

Data Management Plan for “2007 National Roadside Study of Alcohol and Drug Use by Drivers [Supporting Datasets]”

20240116

Basic Information

0. Basic Information

- 0.01 Lead researcher, or lead staff name: Amy Berning
- 0.02 Lead researcher, or lead staff ORCID or other identifier: <https://orcid.org/0000-0002-0614-2793>
- 0.03 Lead researcher contact information: amy.berning@dot.gov
- 0.04 Organization: Office of Behavioral Safety Research (BSR), National Highway Transportation Safety Administration (NHTSA), U.S. Department of Transportation
- 0.05 Other researchers: See cataloging record and reports for other authors
- 0.06 Title of Research Proposal/Project: 2007 National Roadside Study of Alcohol and Drug Use by Drivers
- 0.07 URL: <https://doi.org/10.21949/1529967>
- 0.08 This is an ☒ initial DMP or a ☐ revised DMP.
- 0.09 Today's date (YYYY-MM-DD): 2024-01-16
- 0.10 This DMP was created by Leighton L Christiansen <https://orcid.org/0000-0002-0543-4268>, Data Curator, leighon.christiansen@dot.gov. You may also contact the NTL Data Curator at NTLDataCurator@dot.gov

1. Data Description:

- 1.01) Dataset name: “2007 National Roadside Study of Alcohol and Drug Use by Drivers [Supporting Datasets]”
- 1.02) This dataset supports the conclusions of the 2007 National Roadside Study of Alcohol and Drug Use by Drivers. The three report volumes -- Methodology, Drug Results, and Alcohol Results -- are available in this repository and linked to this record. This dataset records results from 11,120 roadside surveys conducted across the United States in 2007, as described in the Methodology volume. When possible, oral fluid and blood samples were taken and were subjected to laboratory screening and liquid chromatography-mass spectrometry (LC/MS/MS; the term MS/MS is the combination of two mass analyzers in one mass spec instrument) and gas chromatography-mass spectrometry (GC/MS) confirmation respectively for alcohol and several classes of drugs, allowing researchers to estimate a national prevalence of alcohol and other drugs in drivers. The dataset also includes information from a self-administered questionnaire, also described in the Methodology document. All drivers' responses were completely voluntary and anonymous. Re-identification risk of study participants from this dataset is very low. The data files are collected into a single ZIP file.
- 1.03) Data was collected between 2007-07-01 to 2007-11-30. No updates
- 1.04) The data can provide long-term value by helping to mark changes in impaired driving over time.
- 1.05) Data is fully accessible to the public.
- 1.06) The National Transportation Library is now responsible for the long-term preservation of the dataset. All responsibility for data content lies with NHTSA.
- 1.07) Describe how you will check for adherence to this data management plan.

2. Standards Employed: Your DMP should describe the anticipated formats that your data and related files will use. To the maximum extent practicable, and in accordance with generally accepted practices in your field, your DMP should address how you will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future. If you are unable to use platform-independent and non-proprietary formats, you should specify the standards and formats that will be used and the rationale for using those standards and formats. Identify the metadata standards you will use to describe the data.

As general guidance you should address the following prompts, as they apply:

- 2.01) The data are available in the following formats: .CSV, .XLSX, .SAS, .SAV, and .DTA. .CSV is an open format. All others are proprietary.
- 2.02) The proprietary formats allow users to work in one of many current and ubiquitous statistical software programs: .XLSX version, opens with Microsoft Excel or other spreadsheet program; .SAV version, opens with IBM SPSS statistical software; .SAS version, opens with SAS statistical software; and, .DTA version, opens with Stata statistical software.
- 2.03) This is the final version of the data. If future updates or changes are needed, file name date and timestamps will be updated, as well as the README.txt document.
- 2.04) Documentation includes a Data Dictionary, a README.txt, this DMP, and the Methodology Report.
- 2.05) This dataset is described using the DCAT-US Version 1.1 metadata schema in file NHTSA_BSR_NRS_2007_METADATA_20240116_1130.json

3. Access Policies:

- 3.01) This data may be shared with the public.
- 3.02) All respondents answered the questionnaire anonymously. Each response has been given a semi-random identifier. The NTL Data Curation team has reviewed the variables in dataset and find the risk of reidentification of any respondent to be very low.
- 3.03) There are no privacy, ethical, or confidentiality concerns raised from sharing this data.
- 3.04)** Respondents' names were not collected at the time of the survey. Further, each row was given a semi-random identifier.

4. Re-Use, Redistribution, and Derivative Products Policies:

- 4.01) This data is managed by the National Transportation Library through agreement with the Office of Behavioral Safety Research (BSR) of the National Highway Transportation Safety Administration (NHTSA).
- 4.02) This data was transferred to NTL in 2023-07.
- 4.03)** This data is in the public domain.

5. Archiving and Preservation Plans: Describe how you intend to archive your data and why you have chosen that particular option. You may select from a variety of options including, but not limited to:

- Use of an institutional repository
- Use of an archive or other community-accepted data storage facility
- Self-dissemination

You must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability. Whatever archive option you choose, that archive must support the capture and provision of the U.S. Federal Government **Project Open Data Metadata Schema**

<<https://resources.data.gov/resources/dcat-us/v>> (known as DCAT-US Schema v1.1 as of 2020). In addition, the archive you choose must support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and must provide for maintenance of those identifiers throughout the preservation lifecycle of the data. Your plan should address how your archiving and preservation choices meet these requirements.

As general guidance you should address the following prompts, as they apply:

- 5.01) This dataset will be preserved by the National Transportation Library (NTL) in the Repository & Open Science Access Portal (ROSA P). The dataset landing page is at <https://doi.org/10.21949/1529967>
- 5.02) In order to protect digital information and data from loss, NTL employs the “3-2-1” backup rule. NTL maintains:
 - A) Three (3) copies of the electronic files
 - B) Stored on two (2) different kinds of storage media
 - C) With at least one (1) copy stored in a different geographic and geologic region.
 - i) Currently, NTL maintains a copy of its repository content and metadata in the following locations:
 - (1) USDOT- managed Microsoft Azure cloud environment
 - (2) CDC Public Access Platform (Amazon Web Services cloud environment)
 - (3) Removable media (external drive)
 - (4) Backups on the USDOT-managed Microsoft Azure cloud environment are in the disaster recovery site, in a different geographical area than USDOT headquarters. Backups on the CDC Public Access Platform are in the disaster recovery (DR) site on the US West Coast, a different geographic area than CDC headquarters. The disaster recovery site is updated daily. All daily backups of the staging server and weekly backups of the production servers are kept for 45 days.
- 5.03) The NTL will preserve and share the data in perpetuity.
- 5.04) NTL mints DOIs for datasets. The DOI for this dataset is <https://doi.org/10.21949/1529967>
- 5.05) NTL’s ROSA P fully meets the criteria outlined on the **Guidelines for Evaluating Repositories for Conformance with the DOT Public Access Plan** page <<https://doi.org/10.21949/1520563>>.

6. Policies Affecting this Data Management Plan

This data management plan was created to meet the requirements enumerated in the U.S. Department of Transportation's **Plan to Increase Public Access to the Results of Federally-Funded Scientific Research Version 1.1** <<https://doi.org/10.21949/1520559>> and guidelines suggested by the DOT **Public Access website** <<https://doi.org/10.21949/1503647>>, in effect and current as of January 2024.