256-RU2806

Commercial Vehicle Data Quality Support for NJDOT
"Improve SafetyNet Data Quality"

FINAL REPORT
January 2012

Submitted by

Patrick Szary, Ph.D.
Associate Director
Center for Advanced Infrastructure and Transportation
Rutgers, The State University of New Jersey

NJDOT Project Manager
Paul Truban

In cooperation with

New Jersey
Department of Transportation
Bureau of Trucking Services
and
U.S. Department of Transportation
Federal Highway Administration

DISCLAIMER STATEMENT

The contents of this report reflect the views of the author(s) who is (are) responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the New Jersey Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

TECHNICAL REPORT
STANDARD TITLE PAGE

1. Report No. 256-RU2806	2.Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle FINAL REPORT Commercial Vehicle Data Quality Support for NJDOT "Improve SafetyNet Data Quality"		5. Report Date January 2012 6. Performing Organization Code Rutgers/CAIT
7. Author(s) Patrick Szary, Ph.D.		8. Performing Organization Report No.
9. Performing Organization Name and Address Center for Advanced Infrastructure & Transportation (CAIT) Rutgers, The State University 100 Brett Rd Piscataway, NJ 08854		10. Work Unit No. 11. Contract or Grant No. 256-RU2806
12. Sponsoring Agency Name and Address New Jersey Department of Transportation PO 600 Trenton, NJ 08625 Federal Highway Administration U.S. Department of Transportation Washington, D.C.		13. Type of Report and Period Covered Final Report 10/2010-1/2012 14. Sponsoring Agency Code

15. Supplementary Notes

U.S. Department of Transportation/Research and Innovative Technology Administration
1200 New Jersey Avenue, SE
Washington, DC 20590-0001

16. Abstract

The services of the Center for Advanced Infrastructure and Transportation (CAIT) were engaged by the Bureau of Freight Services: Trucking (BFST) of the New Jersey Department of Transportation (NJDOT) in order to promote data quality in the SAFETYNET database and maintain New Jersey state rating in an acceptable level ("green") for all defined State Safety Data Quality (SSDQ) measures through consistent monitoring. NJDOT-BFST and CAIT undertook to perform this quantitative enhancement investigation as a response to the Federal Motor Carrier Safety Administration (FMCSA)–developed methodology for SSDQ, which rates states every month on certain pre-defined qualitative (SSDQ) measures. It was thus the endeavor of this program to consistently maintain high ratings for the state of New Jersey in all these measures.

17. Key Words SafetyNet, Trucking Services, commercial vehicles, data quality	18. Distribution Statement		
19. Security Classif (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No of Pages 12	22. Price

Form DOT F 1700.7 (8-69)

ACKNOWLEDGEMENTS

The author wishes to express appreciation especially to Paul Truban and Shari D. Leichter from NJDOT-BFST in carrying out this research project, without whom this research would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

The FMCSA is responsible for ensuring safety in Commercial Motor Vehicle (CMV) operations through strong enforcement of safety regulations. It is a data-driven organization and to this end maintains the Motor Carrier Management Information System (MCMIS), which contains census, crash, inspection, safety audit, and compliance review data for carriers. Further efforts on its part include continued funding to states to improve the collection and analysis of CMV crash data, the development of the State Safety Data Quality (SSDQ) map, and other initiatives to assist states in improving data reported to FMCSA. The FMCSA developed the SSDQ methodology to compare data quality of state-reported crashes and roadside inspection data to a standard defined by FMCSA, as well as having developed a mapping tool to visualize state-reported data. The map depicts the overall data quality for each state in one of three rating categories—good, fair, and poor. The underlying rating system that is visually depicted serves as an incentive for states to improve their crash and inspection data. To obtain an overall state rating, nine SSDQ indicators under two categories (crash and inspection) are measured, and states receive qualitative overall scores of "good," "fair," or "poor". A state with at least one "good" crash measure, one "good" inspection measure, and no "poor" measures receives a "good" rating. A state with only one "poor" measure will receive a "fair" rating, and any state with two or more "poor" measures will receive a "poor" rating. These nine SSDQ indicators are:

Crash measures

1. Crash record completeness
2. Non-fatal crash completeness
3. Fatal crash completeness
4. Crash timeliness
    - Crash accuracy. In addition, there is an overriding indicator (crash consistency), which evaluates the percentage of state-reported non-fatal crash records.
5. Inspection measures
6. Record completeness
7. VIN accuracy
8. Timeliness
9. Accuracy

The task of this maintenance for the state of New Jersey is the responsibility of the NJDOT, specifically the BFST. Based on SSDQ measures, New Jersey was rated overall "green" as a consequence of consistent monitoring by the BFST staff; however, it obtained a steady yellow" rating from FMCSA in crash accuracy. As of October 2010, the compliance rate for "crash accuracy," the most problematic measurement criterion, stood at 92 percent; in order to obtain a "good" rating, a state must achieve a minimum of 95 percent compliance for all SSDQ measurement criteria. The "crash record completeness" had also gone down at the time from 93 to 92 percent. Although, still rated "green," if left unchecked, this particular parameter could also have brought down

the overall data quality rating for the state.

With this consideration, the primary objective of this project was to promote data quality in the database and maintain New Jersey's rating in an acceptable level ("green") for all defined SSDQ measures through consistent monitoring. The project also provided recommendations for the future enhancement of the project and for consistent monitoring of data quality. Hence, the overall goal of this effort was to assist NJDOT in obtaining a "good" rating, which equals a minimum of 95 percent compliance for all SSDQ measurement criteria.

THE WORK AT-A-GLANCE

The SAFETYNET database maintained by the NJDOT-BFST, stores in excess of 8,000 records on crash data, and up to 50,000 records on inspection data per year. The data maintained in this database currently provides decision making support to the BFST. SAFETYNET consists of two major modules; crash and inspection. For each module data is compiled through different resources. The Commercial Motor Vehicle (CMV) crash event data elements are sent to the BFST from the Bureau of Safety Program (BSP) in an ASCII text file format. At the BFST, the ASCII file is uploaded into SAFETYNET. This process creates a log file that summarizes the results of the import. This file is then reviewed for errors and warnings that may have occurred during in the import of the data. If there are fatal errors listed in the log file, the ASCII text file is corrected using the SAFETYNET Crash File Editor, and the data merging process is rerun to import these corrected records.

If there are no fatal errors listed in the log file—or if they have previously been corrected—the SAFETYNET operator checks for any warnings listed in the log file. In this case, a set of queries is run in an attempt to provide the information missing from the crash records. Queries run by the CAIT staff included those for blank GVWR, blank cargo body type, and blank commercial driver license class. All of these records are then corrected in SAFETYNET. After all errors are corrected, the SAFETYNET matching process is conducted on each new record. Those carriers not automatically matched during this process are searched in the local SAFETYNET census file to allow for a match. After this process is completed, an export file is created by SAFETYNET and is electronically transferred to the MCMIS FTP site.

Similarly, the inspection process starts with collecting data either through Aspen software or a paper inspection form. If the electronic data collection is feasible, then the inspector collects and sends the data to the SAFER database. The staff at the BFST office downloads the inspection records from the SAFER inbox and imports the data into SAFETYNET. After each SAFETYNET import process, a log file is created by the system and is reviewed for errors. If there are fatal errors in the log file, then the records are not imported to SAFETYNET, and the BFST office emails the log file to the New Jersey State Police (NJSP) headquarters. The NJSP review the log files for errors and

warnings. If there is a recurring warning, this matter is discussed with the inspector. If the log files still contain fatal errors, the inspector is asked to correct the ASPEN record. Then, the Aspen file is sent to the SAFER inbox again. If no errors are found after the inspection record is downloaded from the SAFER inbox into SAFETYNET, the "matchware" tool is utilized. If the carrier information is not matched during this process, this information is entered into the local SAFETYNET census file to allow for a match. The SAFETYNET export file is then electronically transferred to the MCMIS FTP site. In essence, the process of "matching" involved cross-referencing information on a particular commercial crash, so that all the vital "signs" in that crash match or add up. It is a process of filling in the missing bits of information in the SAFETYNET database by making use of the following databases at the disposal of the NJDOT-BFST:

- SAFER
- Query Central
- A&I (Analysis and Information Online)
- CDLIS (Commercial Driver's License Information System)
- MCMIS
- VIN Decoder (which utilizes Vehicle Identification Number (VIN)

This missing information could include an intrastate carrier (which do not carry USDOT numbers) wrongly listed with a USDOT number, interstate carriers incorrectly marked as intrastate carriers, crashes with blank zip codes, crashes with passenger vehicles incorrectly marked as hazardous material carriers, etc. Each of these listed databases had its own strengths and weaknesses, particularly specializing in a unique part of the information whole while being deficient in others (e.g., VIN Decoder was thorough in terms of co-relating VINs to the probable gross vehicle weight ratings (GVWR) of that class of vehicles). Used together, these various databases combined to form a powerful tool in the hands of the practitioners. The goal was to exploit the information available in one database to discover the missing information in SAFETYNET, thereby in the end correctly tallying up the USDOT number, VIN, license plate number, commercial driving license (CDL) number, GVWR, driver's name, etc. for any one commercial carrier. This information could also be had from the inspection module of SAFETYNET listed above or can involve directly calling the carriers, their insurance companies, or the drivers themselves. Google searching was also a powerful tool in this endeavor. More attention was obviously focused on the procedure of crash accuracy, since this measure depicted lower scores in comparison with other measurement ratings.

Another task that affected both crash accuracy and completeness was researching and answering DataQs. The DataQs system is an electronic means for filing concerns about

federal and state data released to the public by the Federal Motor Carrier Safety Administration (FMCSA). Through this system, data concerns are automatically forwarded to the appropriate office for resolution. The system also allows filers to monitor the status of each filing. Any user can enter data challenges into the system, including general public users, commercial drivers, motor carriers, FMCSA/state agency users, and FMCSA administrative level users. Commercial drivers may file challenges to their commercial driver data. To obtain their data, commercial drivers may request the information via the pre-employment screening program (PSP) or Freedom of Information Act (FOIA) websites. Motor carrier users may file challenges to data found on their carrier profile, SAFER information, and other online records. The DataQs system is to be used to challenge data issued by FMCSA. Information disseminated by the FMCSA includes data reported to FMCSA through the requirements of federal and state programs. With respect to crash and roadside inspection data, the MCMIS documents the occurrence and results of these events as reported by the states. Any challenges to data provided by state agencies must be resolved by the appropriate state agency, in this case NJDOT-BFST. Once a state office makes a determination on the validity of a challenge, FMCSA considers that decision as the final resolution of the challenge. FMCSA cannot change state records without state consent. Hence, in keeping with this process, the CAIT employees embedded with NJDOT assisted the BFST staff in investigating these challenges, and based upon their inquiries, the BFST staff either upheld the data or corrected the logs in SAFETYNET, thereafter informing the concerned party of the final decision (also accomplished via the DataQs website). Needless to say, this was an extremely thorough, exhaustive, and at times deductive procedure, demanding a good deal of resource time from both the BFST and CAIT resources placed with NJDOT-BFST. It may also be pertinent to add here that this task was both extensive and intensive in its concentration on the execution of the above mentioned requisite procedures and recognized all elements addressed and modified by the BFST staff. To this end, two CAIT employees were working full time on site at NJDOT, Trenton, New Jersey, in broadly speaking to the following two tasks:

- **Learning and training** on the databases used for maintaining crash accuracy records and crash database completeness (listed above). Although learning was a constant exertion, this task was for the most part accomplished in the first half of the project's life-cycle.

- Performing constant **data quality monitoring, assessment, and improvement**. While monitoring was a constant task, assessment was relegated towards the last quarter of the project along with the recommendations on further improvements.

Of course, the continuous collaboration and help of the BFST staff was critical to understanding the databases involved, learning the procedure of downloading crash and inspection data, error handling, inaccuracy recognition, and finally uploading of the corrected and matched data. As an example, many of the databases listed above are maintained by the FMCSA and require separate log-in requisition. This permission was

applied for and obtained by BFST, and until such authority was received by the CAIT employees, they used the privileges of their BFST hosts.

CONCLUSION AND RECOMMENDATIONS

The project could be easily classified as completely successful in achieving its outlined goals, as is amply clear from figures 1 and 2. Figure 1 shows the standing of New Jersey in SSDQ measures as of October 2010, when the project began. It shows crash accuracy at yellow ("fair"), with inspection record completeness and VIN accuracy records being unavailable.

New Jersey: Overall State Rating

Monthly Results *	MCMIS Run Date	UPDATED Overall State Rating	State Data Quality Measures									Overriding Indicator †
			Crash					Inspection				
			Crash Record Completeness	Non-Fatal Crash Completeness	Fatal Crash Completeness	UPDATED Crash Timeliness	Crash Accuracy	New! Inspection Record Completeness	New! Inspection VIN Accuracy	UPDATED Inspection Timeliness	Inspection Accuracy	Crash Consistency
Oct '10	10/22/2010											No Flag
Sep '10	9/24/2010											No Flag
Aug '10 *	8/27/2010							N/A	N/A			No Flag
Jul '10 *	7/23/2010							N/A	N/A			No Flag
Jun '10 *	6/25/2010							N/A	N/A			No Flag
May '10 *	5/21/2010							N/A	N/A			No Flag
Apr '10 *	4/23/2010							N/A	N/A			No Flag
Mar '10 *	3/26/2010							N/A	N/A			No Flag
Feb '10 *	2/19/2010							N/A	N/A			No Flag
Jan '10 *	1/22/2010							N/A	N/A			No Flag
Dec '09 *	12/18/2009							N/A	N/A			No Flag
Nov '09 *	11/20/2009							N/A	N/A			No Flag
Oct '09 *	10/23/2009							N/A	N/A			No Flag

*Monthly Results were evaluated using Rating Criteria prior to September 24, 2010.

Figure 1: New Jersey overall standing in Oct., 18 2010 (project start)

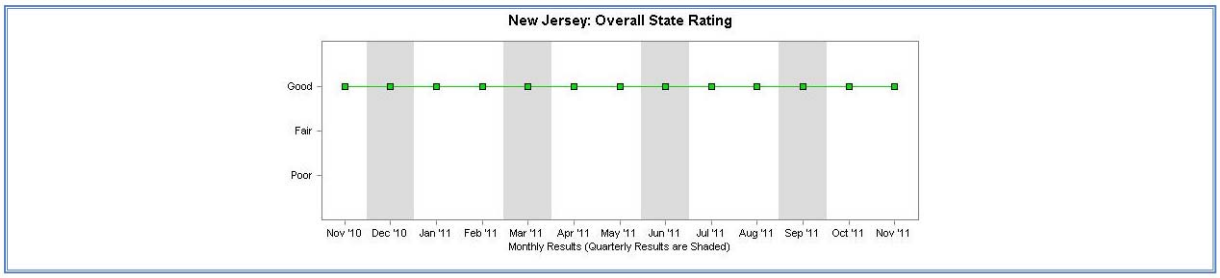
On the other hand, figure 2, shows New Jersey in "green" for all the measures, illustrating the extent of the project's success. In terms of numbers in SAFETYNET:

- The non-match crash data had around 300 entries, which were reduced to 60.
- The number of entries missing vehicle information from crash data was around 850, which was reduced to less than 200.
- The number of blank vehicle license number entries was 330 and was reduced to zero.

New Jersey: Overall State Rating Report

Monthly Results as of: **November 18, 2011**

Considers all nine SSDQ measures and the Overriding Indicator, except measures with a rating of "Insufficient Data." States receive an overall score based on ratings in each of the measures and the Overriding Indicator. (View Methodology)



New Jersey: Overall State Rating

Monthly Results	MCMIS Run Date	Overall State Rating	State Data Quality Measures									Overriding Indicator †
			Crash					Inspection				
			Crash Record Completeness	Non-Fatal Crash Completeness	Fatal Crash Completeness	Crash Timeliness	Crash Accuracy	Inspection Record Completeness	Inspection VIN Accuracy	Inspection Timeliness	Inspection Accuracy	Crash Consistency
Nov '11	11/18/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Oct '11	10/28/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Sep '11	9/23/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Aug '11	8/26/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Jul '11	7/22/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Jun '11	6/24/2011	●	●	●	●	●	●	●	●	●	●	No Flag
May '11	5/20/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Apr '11	4/22/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Mar '11	3/25/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Feb '11	2/25/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Jan '11	2/10/2011	●	●	●	●	●	●	●	●	●	●	No Flag
Dec '10	12/17/2010	●	●	●	●	●	●	●	●	●	●	No Flag
Nov '10	11/19/2010	●	●	●	●	●	●	●	●	●	●	No Flag

Figure 2: New Jersey overall standing as of Nov. 18, 2011 (project end, all green)[1]

Although the work carried out at NJDOT ensured a "green" overall rating for the state of New Jersey, the following recommendations are made so that future slippage of these vital statistics can be avoided:

- Due to the amount of work entailed, as already described above, it is recommended that an employee should be committed to this effort full time in the future.

- The FMCSA in conjunction with the U.S. Department of Transportation (USDOT) organizes many nationwide conferences year-round focused on the maintenance of these vital crash and inspection statistics with the aim of promoting road safety all over the United States. At least the employee responsible for the SAFETYNET effort should be provided with all possible funds and resources so as to enable his/her smooth attendance in these workshops, as an exchange of information and ideas with different practitioners of the same field from other states can be extremely valuable.

- Although the SAFETYNET system is efficient, a lot of the work performed is manual in nature. In this digital age, it's only a matter of time before the information overload overtakes the human ability to input, correct, and upkeep. Hence, with eyes towards the future, a move should be made towards

1 http://ai.fmcsa.dot.gov/DataQuality/dataquality.asp?redirect=overall_s.asp&ns=S&state=NJ&display=M&i=9

automation. The information technology market has some elegant solutions to offer with entire companies committed to computerization of SAFETYNET-like databases (e.g., Oracle and SAP). Some important groundwork has already been done at CAIT with the development of the CISS (Crash and Inspection Safety System) application.

- A majority of the SAFETYNET work executed in the NJDOT-BFST is performed by the respective state police organizations in other states, as the state police records the first information reports for both inspections and crashes. However, as this task is currently the responsibility of the NJDOT and the move towards shifting it to NJSP seems unlikely in the near future as well, it is highly advisable to promote better cohesion between the two departments. Notwithstanding the modalities of such an arrangement, it should also include the Bureau of Safety Program (BSP), as they process a lot of the raw data that takes time to trickle down to the NJDOT-BFST.