Visualizing National Transportation Library Repository Metrics

Aileen Nolan, Data Visualization Fellow, National Transportation Library <aileen.nolan.ctr@dot.gov>

Mary Moulton, Digital Librarian, National Transportation Library <mary.moulton@dot.gov >

rosap.ntl.bts.gov

WHAT IS ROSA P?

The Repository and Open Science Access Portal or ROSA P is the designated institutional repository for research funded by the US Department of Transportation under the USDOT Public Access Plan. It includes full-text electronic publications, datasets, and other resources that are provided freely to transportation researchers, statistical organizations, the media, and the general public.

ROSA P is maintained by the National Transportation Library (NTL), an all-digital federal library servicing the Bureau of Transportation Statistics within the US Department of Transportation. NTL is mandated to:

- Provide national and international access to transportation information
- Coordinate information creation and dissemination
- Offer reference services for the transportation community

ROSA P can be accessed at https://rosap.ntl.bts.gov/.

BACKGROUND

Institutional Repositories (IR) have been a core innovation for digital libraries and institutions involved in disseminating research. IRs provide a way to collect, manage, and preserve research outputs in a manner that is more efficient than traditional methods involving print and microfiche.

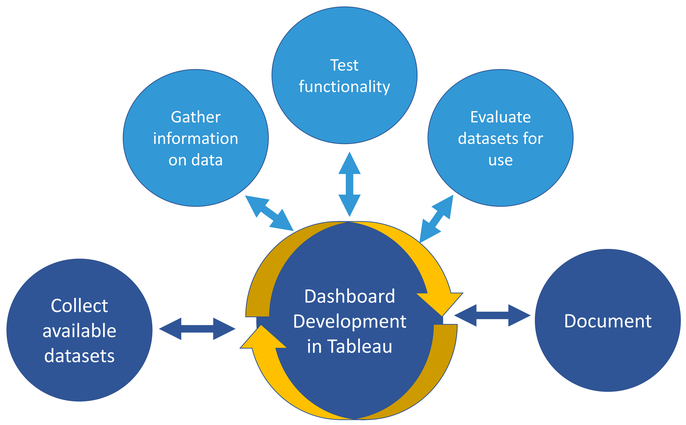
As a national federal library that has been designated as the central repository for USDOT research and a clearinghouse for Government transportation data, NTL has special interest in ensuring its authority and integrity and has committed to creating a trustworthy repository. Beginning in 2016, in the wake of a migration project to move all records to ROSA P, NTL undertook a self-assessment to evaluate the repository according to the criteria established CoreTrustSeal certification, and by extension, ISO 16363: Space data and information transfer systems -- Audit and certification of trustworthy digital repositories. The goal of the project was to identify opportunities to improve and prepare to eventually pursue CoreTrustSeal certification.

NTL is interested in visualizing repository metrics to promote communication between internal stakeholders, experts, and their community of users. By creating a mechanism for guidance and feedback, visualization also supports CoreTrustSeal criteria for trustworthy repositories.

In this project, NTL created a dashboard using Tableau software. Among several, one of the main goals was to establish a baseline on which NTL can evaluate its services. In the long run, this baseline can serve as a way to provide accountability to the public and evidence of value, progress, a tool to identify areas for improvement to internal and external stakeholders .

OBJECTIVE Create a suite of dashboards using data visualization techniques and multiple data sources to track metrics for assessing the institutional repository, ROSA P.

METHOD



CHALLENGES

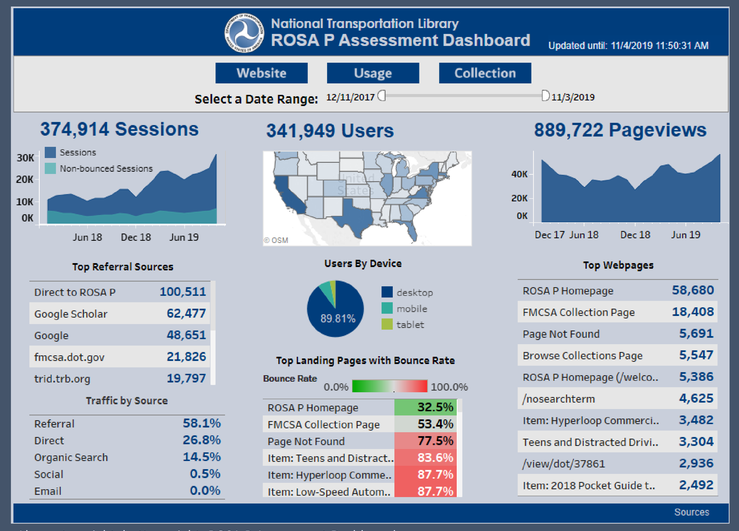
- Issue 1: Data discrepancies
- Issue 2: Accessing raw data
- Issue 3: Decisions on most robust and accurate data

RESULTS

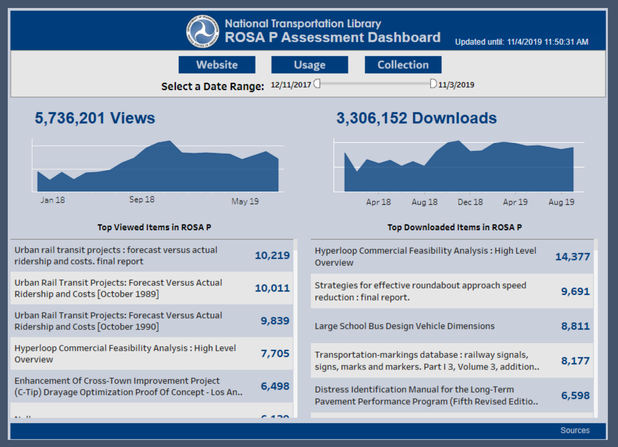
A number of dashboards were created as part of this objective to build a suite of dashboards. Included in this are dashboards on the NTL and ROSA P websites, collections usage, reference questions, and others.

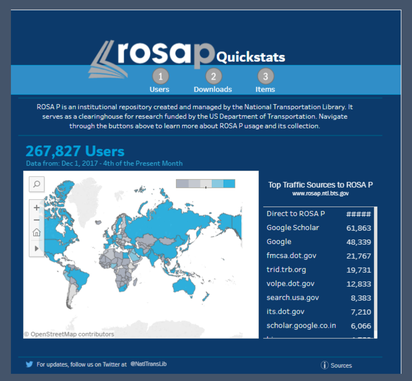
The three biggest images below show the ROSA P Assessment Dashboard. This dashboard combines data from more than five sources including Google Analytics reports, data extracted from the servers that host the ROSA P system, and data exported from NTL's metadata management system. This past year, this dashboard was redesigned to better facilitate data exploration and to improve the experience of getting to the most meaningful information quickly. That redesign included adding in more in-depth information on website traffic, long-term trends over time, and dynamic filters to search the items in ROSA P across many dimensions. It also inspired improvement of the visual design of all the dashboards being created in this suite.

Below the ROSA P Assessment Dashboard images are two examples of other dashboards created for NTL. One is a spin off of the original ROSA P Assessment Dashboard specifically tailored for public usage. This dashboard features a pared down version of the data and a simpler visual design. The other example is an internal dashboard made specifically for other USDOT Operating Administrations (OA) that upload items to ROSA P. These OAs can look at this dashboard to view the status and performance of their collections without having to do any data prep themselves.

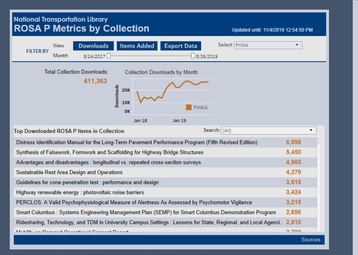


Above, top-right, bottom-right: ROSA P Assessment Dashboard

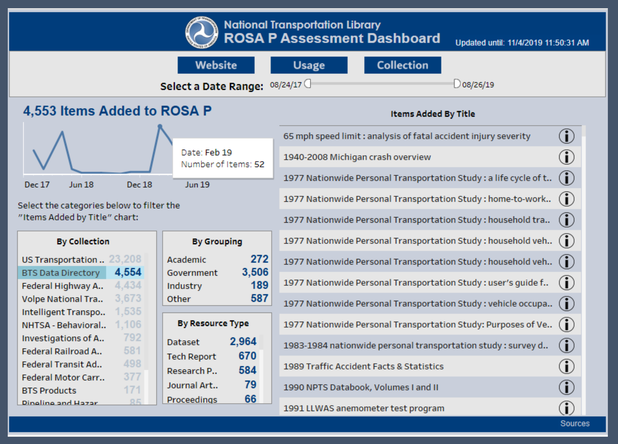




Above: ROSA P Assessment Dashboard for Public Use



Above: OA Collection Dashboard



FUTURE CONSIDERATIONS

Going forward these dashboards will continue to be improved and shared internally. Some dashboards may be shared with the public through embedding into NTL's homepage or the ROSA P website.
Other opportunities to improve include replacing current data sources with data that is extracted from the ROSA P API as JSON files. More improvements could come from finding innovative solutions to automation. Currently we are looking into using the Selenium Python package to access web-based data sources.

ABSTRACT

The Repository and Open Science Access Portal (ROSA P) is the designated institutional repository for research funded by the U.S. Department of Transportation under the USDOT Public Access Plan. Managed by the National Transportation Library (NTL), this repository plays a crucial role in the collection, management, and preservation of research outputs. As part of NTL's efforts to establish ROSA P as a Trustworthy Repository, this project focused on implementing data visualization software to create dashboards to consolidate metrics on users and usage, including key metrics such as aggregated user location, download counts, and item resource type. These visualizations provide a mechanism that is used for internal and external library assessment.

RECOMMENDED CITATION

Nolan, Aileen https://orcid.org/0000-0002-1495-9226 and Mary E. Moulton https://orcid.org/0000-0002-1791-068X. 2019. "Visualizing National Transportation Library Repository Metrics." Transportation Research Board 99th Annual Meeting. Washington, D.C., USA. https://doi.org/10.21949/1506104