Assessing the National Transportation Library's *ROSA P* Using Data Visualization

Aileen Nolan, USDOT Intern and MLIS Candidate 2018, University of Illinois at Urbana-Champaign School of Library & Information Science aileen.nolan@gmail.com

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

U.S. Department of Transportation
Office of the Secretary of Transportation

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 .

rosap
Repository & Open Science
Access Portal

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.
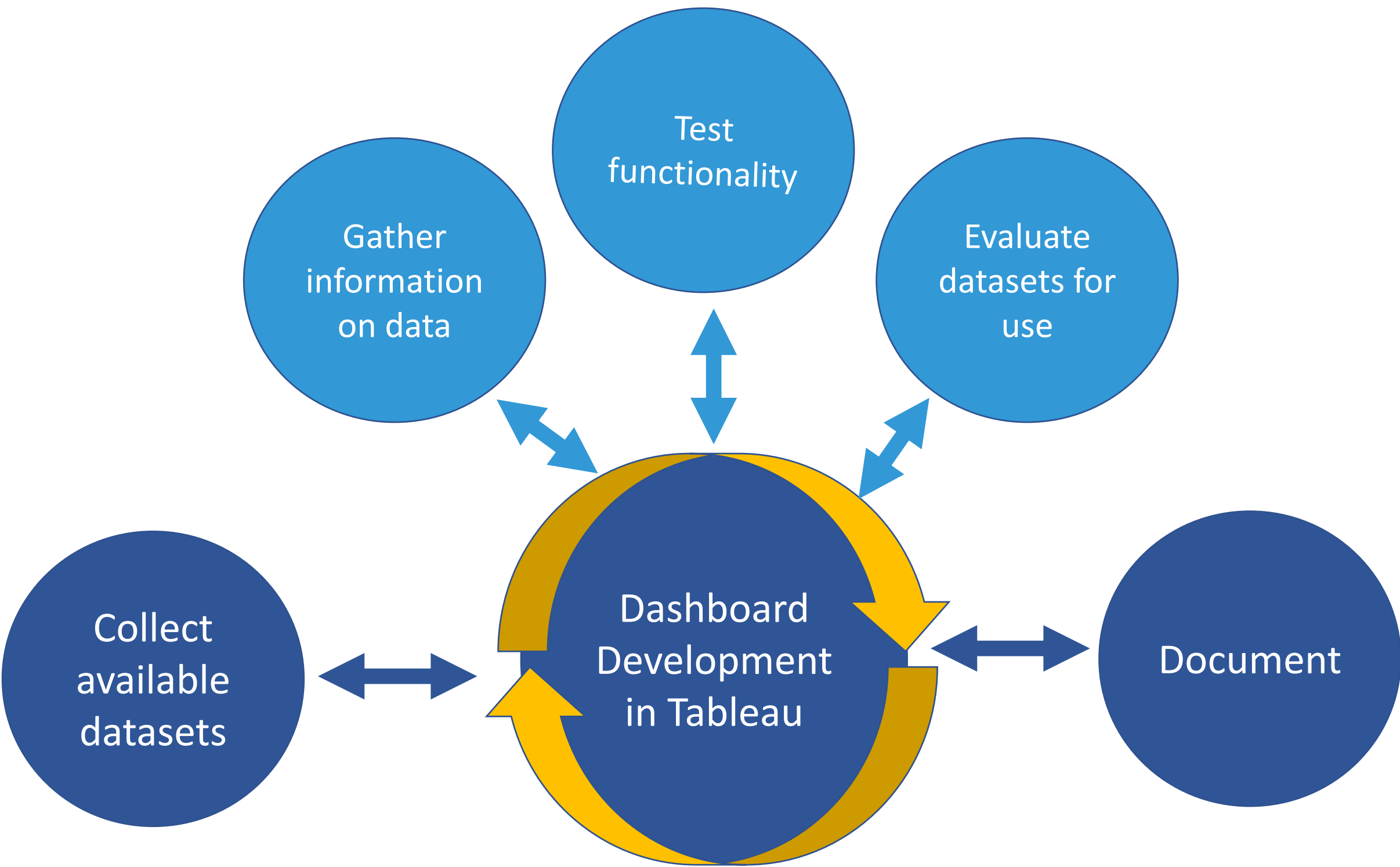
As a national federal library that has been designated as the central repository for USDOT research and a clearinghouse for Government transportation data, NTL has special interest in ensuring its authority and integrity and has committed to creating a trustworthy repository. Beginning in 2016, in the 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 dashboard using data visualization techniques and multiple data sources to track metrics for assessing the institutional repository, ROSA P.

METHOD

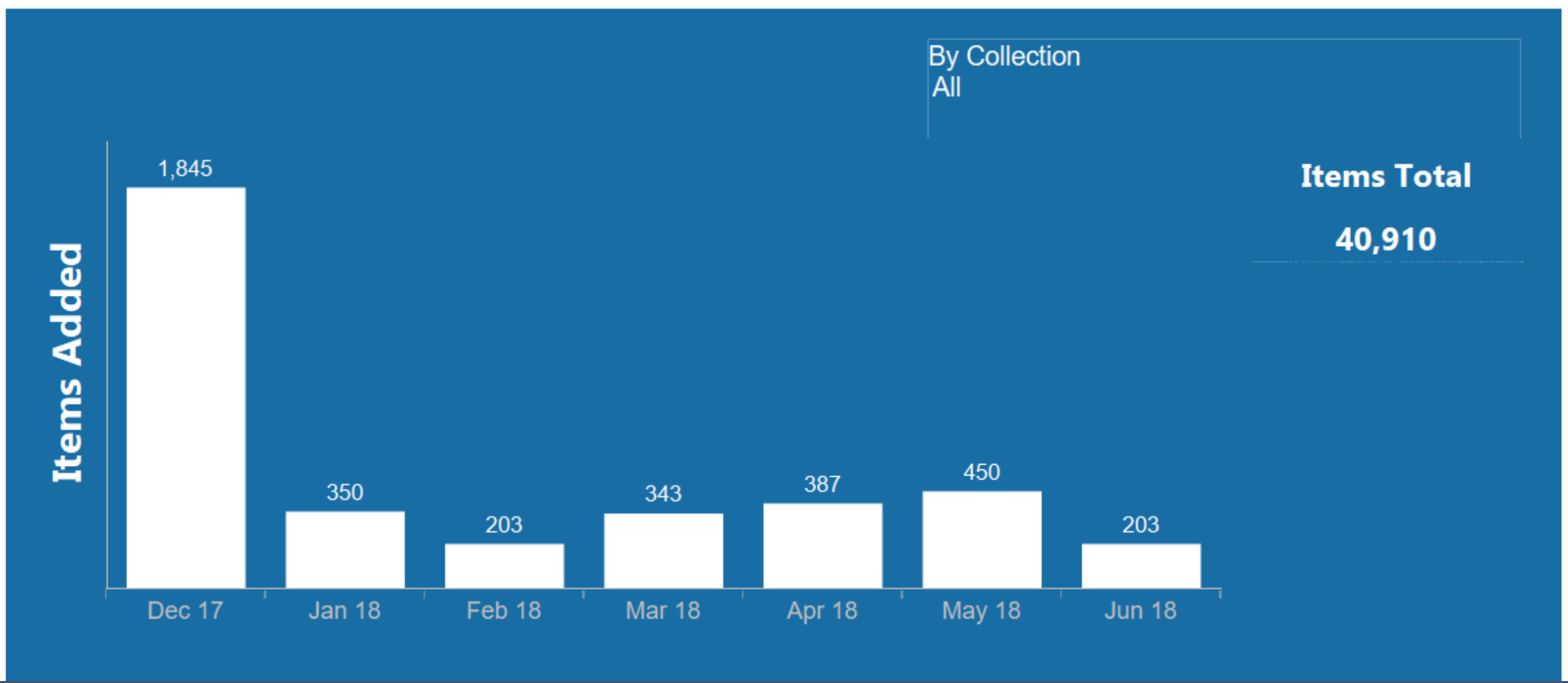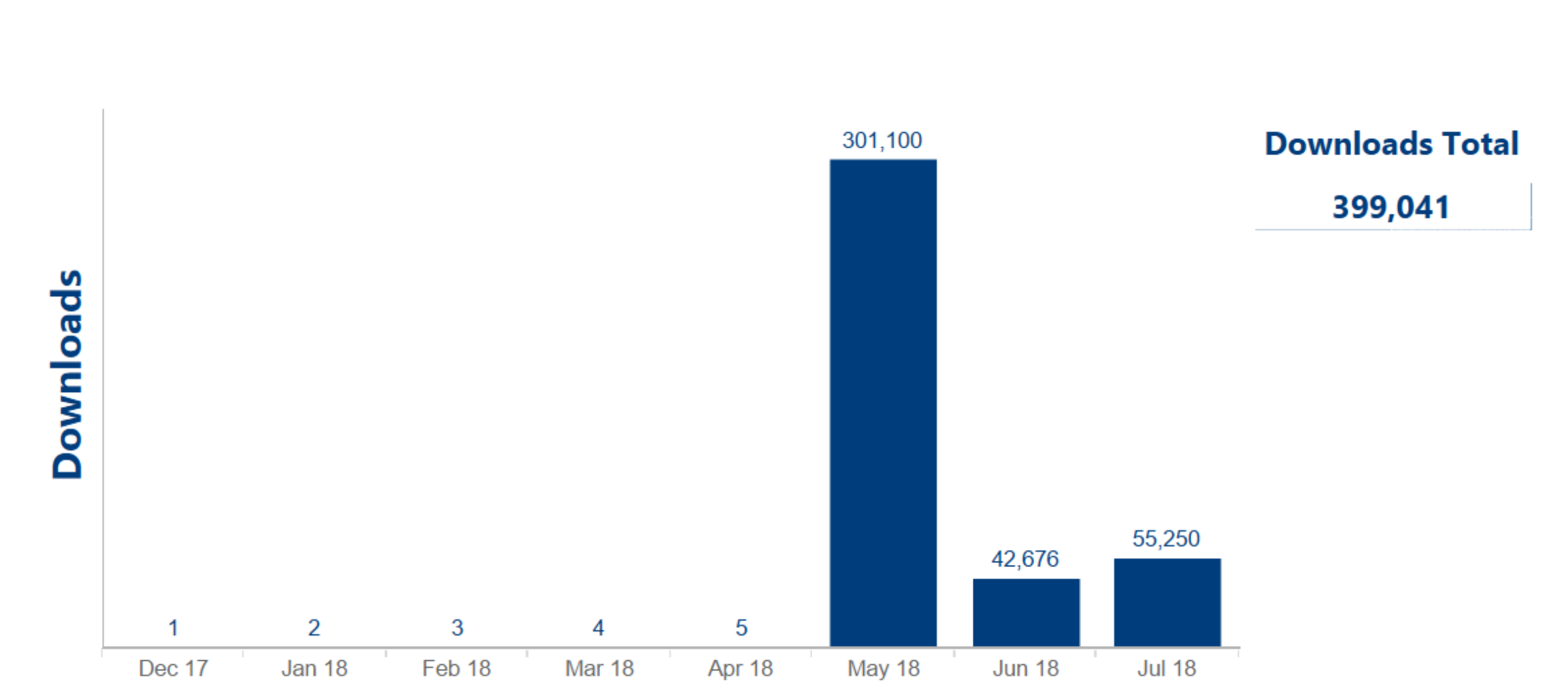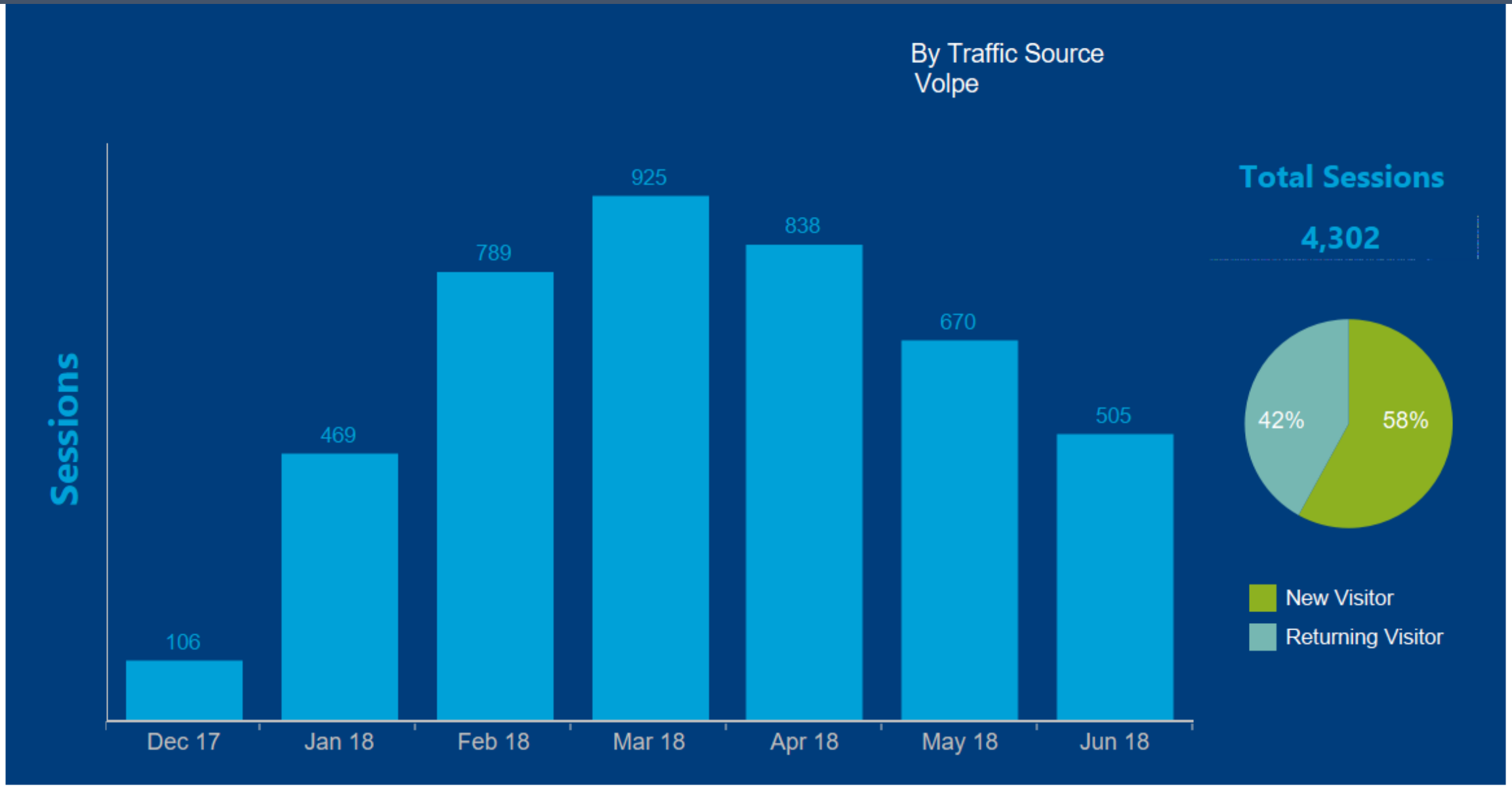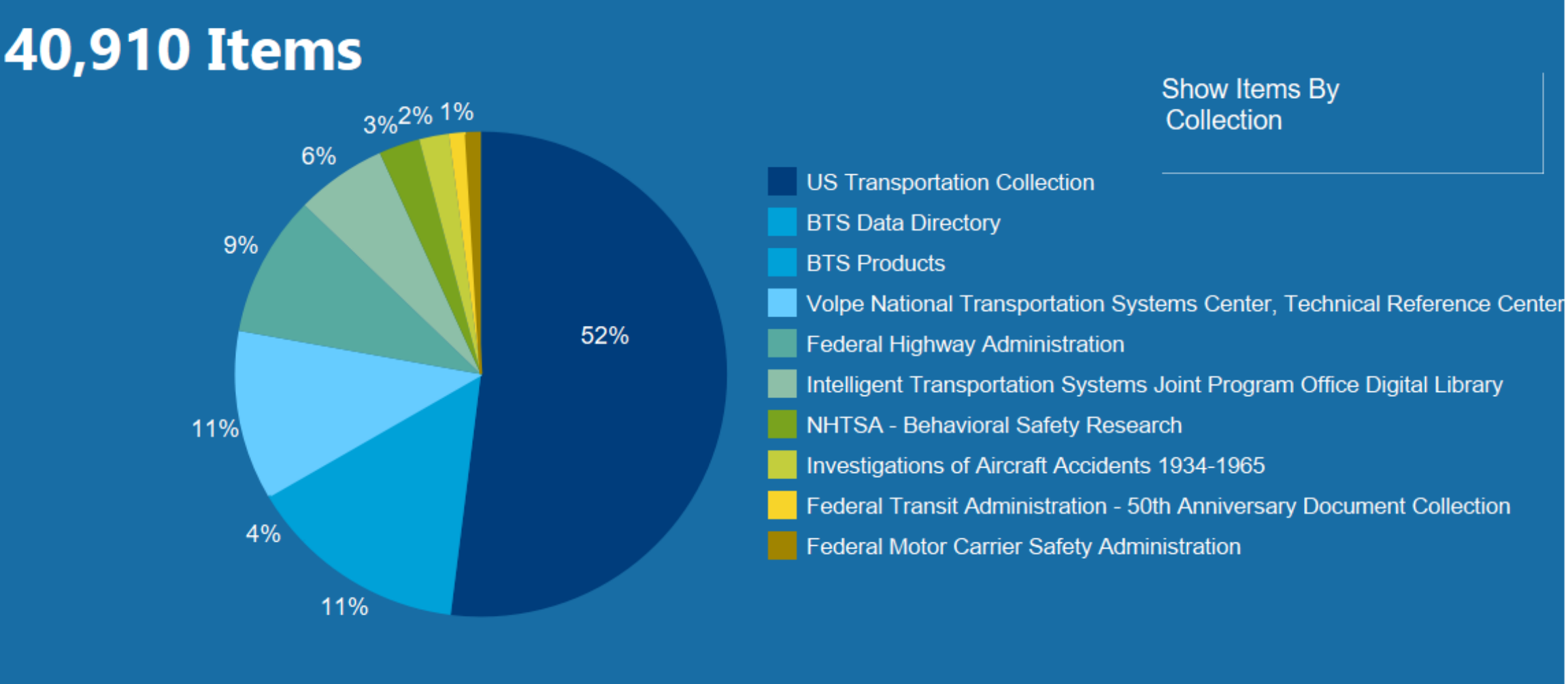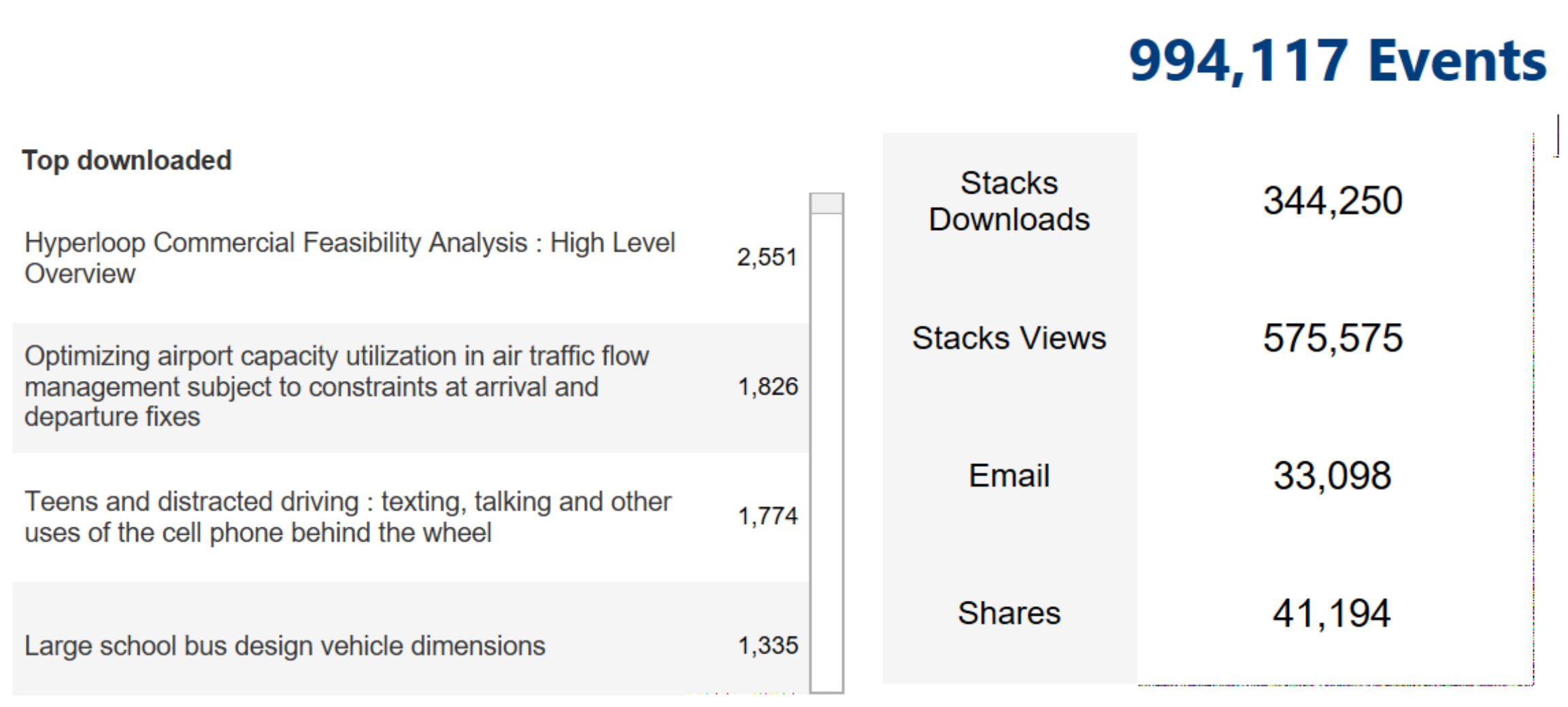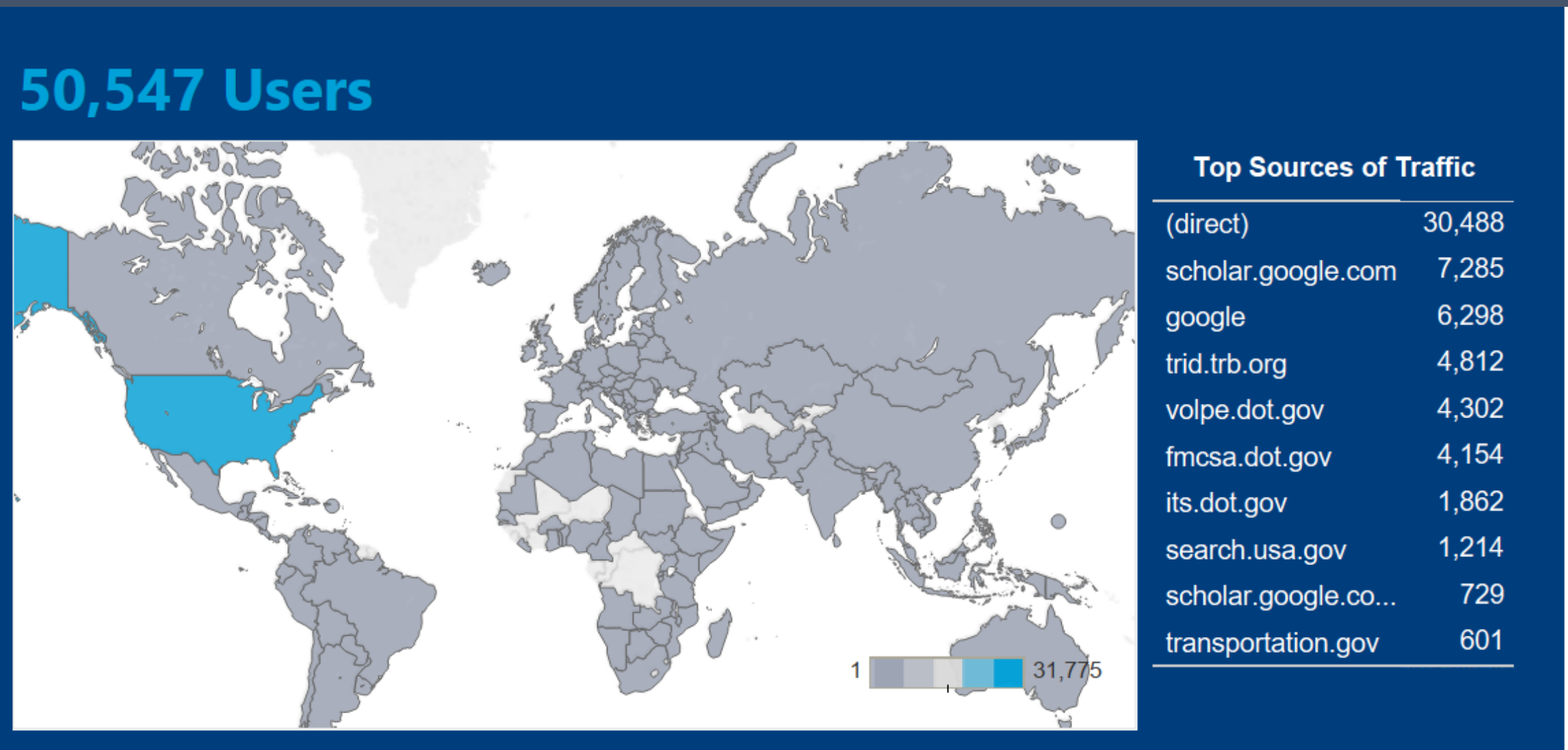### CHALLENGES

Issue 1: Data discrepancies

Issue 2: Accessing raw data

Issue 3: Decisions on most robust and accurate data

RESULTS

The ROSA P Assessment Dashboard employs five data sources. This data includes Google Analytics reports, data extracted from the servers that host the ROSA P system, and data exported from NTL's metadata management system.

The dashboard is organized to communicate metrics based on three areas of interest: Users, Events, and Items. The Users section reports on how many users are visiting the ROSA P website, where they come from, and the website that referred them. The Events section reports on the usage of items in ROSA P. This includes information about the total downloads, shares, emails, and views of the top 100 downloaded items. Finally, the Items section considers all of the resources hosted by ROSA P and visualizes them as a whole using a pie chart. The pie chart can then be broken down by Collection, Resource Type, and Corp Creator. This provides information on what types of items dominate the collection.

The dashboard is comprised of two worksheets, Snapshot and Metrics by Month. Snapshot is intended to serve the broadest possible audience, including the public. It provides a baseline on which to understand what is in our collection and how users are interacting with it. Metrics by Month breaks the interest areas into further detail, zeroing in on measurements broken down by month. This worksheet is intended to serve internal and external BTS and NTL stakeholders by providing one type of baseline on which to evaluate NTL services.

FUTURE CONSIDERATIONS

Going forward this dashboard may be shared with the public through embedding into NTL's homepage or the ROSA P website. Before that happens however, problems with verifying the validity of the data need to be addressed.

Other opportunities to improve the Dashboard include replacing current data sources with data that is extracted from the ROSA P API as JSON files. Additionally, to improve the process of maintaining the dashboard Tableau Server could be implemented. This would involve troubleshooting access problems caused by internal firewalls.

RECOMMENDED CITATION

Nolan, Aileen https://orcid.org/0000-0002-1495-9226 and Mary E. Moulton https://orcid.org/0000-0002-1791-068X. 2019. "Assessing the National Transportation Library's *ROSA P* Using Data Visualization." Transportation Research Board 98th Annual Meeting. Washington, D.C., USA. https://doi.org/10.21949/1503232

50,547 Users

Top Sources of Traffic

(direct)	30,488
scholar.google.com	7,285
google	6,298
trid.trb.org	4,812
volpe.dot.gov	4,302
fmcsa.dot.gov	4,154
its.dot.gov	1,862
search.usa.gov	1,214
scholar.google.co...	729
transportation.gov	601

994,117 Events

Top downloaded

Hyperloop Commercial Feasibility Analysis : High Level Overview	2,551
Optimizing airport capacity utilization in air traffic flow management subject to constraints at arrival and departure fixes	1,826
Teens and distracted driving : texting, talking and other uses of the cell phone behind the wheel	1,774
Large school bus design vehicle dimensions	1,335

Stacks Downloads	344,250
Stacks Views	575,575
Email	33,098
Shares	41,194

40,910 Items

Show Items By Collection

- US Transportation Collection
- BTS Data Directory
- BTS Products
- Volpe National Transportation Systems Center, Technical Reference Center
- Federal Highway Administration
- Intelligent Transportation Systems Joint Program Office Digital Library
- NHTSA - Behavioral Safety Research
- Investigations of Aircraft Accidents 1934-1965
- Federal Transit Administration - 50th Anniversary Document Collection
- Federal Motor Carrier Safety Administration

By Traffic Source
Volpe

Total Sessions 4,302

New Visitor / Returning Visitor (42% / 58%)

Downloads Total 399,041

By Collection
All

Items Total 40,910