information and knowledge network” to support the I- XI list [information required of BTS at 49 USC 6302((b)(3)(B)(vi)]
- Publicize, facilitate, and promote access to information products and services
- Receive contract, grant, and agreement authority to support these goals

NTL and the Public Access Plan

Building on its legislative mandate regarding grey literature produced and funded by USDOT, since 2016 NTL has had a leading role in implementing the USDOT Public Access Plan.¹ This plan was issued in response to the February 22, 2013, Office of Science and Technology Policy (OSTP) memorandum to the heads of executive departments and agencies entitled *Increasing Access to the Results of Federally*

*Funded Scientific Research.*² The memorandum requires all executive departments and agencies spending more than \$100 million/year on R&D to ensure public access to peer-reviewed publications and digital datasets arising from federally funded scientific research. As a result of this OSTP memorandum and the public access plan it spawned, there is now a greater emphasis on acquiring grey literature research outputs by USDOT for NTL to manage, and a greater responsibility to make those research outputs discoverable and accessible to users.

USDOT's Public Access plan requires external authors and the agencies within USDOT to submit to NTL's digital repository all publications and datasets that meet the plan's criteria. Publications are defined as any written scientific research project deliverable that is funded, fully or partially, through a USDOT-managed contract, grant, or other funding agreement. Publications include technical or final research reports, as well as any final peer-reviewed manuscripts accepted for publication. To protect intellectual property, researchers may request a 12-month embargo period for a journal manuscript. Increasingly, grey literature research outputs are accompanied by supporting datasets, data management plans, software, and readme files. These datasets and supplementary files are as valuable a component of the research output being preserved by NTL as the written publications, raising compelling questions about the current boundaries of grey literature.

Datasets are defined as all scientific data collected through research projects funded, either fully or partially, by Federal funds awarded through a USDOT contract, grant, or other agreement or collected by USDOT employees. Data may be numeric or qualitative, images or video, text sequences, samples, physical collections, software, algorithms, and models. Data can be instrument-generated digital data output, researcher-generated databases, data transformations, data linkages, tables,

spreadsheets, or observational data.

Data Management Plans and digital data preservation

In addition to research results, all USDOT-funded research proposals, intramural and extramural, require submission of a Data Management Plan (DMP). A DMP is essential to digital data preservation. It describes how researchers will handle digital data both during and after a research project. A DMP indicates how the research proposal conforms to USDOT policy on the dissemination and sharing of research results. In certain instances, it may also discuss why long-term preservation or public access cannot be justified. These DMPs are preserved alongside the datasets and the written publications within the NTL digital repository, as mandated by USDOT's Public Access Plan.

Similarly, USDOT's Public Access Plan requires all researchers to obtain and report their unique ORCID (Open Researcher and Contributor ID). ORCID provides a unique, persistent digital identifier that distinguishes one researcher from another. USDOT integrates ORCIDs into workflows such as submissions of grants and research results, and includes them in metadata for digital objects in NTL's Repository and Open Science Access Portal, *ROSA P* (<https://rosap.ntl.bts.gov/>).³

***ROSA P* — an open access repository and digital archive**

Named in honor of Rosa Parks and the role of public transportation in the 20th Century American civil rights movement, *ROSA P* was developed in large part as a response to the directives of making grey literature discoverable and accessible as required by the USDOT Public Access Plan. *ROSA P* is an open access repository and full-text digital archive for all modes of transportation and related disciplines. All content in *ROSA P*

is grey literature, in the public domain, and available for reuse without restriction. *ROSA P* contains approximately 47,000 digital objects. Content types include technical reports, datasets, research papers, proceedings, briefs, manuals, and presentations. The collection is comprised of born digital and digitized copies of legacy print materials. *ROSA P* is built on an open source Fedora platform, developed, and maintained by the U.S. Centers for Disease Control. All content is stored in the cloud and backed up in three different geographic locations. Science.gov, Google, Google Scholar, and Bing index and search *ROSA P*. Metadata is regularly harvested by OCLC's Worldcat Digital Collection Gateway. NTL also coordinates indexing of *ROSA P* with TRID (Transport Research International Documentation), the world's largest and most comprehensive bibliographic resource on transportation research information, produced and maintained by the Transportation Research Board of the U.S. National Academies.

NTL migrated content from its existing repository to *ROSA P* in 2017. The transition included review and revision of metadata to accommodate public access documents and linked open data. The public launch of *ROSA P* occurred in January 2018. The migration exposed out of scope and inaccessible documents within this collection of grey literature. The original repository was built with batch uploads that were not scrutinized for compliance with NTL policies. These batches contained duplicates, copyrighted material, incomplete and poor quality scans of print materials, erroneous or missing metadata, and records linking to HTML pages. The original repository contained approximately 55,000 records; *ROSA P* was publicly launched with 32,000 items. The remaining items and records were held back for review. Subsequently, NTL revised its Collection Development and Maintenance Policy, and began a remediation process for problem metadata and content.⁴ NTL also initiated monthly

meetings with satellite USDOT library staff to collaborate on remediation of their collections in *ROSA P*. As of June 2020, more than 6,000 catalog records in *ROSA P* have been modified in some way.

Efforts to improve content quality and collection accessibility can be measured in repository page views. During the month of November 2017, prior to the public launch of *ROSA P*, there were 9,922 page views in the existing repository. There were 35,216 page views in *ROSA P* during November 2018 and 51,368 page views during November 2019.

NTL revised work processes

Implementation of USDOT's Public Access Plan and the transition to a new repository also required revision of NTL work processes for managing grey literature. In 2017, NTL initiated review and documentation of digital submissions, ingest, and cataloging work processes. Content creators or owners currently submit digital objects to NTL by email. NTL recently devised an electronic submissions form that enables capture of required metadata and information about submitters. The electronic submissions form is scheduled to be implemented by the end of 2020. The new process should reduce cataloging errors, increase efficiency, and provide compliance data for USDOT's Public Access Plan. Additional work process changes include application of DataCite digital object identifiers (DOI) to all USDOT publications and datasets, checking documents for accessibility, and improved metrics reporting.

To ensure ongoing availability to users as a public access repository and a long-term archive, NTL strives to ensure its digital collections are in non-proprietary open formats, wherever possible. If a document is no longer accessible, NTL migrates it to a new format to preserve the knowledge it contains and records the change in a public

note. NTL also follows the final Federal rule (83 FR 2912) for the Information and Communication Technology (ICT) Refresh under Section 508 of the Americans with Disabilities Act (ADA).⁵ According to the updated rule, all public facing content submissions published on or after March 23, 2018, must be accessible to users covered by ADA. NTL checks all submitted content for compliance with the ICT Refresh and works with submitters when remediation is necessary. NTL also monitors downloads from *ROSA P* and remediates the 25 most popular documents, to ensure that they are fully accessible to all. Remaining legacy content is made accessible and remediated at the request users.

Scope of NTL digital special collections

NTL's original digital repository was built in collaboration with USDOT satellite libraries at the Volpe National Transportation Systems Center and Federal Highway Administration, and with agreements from the USDOT operating administrations Intelligent Transportation Systems Joint Project Office and National Highway Traffic Safety Administration. These partner organizations supplied digital content and appropriately named collections were established within the repository. Before migrating content from the existing repository, collections for remaining operating administrations were created in *ROSA P* with the goal of enabling transparency and ensuring compliance with USDOT's Public Access Plan. As of June 2020, 7 of 10 USDOT operating administrations are represented by live collections in *ROSA P*. NTL is conducting an aggressive outreach program to ensure that by June 2021, all USDOT operating administrations have live collections in *ROSA P*.

Legacy print materials that have been digitized are collected if they have historic, technical, or national significance. Examples of digital special collections in *ROSA P*

include investigations of aircraft and railroad accidents supplied by the National Transportation Safety Board. “Investigations of Aircraft Accidents 1934-1965” contains digitized copies of U.S. Civil Aeronautics Board reports, searchable by airline name, incident site and date. “Investigations of Railroad Accidents 1911-1993” contains digitized copies of U.S. Interstate Commerce Commission railroad accident reports, searchable by incident site and date. In 2014, NTL collaborated with the U.S. Federal Transit Administration (FTA) on a digitization project to celebrate the 50th anniversary of the Urban Mass Transit Act. Publicly accessible in *ROSA P*, the “FTA 50th Anniversary Document Collection” features approximately 500 reports, memoranda, letters, images and other documents highlighting the history and achievements of the submitting agency.

NTL partners with and accepts content submissions for *ROSA P* from state, metropolitan, local, and tribal agencies; other Federal agencies; University Transportation Centers; professional organizations; and private sector sources. With the increased focus on ensuring its collection of grey literature is preserved and accessible, NTL has proactively engaged partners in the transportation field to attain this goal. Most of NTL’s outreach for these submissions occurs through the National Transportation Knowledge Network (NTKN), a collaborative alliance of volunteer knowledge and information professionals, coordinated by NTL.⁶ NTKN hosts a monthly web conference series, the Transportation Librarians Roundtable (TLR), that provides transportation information professionals with opportunities to learn about the USDOT Public Access Plan and important aspects of it, such as document accessibility, intellectual property management, application of digital object identifiers, and introduction to data management plans.⁷ The NTKN Cooperative Digitization Working Group documents best practices for managing digitization

projects and assembles shared inventories to eliminate duplication of effort.

Conclusions

With government increasingly interested and invested in making the grey literature research outputs it creates and funds discoverable and available to users now and in the future, new thought and effort is being placed in grey literature—what it is and how we acquire, manage, preserve, and serve it. The work being undertaken by the National Transportation Library and by USDOT in general is a clear indication of the value placed on this type of content. Despite the name grey literature, it cannot be treated as something in the shadow of other types of scholarly communication, but as an equal and important strand of the network of scholarly communication.

Notes

1. U.S. Department of Transportation (2015) Plan to Increase Public Access to the Results of Federally Funded Scientific Research Results. Retrieved June 16, 2020, from: <https://rosap.ntl.bts.gov/view/dot/29637>
2. Office of Science and Technology Policy (2013) Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research. Retrieved June 16, 2020, from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
3. U.S. Department of Transportation (2020) ROSA P. Retrieved June 16, 2020, from: <https://rosap.ntl.bts.gov/>
4. U.S. Department of Transportation (2018) National Transportation Library

Collection Development and Maintenance Policy, version 1.3. Retrieved June 16, 2020, from:

<https://ntl.bts.gov/sites/bts.dot.gov/files/docs/ntl/219951/ntlcollectiondevelopmentpolicy132018-01-22.pdf>

5. U.S. Access Board (2018) Corrections to the Information and Communication Technology (ICT) Final Standards and Guidelines. Retrieved June 16, 2020, from:
<https://www.access-board.gov/guidelines-and-standards/communications-and-it/about-the-ict-refresh/corrections-to-the-ict-final-rule>

6. U.S. Department of Transportation (2020) National Transportation Knowledge Network (NTKN). Retrieved June 16, 2020, from:
<https://transportation.libguides.com/ntkn>

7. U.S. Department of Transportation (2020) Transportation Librarians Roundtable. Retrieved June 16, 2020, from: <https://transportation.libguides.com/TLR>

