

# SHRP2 Naturalistic Driving Study and some other DOT data activities

2016 NSF BIGDATA PI MEETING  
APRIL 20TH-21ST – WASHINGTON, DC

Charles Fay SHRP 2 Research Manager  
Safety Training & Analysis Center (STAC)

Office of Safety Research & Development  
Turner Fairbank Highway Research Center  
Federal Highway Administration  
US Department of Transportation

[charles.fay@dot.gov](mailto:charles.fay@dot.gov)



## What are some big data challenges ?

- Federal, State, and local transportation agencies' – different roads-different agency, disparate systems; agencies' staff experience with big data today is limited
- Big data management and analytics novel territory – practitioners at agencies; researchers as well
- A connected-automated highway system – requires different areas within a DOT and across agencies to communicate/ share resources, including data collection and analytics. Not trivial endeavor to integrate just the roadway part of the system – not to mention vehicles and other users
- Standardization of data and QA; non proprietary/ interoperable format
- Anticipated volume of data with connected vehicle
- Protection of personally identifiable information (PII)

## Data sets that are available for the community to use

- **SHRP2 Naturalistic Driving Study (NDS) data** – a ‘big’ resource for DOTs and others~ \$70M investment for the data collection – max use/benefit, while protecting PII.
  - 2PB; 1Mhrs video, sensor - time series data(varying resol); veh and network data recorded 1-2 yr; 5.4M trips; 30M+ miles; six US sites; 1465 crashes, 2800 near crashes; 3500 participants-demographic, assessment testing; PII and lots of it(face, O/D within trillions GPS pts). Linkable to Roadway Information Database = context for trips (GIS-disparate sources-able to link additional data)
  - National Academies/ TRB – stewards for Phase1 till 2019/20 – interested partners for Phase 2?
  - InSight website, InDepth data access-DUL/IRB/PII in approved-secure enclave (e.g., STAC). <https://www.fhwa.dot.gov/research/resources/stac/3126a-DUL> in one year; over 1200 users of website; original data needs to be

# Data sets & DOT data activities to be aware of

Research Data Exchange (RDE) [www.its-rde.net](http://www.its-rde.net).

- Open Source Application Development Portal(OSADP) [www.itsforge.net](http://www.itsforge.net)
- Future data - Connected Pilot project <http://www.its.dot.gov/pilots/>

<https://www.fhwa.dot.gov/2015datapalooza/> - ex “volume” and “variety” aspects of big data analysis from the Department.

- FHWA/NREL collaboration to collect household travel survey data  
[https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.1\\_Murakami.pdf](https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.1_Murakami.pdf)
- Battelle work on fusing administrative records with surveys  
[https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.4\\_Pierce.pdf](https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.4_Pierce.pdf)
- FAA/MITRE work  
[https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.5\\_Sawyer.pdf](https://www.fhwa.dot.gov/2015datapalooza/presentations/PolicyDev.5_Sawyer.pdf)

## DOT theme in this domain ex. quantifying and tracking risk

**Building our data integration muscle.** Pull together disparate, multi-modal and cross-governmental data sets to facilitate safety risk management.

**Building our data analytics technology.** Find low-cost, innovative ways to analyze the integrated data sets and build novel data products that inform leadership.