



# Introducing *ROSA-P*: The National Transportation Library's Repository and Open Science Access Portal

Mary Moulton, Digital Librarian 0000-0002-1791-068X

Leighton Christiansen, Data Curator 0000-0002-0543-4268

Laura Farley, NTL Fellow 0000-0002-9980-6850

Xin Wang, Systems Librarian

US DOT National Transportation Library

Federal Depository Library Conference

Arlington VA

October 17, 2017

# *ROSA-P* Repository and Open Science Access Portal

<https://rosap.ntl.bts.gov/>

---



By [http://www.loc.gov/rr/print/list/083\\_afr.html#ParksR](http://www.loc.gov/rr/print/list/083_afr.html#ParksR),  
Fair use, <https://en.wikipedia.org/w/index.php?curid=3034067>



U.S. Department of Transportation

Office of the Assistant Secretary for  
Research and Technology

# About NTL

---

Established in 1998, we provide access to:

- Digital collections
- Data services
- Reference and research services
- Networking

We are an **open access** digital repository.

All items are in the **public domain** and available for reuse **without restriction.**

# NTL Mandates

---

- **Transportation Equity Act for the 21st Century (1998)**
  - “establish and maintain a National Transportation Library, which shall contain a collection of statistical and other information needed for transportation decision making at the Federal, State, and local levels.”
- **MAP-21 (2012)**
  - Acquire, preserve and manage transportation information and information products and services for use by DOT, other Federal agencies, and the public
  - Central repository for DOT research results and technical publications
  - Central clearinghouse for transportation data and information of the Federal Government
- **White House Office of Science and Technology Policy memo (2013)** requiring 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

# NTL Digital Repository and Services

---

- Digital Repository
  - Purpose: serve as central clearinghouse and archive for public access transportation information
  - 32,000+ items
  - Special features: access to specialized collections
    - Volpe Reports
    - Directory of Transportation Data Sources
    - NHTSA Behavioral Safety
    - ITS JPO Electronic Library
  - Services: digitization activities, preservation & ensuring perpetual access; permanent URLs, links incorporated into TRID records



# NTL Digital Repository Migration

---

## Current

- Homegrown OAI-PMH compliant repository (Open Archives Initiative Protocol for Metadata Harvesting)
- In-house hardware
- Modified Dublin Core metadata
- Indexed by popular search services Google and Bing
- Self governance using OAIS model

## *ROSA-P*

- Open source platform
  - CDC Stacks platform (Islandora/Drupal/Fedora)
- Cloud storage – stable and responsive
- Revised metadata accommodates public access documents and linked open data
- Additional indexing by Google Scholar
- DOIs minted for every report and dataset in NTL repository – persistent, uniquely identifies each object
- Compliance with core trustworthy repository requirements (ISO 16363) via self-audit and peer review

# National Transportation Library

Home Collections Recent Additions Public Access Submit Content

About ROSA P



All Collections

Enter keyword or phrase...

Search

Advanced Search

## Welcome to ROSA P

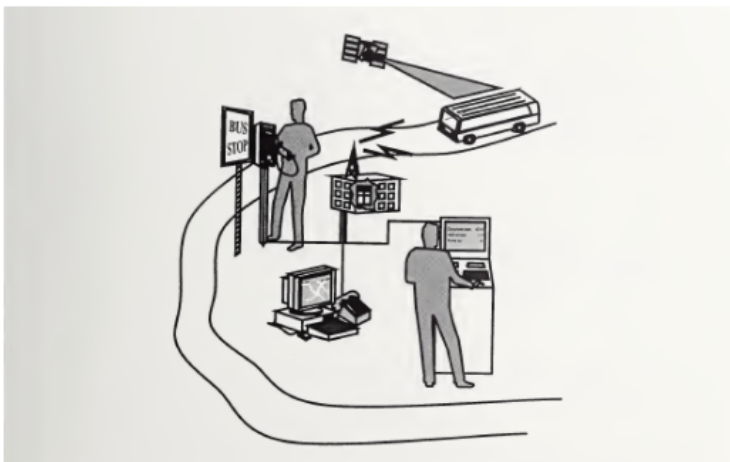
A digital library for transportation research

### Featured Article



#### Connected Vehicles

Coming soon to a road near you



## In the Spotlight

### Intelligent Vehicle Highway Systems

Early research in intelligent transportation systems from the Federal Transit Administration 50th

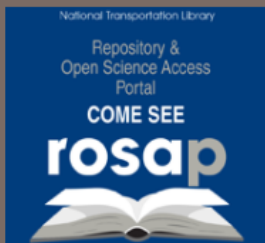
- Assessment of Computer Dispatch Technology in the Para...
- German "Smart-Bus" Systems: Potential for Application ...
- Advanced Traveler Aid Systems for Public Transportatio...

Spotlight Archive

### Stay Connected



Put a ROSA P badge on your website



View All ROSA P Badges

### Recently Added



Advanced Traveler Aid Systems for Public Transportation: The Intelligent Transit Mobility System (ITM...  
Added: 08/25/2017



German "Smart-Bus" Systems: Potential for Application in Portland, Oregon - vol. 1: Technical Report  
Added: 08/25/2017

### Featured Collection



Federal Transit Administration - 50th Anniversary Do...  
Historical resources digitized to celebrate the 50th anniversary of the Federal Transit Administration

### Most Popular



Cost Reduction and Service Improvements from Contracting in Rural, Small Urban, and Suburban Areas



Kentucky Peer Exchange : Strategic Highway Safety Plans (SHSPs) - An RSPCB Peer Exchange



Vortex Advisory System Safety Analysis : Volume 1. Analytical Model



Onboard Monitoring and Reporting for Commercial Motor Vehicle Safety Final Report



Access : Brokering Paratransit Services to the Elderly and Handicapped in Allegheny County

# Collection Policy

---

- NTL collects and maintains significant research and statistical publications in its Digital Repository
- Resources in the collection have value to federal, state, and local transportation decision makers, transportation analysts, and researchers



# Cataloging Policy

---

- NTL records are structured around a modified Dublin Core metadata schema. A minimal cataloging standard was adopted in 2013 that satisfies OCLC K-level standard.
- Drivers
  - Searchable full text repository
  - Reduced staffing level due to sequestration
  - Government mandates
  - Opportunity to leverage existing bibliographic tools (TRID, TRT)
- Goal: Improve cataloging efficiency while maintaining consistent and accurate metadata in NTL's digital collections
  - Identify minimum requirements for functionality and interoperability amongst NTL's user interfaces and digital content platforms (current and future)
  - Meet our legislative mandates
  - Eliminate duplication of effort
  - Develop new cataloging guidelines for NTL's digital repository

# Trustworthy Digital Repositories

---

- Method needed to measure and communicate the value of digital repository to stakeholders
- Digital repositories are not websites; require different analytics
- Trustworthiness is a valid qualitative assessment
- Digital repositories should either be certified or support a repository standard

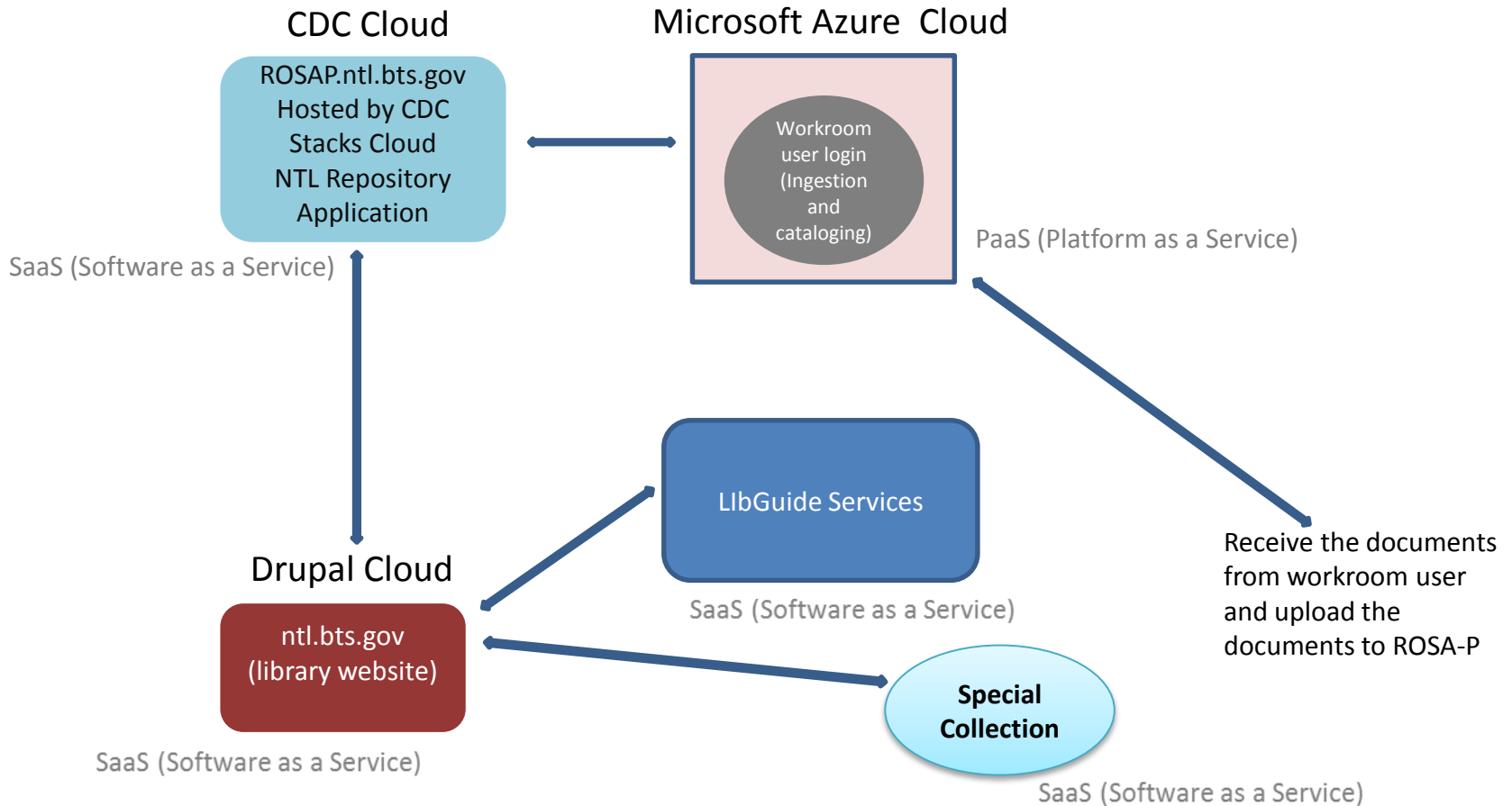


---

# NTL Systems



# NTL Systems Architecture



# NTL Migration to Cloud

---

- NTL Repository in Azure will be for backend processing works
- ROSA\_P is the public access site for NTL repository in CDC Cloud
- NTL home site is in Drupal cloud
- Special Collection is in MicroSearch Private Cloud
- NTL Reference services utilize LibGuide Cloud Services



---

# User Interface



# Searching ROSA P

**Cost Effective Wildlife Crossing Structures which Minimize the Highway Barrier Effects on Wildlife and Improve Highway Safety along US 64, Tyrrell County, NC**



**Final Report**

Produced and Written by

Daniel J. Smith, Ph.D., A.I.C.P., Principal Investigator  
Research Associate/Adjunct Graduate Faculty

Department of Biology  
University of Central Florida  
4000 Central Florida Blvd.  
Orlando, FL 32816

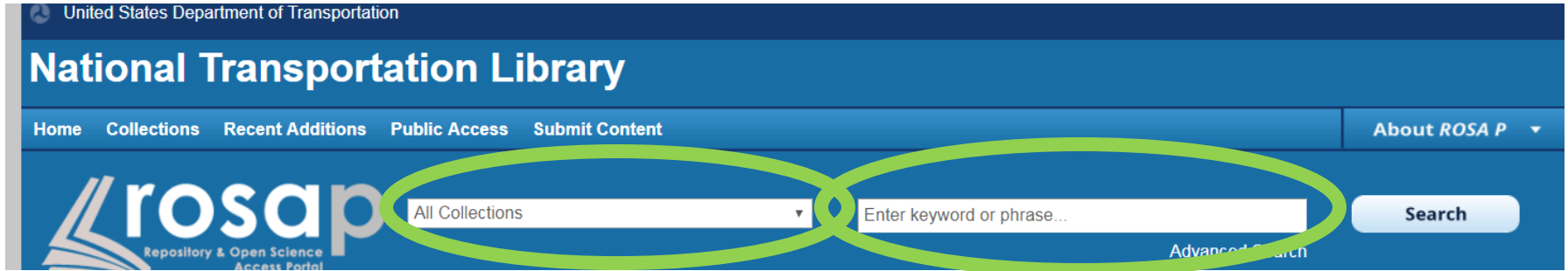
April 2011

1. Homepage: Basic Search
  - Collections dropdown
  - Keyword or phrase
2. Collections:
  - 6 topic groups
  - 2 USDOT Libraries
  - 4 USDOT Operating Administrations
3. Advanced Search
  - Limit by Collection
  - 3 search categories

Cost effective wildlife crossing structures which minimize the highway barrier effects on wildlife and improve highway safety along US 64, Tyrrell County, NC. FHWA/NC/2009-26. 2011.



# Basic Search



## What's New:

- Collection search on the homepage

## Pros:

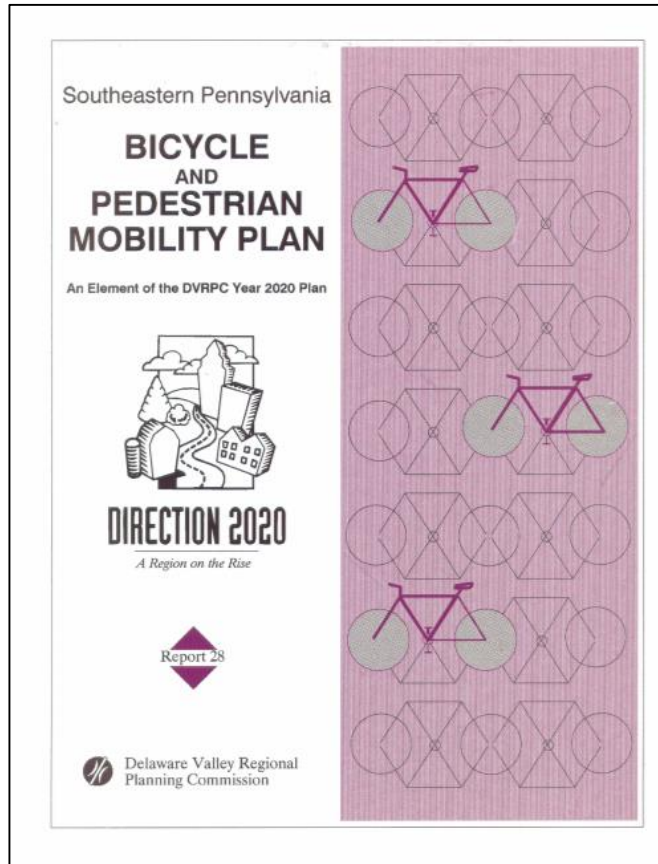
- Quick
- Broad

## Cons:

- Relevancy
- Volume of results



# Basic Search: Example



Southeastern Pennsylvania bicycle and pedestrian mobility plan : an element of the DVRPC year 2020 plan direction 2020. Delaware Valley Regional Planning Commission. 1995.

<https://rosap.ntl.bts.gov/>

- Keyword or phrase: bicycle
    - 3,769 results
  - Results: Page 1 - 20 PDFs
    - Handbooks
    - Reports
    - Pamphlets
    - Manuals
    - Journal Articles
    - Policy Statements
    - Research Papers
- Date Range: 1978-2017



# Basic Search: 3,769 results? That's too many!

United States Department of Transportation

## National Transportation Library

Home Collections Recent Additions Public Access Submit Content About ROSA P

**rosap** Repository & Open Science Access Portal

All Collections  Search Advanced Search

Search Results for "Bicycle"

1 - 20 of 3769 results  50 results

Sort By:

« Previous  2 3 4 ... 189 Next »

**Narrow Results:**

**Resource Type**

- Bibliography (11)
- Booklet/Pamphlet (27)
- Brief (6)
- Dataset (404)

**Subject**

- Accident Data (8)
- Advanced Public Transp...(10)
- Advanced Traffic Manag...(21)
- Advanced Traveler Info... (19)

**Bicycle commuting handbook**

*Published:* 1995

*Description:* Responds to common questions about bicycle commuting and bicycle safety.

*File Type:* [PDF - 120.07 KB]

**bicycle route planning.**

*Published:* 2017-03-01

*Description:* Active transportation modes such as walking and biking are key elements of sustainable : transportation systems. In order to promote biking as an alternative form of transportation, a : holistic approach to improving the quality of the biking experie...



# Collection Search

United States Department of Transportation

## National Transportation Library

Home Collections Recent Additions Public Access Submit Content About ROSA P

rosap  
Repository & Open Science  
Access Portal

All Collections

- All Collections
- BTS Data Directory
- Federal Highway Administration
- Federal Highway Administration Research Library
- Federal Motor Carrier Safety Administration
- Federal Railroad Administration
- Federal Transit Administration
- Federal Transit Administration - 50th Anniversary Document Collection
- Intelligent Transportation Systems Joint Program Office Electronic Library
- National Transportation Data Archive
- NHTSA - Behavioral Safety Research
- US Transportation Collection

Enter keyword or phrase... Search

Advanced Search

Collections present information... months, so return to this site to see what's new.

BTS Data Directory  
Transportation data repository  
USDOT - the Federal

National Transportation Data Archive  
Historical national data... published in

## Pros:

- Focused search
- Smaller number of returns

## Cons:

- Knowledge of collection content

# Collection Search: Example



- Collection: NHTSA – Behavioral Safety Research
- Keyword or phrase: teenager – 259 results
- Results: Page 1 - 20 PFDs
  - Tech Reports
  - Research Papers
  - Journal Articles

*Insurance Institute for Highway Safety, vol. 43, no. 4, June 9, 2008.*

<https://rosap.ntl.bts.gov/>



U.S. Department of Transportation  
Office of the Assistant Secretary for  
Research and Technology

# Advanced Search

United States Department of Transportation

## National Transportation Library

Home Collections Recent Additions Public Access Submit Content About ROSA P

**rosap**  
Repository & Open Science  
Access Portal

### Advanced Search

Select up to three search categories and enter corresponding keywords using the fields on the right. Refer to the [Help](#) section for more detailed instructions.

All Collections

Title

AND

Title

AND

Title

Enter keyword or phrase...

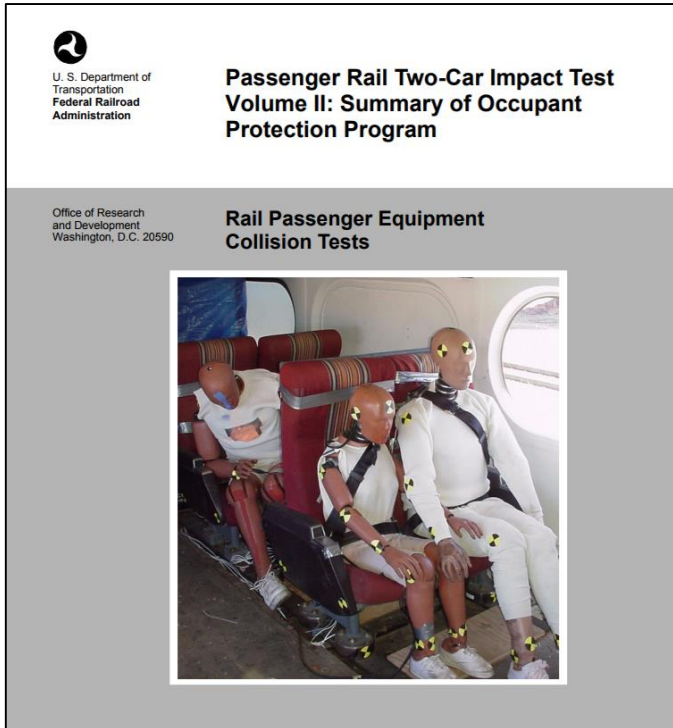
Enter keyword or phrase...

Enter keyword or phrase...

Clear All Simple Search Search

- Limit by: Collection
- 3 Search Categories:
  - Title
  - Abstract
  - Personal Creator
  - Corporate Creator
  - Publication Date
  - DOI
  - Contract Number
  - Report Number
  - Keywords

# Advanced Search: Example



Passenger rail two-car impact test. Volume 2 : summary of occupant protection program. VanIngem-Dunn, C. 2002.

- Collection: BTS Data Directory
- Title: passenger  
AND
- Publication Date:  
1990 to 2000
- Results: Page 1 - 11 PDFs
  - Research Papers
  - Tech Reports
  - Journal Articles
  - Datasets

[Back to Previous Page](#)

## Behavioral distinctions : the use of light-duty trucks and passenger cars

Published Date: 2000-12-01

Language: English

[Viewer](#) [Details](#) [Supporting Files](#) [Related Documents](#) [You May Also Like](#)

Details

Creators: Kockelman, KM ; Zhao, Y ;

Resource Type: Journal Article

Format: PDF

Keywords: [-]  
Behavior  
Light Trucks  
Passenger Cars  
Pickup Trucks  
Sport Utility Vehicles  
Travel Behavior  
Travel Patterns

Description: URL: <http://ntl.bts.gov/lib/10000/10900/10905/3kockelman.pdf>

Pickup trucks, sport utility vehicles (SUVs), and minivans are classified as light-duty trucks and are subject to different safety and emissions protections. According to production and purchase trends, Americans are buying more light-duty trucks. This research explores the subtle differences in ownership and use patterns between light-duty trucks and passenger cars using National Household Travel Survey (NHTS) data. The results, based on a variety of model specific variables, show that light-duty trucks are driven longer distances with more people aboard and is purchased by wealthier households, while passenger cars are driven shorter distances, are purchased by smaller households, and carry fewer people.

Collection(s): [BTS Data Directory](#)

# Use Case

## Linked:

- Creators
- Resource Type
- Keywords
- Collections



# Datasets: Open Access Initiative



## Analytical procedures for determining the impacts of reliability mitigation strategies.

**Published:** 2013

**Description:** Reliability of transport, especially the ability to reach a destination within a certain amount of time, is a regular concern of travelers and shippers. The definition of reliability used in this research is how travel time varies over time. The ...

**File Type:**  [PDF - 9.53 MB]



## Analytical procedures for determining the impacts of reliability mitigation strategies [supporting datasets]

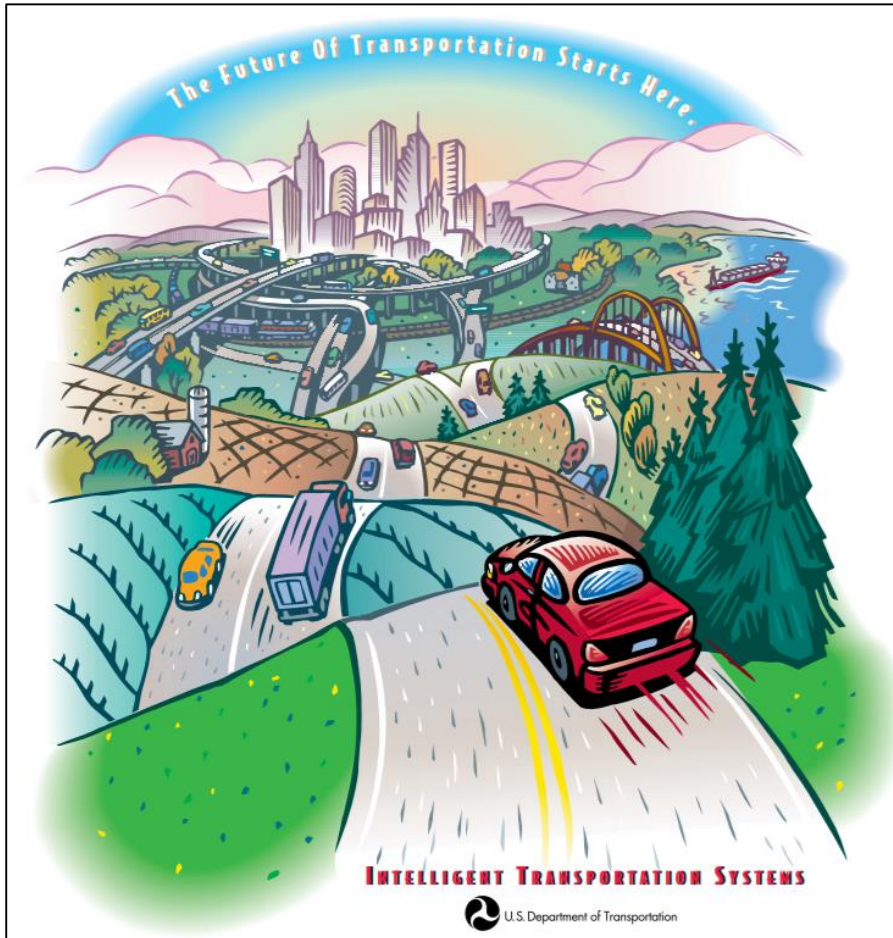
**Published:** 2012-11-30

**Description:** This zip file contains comma separated value (.csv) files of data to support SHRP 2 report S2-L03-RR-1, Analytical procedures for determining the impacts of reliability mitigation strategies. Zip size is 1.83 MB. Files were accessed in Microsoft Excel...

**File Type:**  [ZIP - 1.58 GB]



# What's Next?



- Curated searches
- Record cloning
- Linking related records
- Exporting to Excel

The future of transportation starts here. United States. Joint Program Office for Intelligent Transportation Systems. 1998.

<https://rosap.ntl.bts.gov/>



U.S. Department of Transportation  
Office of the Assistant Secretary for  
Research and Technology

---

# The Good, The Bad, & The Metadata



# NTL Digital Repository Metadata

---

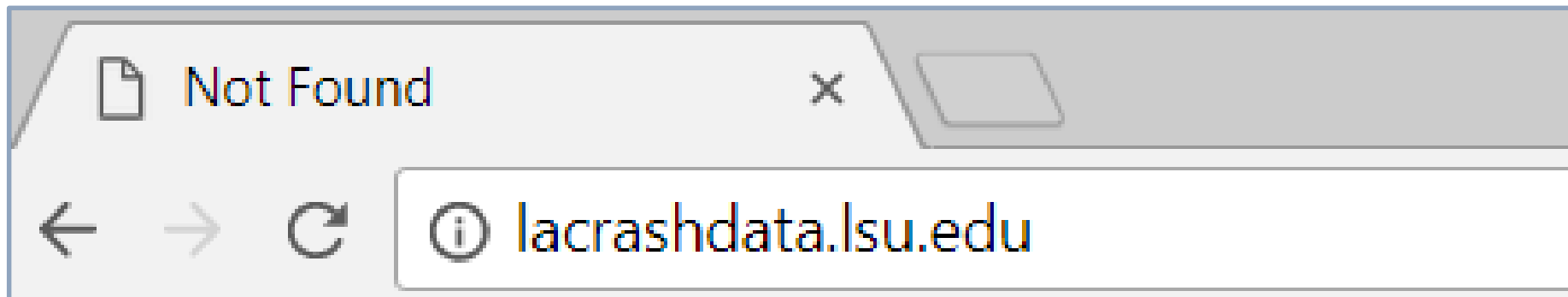
## Legacy

- NTL Minimal Cataloging Standard (OCLC K-Level) (2014 forward)
- Modified Dublin Core metadata
- Indexed by popular search services Google and Bing
- Some bad habits (upcoming slide)

## Developing

- Revising metadata to Project Open Data Metadata, as well as PREMIS and other data-specific metadata fields
- Goals: Public Access & Linked Open Data
- Additional indexing by Google Scholar
- DOIs minted for landing pages in ROSA P
- Better habits and workflows

# 404



## Not Found

- 3200+ Rotted Links
- Non-standard ingestion process
- Repairs: some easy, some hard, some impossible
- Time intensive

HTTP Error 404. The requested resource is not found.



United States Department of Transportation

# National Transportation Library

Home Collections Recent Additions Public Access Submit Content



All Collections ▾

Enter keyword or phrase...

← Back to Previous Page

## Connected vehicles and your privacy.

Published Date: 2016

Language: English

Viewer Details Supporting Files Related Documents You May Also Like

PDF - 1.20 MB Download

Page: 1 of 2 Automatic Zoom

Photo Source: USDOJ

**CONNECTED VEHICLES AND YOUR PRIVACY**

LIMIT 35

No Personal Information Shared

# Using DOIs in ROSA-P



# Using DOIs in our Records

bts NTL Workroom Data Entry - Google Chrome

Secure | <https://ntlsearch.bts.gov/ntl/workroom/holderentry.jsp?DOCID=56951>

Data Entry 1 2 3 4 Overview LC Authority Search

56951 - Approved

Title Connected vehicles and your privacy.

|                       |                      |             |                           |                      |
|-----------------------|----------------------|-------------|---------------------------|----------------------|
| TRIS                  | <input type="text"/> | TRIS >> NTL | OCLC                      | <input type="text"/> |
| ISBN                  | <input type="text"/> |             | ISSN                      | <input type="text"/> |
| Conference Begin Date | <input type="text"/> |             | Conference End Date       | <input type="text"/> |
| Copyright Date        | <input type="text"/> |             | Date Captured             | 16-02-11             |
| Copyright             | <input type="text"/> |             | Digital Object Identifier | <input type="text"/> |
| Contract Number(s)    | <input type="text"/> |             | Contracting Officer       | <input type="text"/> |
| Report Number(s)      | FHWA-JPO-16-268      |             | Edition                   | <input type="text"/> |

Supplier

Collection(s) Intelligent Transportation Systems Joint Program Office{103}

Staff Notes

Status Approved

Status Comment

Date Modified 2016-02-11 14:06:37.0 Modified By megan.tietjen

Date Created 2016-02-11 12:34:03.0 Created By megan.tietjen

Comment Date Comment By

Prev Next Finish Reset Cancel



# Minting DOIs

```
ROSAP_clean_start_DOI_to_XML_2017_08_22_1510.xml
1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <ntl_rosap_doi_xml_generator_20170821 xmlns:xsi=
  "http://www.w3.org/2001/XMLSchema-instance">
3   <record>
4     <title>Airworthiness Approval of Global Positioning System (GPS) Navigation
      Equipment for Use as a VFR and IFR Supplemental Navigation System</title>
5     <creators>Rebelo, Jorge M. ; Benvenuto, Pedro P.; </creators>
6     <publisher>Federal Aviation Administration Technical C
7     <publication_date>1994</publication_date>
8     <site_url>http://phony_url</site_url>
9     <product_type>Text</product_type>
10    <product_type_specific>Tech Report</product_type_speci
11    <contact_name>National Transportation Library</contact
12    <contact_org>National Transportation Library</contact
13    <contact_email>answers@dot.gov</contact_email>
14    <contact_phone>800-853-1351</contact_phone>
15    <product_nos>FAA AC 20-138</product_nos>
16    <other_identifying_nos>3</other_identifying_nos>
17  </record>
18  <record>
19    <title>Truck safety : Motor Carriers office's activities to reduce fatalities
      are likely to have little short-term effect</title>
20    <creators>[Yung, Jessie] ; [Le Frois, Gregory A.] ; [Lee, Ming-Shuin] ;
      [Chang, Mark]; United States. Federal Highway Administration ; HNTB
      Corporation ; Booz Allen Hamilton for 2001-</creators>
21    <publisher>United States. General Accounting Office</publisher>
```

- Annual Subscription Service provided by DOE OSTI
- XML file with metadata subset to mint
- Short turn around depending on file size

# Metadata Lessons

---

## Problematic File Types

- HTML (Web pages)

## Employ Standards

- Dates: ISO 8601
  - YYYYMMDD or
  - YYYY-MM-DD
- Geographic Abbreviations: ISO 3166-2
  - California or Coahila?
    - US-CA; MX-CA
  - Hawai'i or Hidalgo
    - US-HI; MX-HI

## Potential Improvements






- Always ingest local copy of publications and reasonably sized files
  - LOCKSS
- Fixity Hashing & Checks
- Master Files vs. Use Files
- Document Workflows



# Next Step: Data Packages

---

A “Data Package” is the dataset, the DMP, and all other documentation needed to contextualize the dataset for any and all users.

-  [bts\\_ota\\_border\\_crossing\\_entry\\_Data\\_2017\\_04\\_12.csv](#)
-  [bts\\_ota\\_border\\_crossing\\_entry\\_DMP\\_2017\\_04\\_12.txt](#)
-  [bts\\_ota\\_border\\_crossing\\_entry\\_Metadata\\_2017\\_04\\_12.json](#)
-  [bts\\_ota\\_border\\_crossing\\_entry\\_PortCodes\\_2017\\_04\\_11.csv](#)
-  [bts\\_ota\\_border\\_crossing\\_entry\\_README\\_2017\\_04\\_12.txt](#)

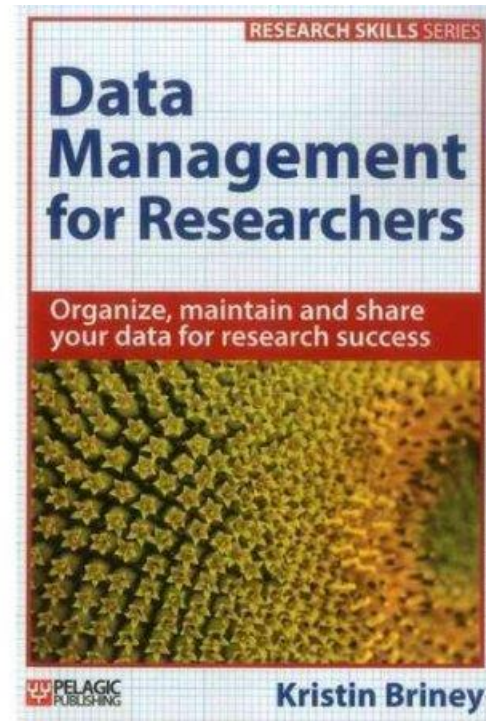
# Data Package Elements

## NTL Dataset Data Package Guidelines

- **Dataset**
  - .csv or other open format
- **Readme.txt**
  - Includes Data Dictionary
  - Notes standards used
  - Defining Zero, Null, & Unknown
  - FAQs and other notes
- **Metadata file** in Project Open Data .json and/or other schema
- **Data Management Plan (DMP)**
- *Code or scripts* used in data analysis
- *Supporting files, tables, etc.*

(**Bold** = Required; *Italics* = Optional, or Required if Applicable)

Adapted from Best Practices described in:



# Element: Data

| 1  | Port Name,State Abbreviation,Port Code,Border,Date,Measure,Value,Location                              |
|----|--|
| 2  | Alcan,AK,3104,US-Canada Border,1996-01,Personal Vehicle Passengers,2011,"Alaska (62.61, -141.0) "      |
| 3  | Alcan,AK,3104,US-Canada Border,1996-01,Personal Vehicles,965,"Alaska (62.61, -141.0) "                 |
| 4  | Alcan,AK,3104,US-Canada Border,1996-01,Trucks,428,"Alaska (62.61, -141.0) "                            |
| 5  | Alcan,AK,3104,US-Canada Border,1996-01,Bus Passengers,9,"Alaska (62.61, -141.0) "                      |
| 6  | Alcan,AK,3104,US-Canada Border,1996-01,Buses,3,"Alaska (62.61, -141.0) "                               |
| 7  | Alcan,AK,3104,US-Canada Border,1996-01,Empty Rail Containers,0,"Alaska (62.61, -141.0) "               |
| 8  | Alcan,AK,3104,US-Canada Border,1996-01,Empty Truck Containers,0,"Alaska (62.61, -141.0) "              |
| 9  | Alcan,AK,3104,US-Canada Border,1996-01,Loaded Rail Containers,0,"Alaska (62.61, -141.0) "              |
| 10 | Alcan,AK,3104,US-Canada Border,1996-01,Loaded Truck Containers,0,"Alaska (62.61, -141.0) "             |
| 11 | Alcan,AK,3104,US-Canada Border,1996-01,Pedestrians,0,"Alaska (62.61, -141.0) "                         |
| 12 | Alcan,AK,3104,US-Canada Border,1996-01,Train Passengers and Crew,0,"Alaska (62.61, -141.0) "           |
| 13 | Alcan,AK,3104,US-Canada Border,1996-01,Trains,0,"Alaska (62.61, -141.0) "                              |
| 14 | Alexandria Bay,NY,708,US-Canada Border,1996-01,Personal Vehicles,29758,"New York (44.35, -75.98) "     |
| 15 | Alexandria Bay,NY,708,US-Canada Border,1996-01,Trucks,15546,"New York (44.35, -75.98) "                |
| 16 | Alexandria Bay,NY,708,US-Canada Border,1996-01,Bus Passengers,1505,"New York (44.35, -75.98) "         |
| 17 | Alexandria Bay,NY,708,US-Canada Border,1996-01,Loaded Truck Containers,160,"New York (44.35, -75.98) " |

CSV: Open, Non-Proprietary Format  
Backward and Forward Compatible  
Long-Term Preservable

# Element: ReadMe

---

## ReadMe.txt

- Includes Data Dictionary
- Notes standards used
- Defining Zero, Null, & Unknown
- FAQs and other notes

## Data Dictionary Example

E. Name: date

Full name: Date

Variable Definition or Description: Year and Month of count of border crossing value.

Data format: YYYY-MM

Field length: 7 characters

Representation of null values: No null values

Note: Date represented in ISO 8601 extended format for Year and Month (YYYY-MM)

## ReadMe Template: Contents

- A. General Information
- B. Sharing/Access Information
- C. Data and Related File Overview
- D. Methodological Information
- E. Data-Specific Information
- F. Appendices

# Element: Metadata

## Project Open Data Metadata V1.1 .json

Validate JSON

Clear

### Results

valid JSON

bts\_ota\_border\_crossing\_entry\_Metadata\_2017\_04\_12.json

1  
2  
3

```
{
  "title": "Border Crossing/Entry Data",
  "description": "The Bureau of Transportation Statistics (BTS) Border Crossing/Entry Data provides summary statistics to the public for incoming crossings at the U.S.-Canadian and the U.S.-Mexican border at the port level.",
  "keyword": "border crossings, ports, trucks, trains, containers, buses, personal vehicles, passengers, pedestrians, canada, mexico",
  "modified": "20170131",
  "publisher": "U.S. Department of Transportation, Bureau of Transportation Statistics, Office of Transportation Analysis",
  "contactPoint": {
    "@type": "vcard:Contact",
    "fn": "National Transportation Library Data Curator",
    "hasEmail": "ntldatacurator@dot.gov"
  },
  "identifier": null,
  "accessLevel": "public",
  "accessLevelComment": null,
  "bureauCode": "021:04",
  "programCode": "021:053",
  "accessURL": "https://data.transportation.gov/Research-and-Statistics/Border-Crossing-Entry-Data/keg4-3bc2",
  "webService": null,
  "format": "csv",
  "license": "Public Domain",
  "spatial": "United States",
  "temporal": "19960101/20170131",
  "accrualPeriodicity": "R/P3M",
  "landingPage": "https://transborder.bts.gov/programs/international/transborder/TBDR_BC/TBDR_BC_Index.html",
  "language": "en-US",
  "issued": "20170328",
  "dataQuality": "yes",
  "references": "bts_ota_border_crossing_entry_README_2017_04_12.txt,
```

24  
25  
26  
27

# Element: DMP

U.S. Department of Transportation  
National Transportation Library

NTLDataCreator@dot.gov  
(202) 366-DATA

### DATA MANAGEMENT PLANNING "GREENOUT" BINGO SELF-ASSESSMENT TOOL

This tool is designed to help assess your current data management practices. Pick a dataset or research project, sit down with your data collection team, and discuss each prompt below. This tool can guide DMP improvement by revealing best practices that you are not using or information that you may not have gathered.

Using colored pencils, highlighters, or the highlighter tool in your document reader, shade your answer in each prompt-box as indicated. The greener the card, the better. But don't force it: Good data management is a process. NTL can help. Write your questions on the back, and contact us for help.



**Red:** Not current practice, Unknown, or No.  
**Yellow:** We are planning on this or are working on it.  
**Green:** Practice enacted, or Yes.  
**Blue:** Not Applicable.

Dataset Name: North American Transborder Freight Data  
URL: [https://www.transportation.gov/management/ntl-repository/transborder/BTSR\\_GA.html](https://www.transportation.gov/management/ntl-repository/transborder/BTSR_GA.html)

Evaluators: Leighton Christiansen

Date: 2016-08-15

| Data Description  | Standards Employed   | Access Policies  | Re-use & Derivative Policies   | Archiving & Preservation Plans   |
|---|--|--|--|--|
| Where it is stored, described, and linked to a research project or data program | If data is created in proprietary formats, digital in open formats are also provided   | How is public access?  | Are intellectual property rights to the dataset formally established?  | A data repository dedicated to long-term preservation has been chosen for this data? |
| The format of data generated by the dataset or program (e.g., access)           | What format is intended for use?   | What data will be shared, and how, if known and recorded?            | Are there any restrictions on the most open license possible, such as Public Domain or Creative Commons (CC BY)? | A minimum data retention period has been agreed upon (with chosen archive)?          |
| The type of data it records and the frequency for that                          | Directory and file naming conventions are documented & used at time of data collection | Researchers have been notified of the data and its location          | If there is a data embargo period, is as short as possible?  | Permanent identifiers such as DOI are used to link to the data                       |
| How long will the data be organized and available to be accessed?               | Metadata is used to contextualize the data, making it comprehensible to all and others | Researchers' sensitive information (PII) is protected and documented | Methods exist to ensure the data is recoverable  | Back-up and disaster recovery plans are in place                                     |
| How responsible for managing data are you and others?                           | Published metadata schemas are employed, and are standard to the field                 | Exchange partners are notified of the data and its location          | Special handling is used with the data   | Staff has been assigned to migrate   |

Based on NTL's "Creating Data Management Plans" <http://ntl.gov>

### 2. Data Description:

| Data Description Narrative Evaluation Prompts   | Explained Fully                     | Partially Explained      | No Information                      | Not Applicable                      |
|---|-------------------------------------|--------------------------|-------------------------------------|-------------------------------------|
| 2.01 The DMP names the data, data collection project, or data producing program.  | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.02 The DMP describes the purpose of the research or data collection.  | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.03 The DMP describes the data generated in terms of nature and scale (e.g., numerical data, image data, text sources, video, audio, database, modeling data, source code, etc.).  | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.04 The DMP describe methods for creating the data (e.g., simulated; observed; experimental; software; physical collection; sensors; satellite; enforcement activities; researcher-generated databases, tables, and/or spreadsheets; digital data such as images and video; etc.). | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.05 The DMP discusses the period of time data will be collected and frequency of update.   | <input type="checkbox"/>            | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/>            |
| 2.06 The DMP describes the relationship between the new data collected for this effort and any existing data also used.   | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/>            | <input checked="" type="checkbox"/> |
| 2.07 The DMP lists potential users of the data.   | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.08 The DMP discusses the potential value the data have over the long-term for not only U.S. DOT, but also for the public.   | <input type="checkbox"/>            | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/>            |
| 2.09 If the DMP contains a request permission to not make the data publicly accessible, it explains the rationale for lack of public access.  | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/>            | <input checked="" type="checkbox"/> |
| 2.10 The DMP indicates the party responsible for managing the data.   | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| 2.11 The DMP describes how project leads will check for adherence to this data management plan.   | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| Total of checked boxes for each column, out of 11:  | 6                                   | 1                        | 2                                   | 2                                   |

**Evaluation questions:**

- Did a majority of the prompts rate "Explained Fully"?
- Do you have a complete picture of: What the data will be gathered; How much data to expect; and, Who is responsible for managing data, and how the data will be managed?

If not, this section may not be sufficiently detailed.

bts\_ota\_border\_crossing\_entry\_DMP\_2017\_04\_12.txt

43 1. Data Description:

44 The dataset is known as "Border Crossing/Entry Data." The Bureau of Transport (BTS) Border Crossing/Entry Data provides summary statistics to the public for crossings at the U.S.-Canadian and the U.S.-Mexican border at the port level available for trucks, trains, containers, buses, personal vehicles, passenger pedestrians. Data fields are numeric counts and textual sequences. These data U.S. Customs and Border Protection (CBP) at individual ports of entry, then tabulated by staff of CBP's Office of Field Operations. CBP uses a mixture of automated procedures to collect the data. The data reflect the number of vehicles, containers, passengers or pedestrians entering the United States. U.S. Customs and Border Protection does not collect comparable data on outbound crossings. Users seeking outbound vehicles may therefore want to review data from individual bridge or state governments, or the Mexican and Canadian governments. The Border Crossing/Entry Data dataset begins with the January 1996 data month.

45 Potential users of the data include: DOT staff, Congressional staff, state DOTs, international organizations, universities and other researchers.

46 The Border Crossing/Entry Data dataset has significant long-term value. Custom border crossing data for a variety of purposes, including trade corridor transportation infrastructure planning, marketing and logistics plans. This data has been used in other BTS products (TSAR, Pocket Guide, etc.). The Border Crossing/Entry Data has been collected since 1996, and thus provide a valuable continuous dataset to interested in transportation access the land borders of the United States.

47 Sean Jahannir and Steven Beningo are responsible for managing these data. We will update this data management plan by conducting annual reviews of this DMP.

48

49 2. Standards Employed:

50 The data have been sent to BTS by U.S. Customs and Border Protection via e-mail format since 2005. Prior to 2005, the data were provided to BTS on a CD-ROM interactive interface. Since Excel .xls is a proprietary file format owned by Microsoft Corporation, BTS will save this dataset as a .csv file, and open source, non-proprietary format, for long-term preservation and distribution to the public.

51

52 Each new version of the dataset will include the Year and Month of the update name, on the documentation webpage, and in the bts\_ota\_border\_crossing\_entry\_DMP\_2017\_04\_12.txt that accompanies the dataset in the NTL repository and other web interfaces.

53

54 The formats used in this dataset are standard. Dates are recorded based on the standard. Metadata is recording use Project Open Data Metadata version 1.1. Metadata will be stored and accessible from the NTL repository, data.transportation.gov, and the Border Crossing/Entry Data website.

55

56 Documentation is easily accessible on the Border Crossing/Entry Data website. The website documentation includes a list of frequently asked questions, coverage, availability definitions and notes. Open and machine readable documentation will also include such as: bts\_ota\_border\_crossing\_entry\_PortCodes.csv; bts\_ota\_border\_crossing\_entry\_README.txt; bts\_ota\_border\_crossing\_entry\_DMP\_2017\_04\_12.txt; bts\_ota\_border\_crossing\_entry\_Metadata.json, and others as appropriate.

# Element: Codebooks & Supporting Files

---



The image shows a screenshot of a CSV file named 'bts\_ota\_border\_crossing\_entry\_PortCodes\_2017\_04\_11.csv'. The file contains a list of 19 rows, each representing a port code, its name, and the state it is located in. The first row is a header row with the columns 'Port Code, Port Name, State'. The subsequent rows list specific port codes and their corresponding names and states.

| Port Code | Port Name     | State |
|-----------|---------------|-------|
| 3104      | Alcan         | AK    |
| 3106      | Dalton Cache  | AK    |
| 3102      | Ketchikan     | AK    |
| 3103      | Skagway       | AK    |
| 2601      | Douglas       | AZ    |
| 2602      | Lukeville     | AZ    |
| 2603      | Naco          | AZ    |
| 2604      | Nogales       | AZ    |
| 2608      | San Luis      | AZ    |
| 2606      | Sasabe        | AZ    |
| 2502      | Andrade       | CA    |
| 2503      | Calexico      | CA    |
| 2507      | Calexico East | CA    |
| 2506      | Otay Mesa     | CA    |
| 2504      | San Ysidro    | CA    |
| 2505      | Tecate        | CA    |
| 3302      | Eastport      | ID    |
| 3308      | Porthill      | ID    |

# Where will we see Data Packages?

The screenshot shows the National Transportation Library (rosap) website. The page title is "The Impact of Family and Peer Protective Factors on Girls' Violence Perpetration and Victimization". The page includes a navigation menu with "Home", "Collections", "Recent Additions", "Public Access", and "Submit Content". The "Supporting Files" tab is highlighted, and a red arrow points to it. Below the tabs, a list of supporting files is displayed:

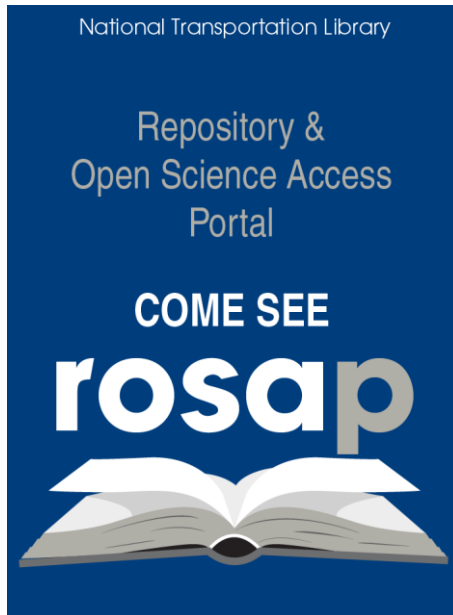
- bts\_ota\_border\_crossing\_entry\_Data\_2017\_04\_12.csv
- bts\_ota\_border\_crossing\_entry\_DMP\_2017\_04\_12.txt
- bts\_ota\_border\_crossing\_entry\_Metadata\_2017\_04\_12.json
- bts\_ota\_border\_crossing\_entry\_PortCodes\_2017\_04\_11.csv
- bts\_ota\_border\_crossing\_entry\_README\_2017\_04\_12.txt

The page also features a search bar, a "Download" button for a PDF file (372.15 KB), and a "Back to Previous Page" link.



# Questions or comments?

---



- **Ask-a-Librarian**  
[Answers@dot.gov](mailto:Answers@dot.gov)  
1-800-853-1351
- **Data Services**  
[NTLDataCurator@dot.gov](mailto:NTLDataCurator@dot.gov)
- **Digital Library Services**  
[NTLDigitalSubmissions@dot.gov](mailto:NTLDigitalSubmissions@dot.gov)

