Connecticut Crash Data Improvement Program

Final Report

Prepared for:

STATE OF CONNECTICUT

By

Federal Highway Administration
Office of Safety
Under Contract DTFH61-10-D-00022
Task Order T-10-004

Prepared By:

Robert A. Scopatz, Ph.D.
Data Nexus, Inc

Jack Benac
Jack D. Benac, LLC

Nancy Lefler
Vanasse Hangen Brustlin, Inc (VHB)

Robert Pollack
Federal Highway Administration

Submitted:
May 7, 2012

NOTICE

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

The conclusions and opinions expressed in this document are those of the authors. They do not necessarily represent those of the State of Connecticut, the Connecticut Department of Transportation, or any political subdivision of the state or federal government.

Table of Contents

List of Tables

List of Figures

Executive Summary

In October 2011, the Federal Highway Administration (FHWA) Crash Data Improvement Program (CDIP) Technical Assistance Team (TAT) conducted a comprehensive data quality assessment of the Connecticut crash data and recommended measures to improve the quality of the crash data reporting system. The CDIP is designed to help States develop and improve methods of assessing the quality of their crash data. The CDIP process examines the quality characteristics of timeliness, accuracy, completeness, consistency/uniformity, integration, and accessibility. The TAT uses a systematic approach to assess the crash data. Each step of crash data processing is examined because different personnel and agencies are typically responsible for collecting, processing, input, storing, and distributing crash data. The focus of the assessment is to establish performance measures (data quality metrics) which allow a State to assess how well each component of the crash data system functions. As strengths and weaknesses are discovered, the State is better able to address process deficiencies. This report provides Connecticut with recommendations on how to develop performance measures and improve performance for each of the components of crash data quality.

The assessment began with the CDIP TAT review of documentation describing the existing State traffic records practices. Connecticut's documentation consisted of the State's 2007 Traffic Records Assessment, a CDIP pre-visit questionnaire completed by Connecticut Department of Transportation (CTDOT) personnel, planning documents (including the Highway Safety Improvement Program [HSIP], and the Traffic Records Strategic Plan), and the State's grant submissions to the National Highway Traffic Safety Administration (NHTSA) under the Section 408 program. The TAT also received crash data files from CTDOT for calendar years 2006 through 2008. The TAT analyzed these files to further evaluate data quality and provide examples of data quality metric calculations. The TAT conducted onsite interviews with crash database administrators and other key individuals at CTDOT, the Connecticut State Police (CSP), the University of Connecticut (UConn), the Department of Public Health (DPH), the Stamford Police Department, two Metropolitan Planning Organizations (Council of Governments for the Central Naugatuck Valley [COGCNV], and the Capitol Region Council of Governments [CRCOG]), and other individuals to discuss and validate the TAT's understanding of Connecticut's current crash data processes, quality performance measures, and analyses.

The TAT is grateful for the generous contributions of Maribeth Wojenski, Joe Cristalli, and Thomas Maziarz from CTDOT for expert guidance, planning assistance, logistical arrangements, and support in conducting this CDIP. The TAT would also like to thank Amy Jackson-Grove (Administrator), Michelle Hilary (Assistant Administrator), Robert Ramirez, Robert Turner, David Nardone, of the FHWA Connecticut Division Office, Keith Sinclair of the FHWA Resource Center, and Mario Damiata of the NHTSA Region 1 Office for participating in the CDIP site visit and the executive briefing at the end of the CDIP site visit.

Connecticut is to be commended for the progress they have made in the last several years to improve their crash data system. The CSP has implemented electronic data collection and report submission to CTDOT. Several local law enforcement agencies (LEAs) are using a variety of software packages for collecting crash data in the field. Although none of these agencies are currently submitting data electronically to CTDOT, such submissions are likely to be supported in the near future. Under contract with UConn, CTDOT is developing a new crash data repository that holds the promise of collection and storage of a complete, multi-year crash database along with improved access for users to analytic resources. Roll-out of a Model Minimum Uniform Crash Criteria (MMUCC) compliant, fully electronic crash report for use by LEAs is being pursued by CTDOT and its many safety partners through the auspices of the Traffic Records Coordinating Committee (TRCC). These planned activities are mentioned here for the purpose of setting the stage for the CDIP TAT's recommendations in the remainder of this report. There is much to be hopeful about for the future of Connecticut's crash data system. The TAT has concentrated its recommendations on the plans and data quality metrics that Connecticut will need in order to bring the plans to completion and document the success of the many projects.

CTDOT is the crash data custodian and also serves as the provider of location data and location-based highway safety analysis for the State. Unfortunately, there are systemic barriers to reaching a high level of data quality—chiefly with regard to crash data timeliness and completeness. This has to do with the large backlog of reports to be entered into the CTDOT crash database, and the fact that CTDOT's data entry process results in an incomplete record since only a subset of data elements are entered into the file. There appear to be no significant problems with missing reports (i.e., under-reporting by LEAs) or with overall accuracy; however, this conclusion is not supported by actual measurement (i.e., data quality metrics). Data accuracy is improved through an edit check process during data entry at CTDOT; however, Connecticut's data are of unknown accuracy and overall completeness since the data are not subject to a formal, comprehensive data quality management process that includes relevant performance measures (i.e., data quality metrics). The TAT's recommendations address this need as well.

The recommendations provided in this report are based on the TAT's assessment of Connecticut's current crash database and the deficiencies identified within the current system. The TAT believes Connecticut would be best served by incorporating the CDIP recommendations into its Strategic Plan for Traffic Records Improvement and by addressing specific recommendations in a Business Plan for crash records, as described in this report.

1 Introduction

The State of Connecticut has a relatively loose set of laws defining custodial responsibility for crash reports. Connecticut ***General Statute Section 14-108a*** on the uniform investigation of accident reports states that,

> (a)(1) The Commissioner of Transportation shall prescribe for the Division of State Police within the Department of Public Safety and for each police department and officer and other suitable agencies or individuals a uniform investigation of accident report, in such form as the commissioner shall prescribe, which form shall be followed in filing all such reports.
>
> (2) In each motor vehicle accident in which any person is killed or injured or in which damage to the property of any one individual, including the operator, in excess of one thousand dollars is sustained, the police officer, agency or individual who, in the regular course of duty, investigates such accident, either at the time of or at the scene of the accident or thereafter, by interviewing the participants or witnesses, shall, within five days after completing such investigation, complete and forward one copy of such report to the Commissioner of Transportation. Such report shall call for and contain all available detailed information to disclose the location and cause of the accident, the conditions then existing, the persons and vehicles involved and the names of the insurance companies issuing their automobile liability policies, as well as the enforcement action taken. The Commissioner of Transportation shall forward to the Commissioner of Motor Vehicles one copy of each report of any accident involving a school bus. The Commissioner of Motor Vehicles may inquire into or investigate any accident reported pursuant to this subsection and may request the assistance of the Division of State Police within the Department of Public Safety for such purposes.

The Connecticut Department of Transportation (CTDOT) website has the following statement:

> The Accident Records and Statistics Section of the Office of Inventory and Forecasting is responsible for the codification, maintenance and compilation of motor vehicle traffic accident data.

To our reading, this places the custodial responsibility for crash data management clearly within the purview of the CTDOT. In the past, however, the CTDOT has followed an internal legal interpretation that allowed the agency to pursue a course that did not serve its own needs well, nor did it serve the needs of external users well. CTDOT has shared the available data with

agencies that request it; however, the dataset is inadequate for many purposes because it is incomplete and out of date. The reasons for that earlier decision are no longer relevant, but almost certainly had to do with the cost of creating and maintaining a crash database when each report had to be entered manually into the system. The old way of thinking has been largely eclipsed within the CTDOT based on a top-level commitment to data-driven approaches to safety management, and the recognition that modern technology can make crash data management both affordable and better able to serve the needs of a broadly defined user community – including local agencies and safety advocates from outside of government.

Because of the historical short-comings of the system, it is clear that in comparison to other States' experience, Connecticut crash data are an underutilized resource. The CTDOT file contains only a partial record lacking key data elements that many safety partners would want. There is no archive of scanned images for engineers to access. Location-based analyses are supported because of the ongoing attention to coding crash locations and ensuring adequate linkage to a roadway inventory file; however, this support extends only to the most basic types of network screening. Modern methods requiring more extensive data are not yet supportable under the system at present. Even more worrisome is that the backlog of approximately 14 months represents a barrier to recommended practices in managing data quality, and reduces the utility of the data for safety evaluation. The data simply are too old to be of much use or to be easily improved in accuracy or completeness by the time errors are noted.

CTDOT and its partners in the State are on the verge of creating a modern and fully capable crash reporting system that will serve the broader safety community. CTDOT's new crash database is a modern design. It is capable of storing complete crash reports (including all coded fields plus the narrative and diagram) when those reports are submitted electronically. About 35 percent of reports (those from the Connecticut State Police [CSP]) are in the system in this complete format. The roadway inventory and digital network of roadways are projects that will help CTDOT take network screening and safety program evaluation to the next level – implementing the full range of analytic techniques described in the Highway Safety Manual. The creation of a crash data repository at the University of Connecticut (UConn) will facilitate electronic data transfer and make the data more accessible to a broader range of users. All of these things are hopeful signs that lead the Technical Advisory Team (TAT) to believe that the next five years will see dramatic improvements in the crash data system in Connecticut. The remainder of this report provides specific descriptions of the current status of the system and the TAT's recommendations for improvement.

1.1 Crash Data Processing

While Connecticut statutes related to crash data do not reference the word "custodian," CTDOT fulfills several of the roles typically assigned to the crash data custodian, such as:

- Serves as the official site for submission of reports by law enforcement agencies (LEAs).
- Responsibility to determine the content of the statewide standardized crash report form (Connecticut Uniform Police Accident Report, form PR-1).
- Responsibility to produce summary reports of crash experience in the State.
- Reporting to Federal oversight agencies under the Highway Safety Improvement Program (HSIP) for the Federal Highway Administration (FHWA) (23-USC-148) and for the National Highway Traffic Safety Administration (NHTSA) under the Highway Safety Act (23-USC-Chapter 4, Section 402).
- Maintenance of a statewide crash database.

The first three of these roles are mandated by State law (Connecticut General Statute Title 14, Section 14-108a) (http://www.ct.gov/dot/cwp/view.asp?a=1383&q=259794). The fourth role, reporting to FHWA and NHTSA, is defined by Federal law. The final role, maintaining a statewide crash database, arises from the need to meet CTDOT's responsibilities with respect to monitoring and improving safety on the State's roadways. In order to comply with applicable Federal laws and regulations, CTDOT is nominally recognized by FHWA, NHTSA, and the Federal Motor Carrier Safety Administration (FMCSA) as the official crash data custodian.

CTDOT receives approximately 110,000 crash reports each year from LEAs. Approximately 65 percent of the reports are submitted on paper forms and the remaining 35 percent are submitted electronically. The latter are from the CSP. An unknown number of local agencies are using various electronic crash data collection software, but none of their reports are submitted electronically to CTDOT. Rather, they are printed and mailed to CTDOT.

CTDOT performs data entry on all the paper reports. Only a subset of the fields from the paper crash reports is entered into the CTDOT crash database. CTDOT clerks perform location coding and quality control on all reports, including those submitted by the CSP. The full record of electronic crashes submitted by CSP is retained in the CTDOT crash database, including the narrative and diagram. As such, approximately 35 percent of crash records are "complete" and the remaining 65 percent consist of partial data containing only those fields of importance to CTDOT. This file is not designed to meet the needs of other users outside of CTDOT. It was clear from the interviews during the site visit that many users' needs from within CTDOT are also not met by this truncated file.

Paper reports are not scanned to create an image file for users to access. The paper reports are stored for a year and then purged. The electronic report data can be used to generate an electronic PR-1 image for review. These are maintained indefinitely. Until all reports are collected and submitted electronically, it appears that CTDOT does not have the resources internally to create an image archive based on scans of paper crash reports.

Electronically submitted data are held in a queue until such time as the manual data entry of paper reports has reached the same month and year. The current month being entered is August of 2010—approximately 14 months after the crash event.

CTDOT maintains the centralized crash database for ten years. Data are currently available back to 2003. The Traffic Accident Viewing System (TAVS) can be used to access reports back to 1999.

The UConn Connecticut Transportation Institute (CTI) in the School of Engineering is developing a crash data repository under contract to CTDOT. UConn currently has the CTDOT data subset back to 1995. A new Memorandum of Understanding between UConn and the CSP will provide data directly to the repository rather than having to wait until CTDOT has completed data entry. UConn is not currently accepting data electronically from local LEAs. The concept of operations for the UConn repository includes the possibility of storing complete records for all crashes—beyond the subset of data elements captured by CTDOT. However, this possibility does not apply to reports submitted on paper.

Following a multi-agency review of the current version of the crash report, a project is underway to simultaneously implement a MMUCC-compliant form and provide free electronic data collection software and interface for all LEAs based on TurboTax®. (Note: the TAT was informed that this system is literally based on the software used in the popular TurboTax® program.) It was reported that many agencies that already have electronic crash data collection capabilities are waiting for this new system and form implementation before exploring electronic data submission to either CTDOT or the UConn repository. The timeline for implementation is not definite but a 2013 date was mentioned as the target for rollout.

Data quality is managed at two points in the process from the initial crash event through final posting of data at CTDOT. LEAs generally include a supervisory review of the reports prior to submission to CTDOT. As these reports are entered into the statewide crash database at CTDOT, edit checks are run to validate data types, data value ranges, and logical agreement among two or more data fields. The CSP has implemented all of the CTDOT edit checks in their field data collection software. It was reported that other LEAs have not implemented the edit checks or have implemented them and later turned them off—sending the data to CTDOT without being checked. The CSP system also has the ability to automatically complete driver

and vehicle fields on the crash report form based on input of the driver's license number and plate number, respectively. It is unknown which of the (25 or so) other vendor products in use by local LEAs have similar capabilities. It is known that at least some of these other systems have the capability to check report data for errors, but that some of the LEAs ignore or defeat the error checking.

Because CTDOT currently focuses on about 1/3 of all fields on the crash report form, any suite of edit checks that relies solely on CTDOT's edits is likely to be inadequate for the task of improving data quality for all key fields on the report.

1.2 Crash Data Processing Recommendations

- CTDOT needs to accept and act upon its clearly defined role as crash data custodian. We have noted in the Introduction that it is clear that CTDOT has experienced a change in approach regarding the crash data and that custodial responsibilities are now acknowledged and welcomed by the agency. In a difficult fiscal climate, CTDOT is to be commended for this recognition and encouraged through assistance from other safety partners, including FHWA and NHTSA.

- Establish a centralized crash database including all fields from the crash report form as the custodial record in the State. The Traffic Records Coordinating Committee (TRCC) should assist in determining who has access to sensitive data within the full record. Ideally, this recommended subcommittee would include representatives from agencies that already handle sensitive data, such as CTDOT, the Department of Public Health (DPH), CSP, the Department of Motor Vehicles (DMV), the courts, and others. This subcommittee's role could extend, in the future, to providing similar advice on controlling user access to merged datasets.

- Develop a scanned image archive for all crash reports (paper and electronic) to enable long-term storage for search and retrieval to support safety analyses. The creation of images from electronically submitted reports is a relatively inexpensive and trivial task in comparison to the costs and effort involved in creating and storing scanned images of paper reports. For this reason, we recommend that CTDOT seek a permanent solution for image archiving that is based on the planned move to 100 percent electronic reporting. An interim solution is needed for paper reports, but this should be viewed as temporary and not worthy of capital outlay (in terms of purchasing sufficient scanners to manage the work in house). It would be reasonable to give lower priority to the scanning of existing paper reports than to the creation of an image archive for electronic reports.

- Consider implementing supplemental data entry via a contract in order to reduce the backlog more quickly. In the Business Plan (recommended below), CTDOT is

specifically called upon to produce a plan for reducing the backlog. It was clear during the TAT's interviews and from the pre-visit questionnaires answered by CTDOT that the State cannot maintain the status quo of a large proportion of crashes being entered manually. The resources simply do not exist that could reliably maintain the current manual system for the long term. Clearing the present backlog is important because CTDOT needs the data for 2010 and 2011 to fulfill its role in managing safety improvement. Long term it is absolutely critical that electronic data collection and submission are achieved.

- Maintain data on crash records indefinitely – as long as there is a defined need for the older data. Once the data are in electronic form, the cost of indefinite maintenance is negligible. So long as users have a need for the data, maintaining it in an accessible form should be a trivial undertaking.

- Develop a Business Plan[1] for the future of crash data collection and management. This plan should address five critical needs/projects:

    1. The creation of the crash data repository and specification of its desired functionality/contents.
    2. Development of the digital roadway network in a statewide base map incorporating the CTDOT linear reference system (LRS) as described in the Location section of this report.
    3. Achievement of 100 percent electronic data collection and electronic submission of crash reports to CTDOT and the crash data repository.
    4. Reduce the data entry backlog.
    5. Creation of a Model Minimum Uniform Crash Criteria (MMUCC)-compliant crash report form.

  The Business Plan should include timeframes, milestones, and priorities.

  In follow-up conversations with CTDOT, the desirability of adding the fifth element to this Business Plan was discussed; that of creating a new crash report form that is MMUCC-compliant. The State hopes to accomplish this goal in conjunction with creation of a new electronic data collection system for crashes (through the Capitol Region Council of Governments' [CRCOG] efforts) that could be offered to all LEAs. The Business Plan should thus include statements about this planned form update and the tie-in to 100 percent electronic data collection showing the points of interaction between the two efforts.

[1] The CDIP TAT provided CTDOT copies of business plans from Michigan and South Carolina. These may serve as examples that CTDOT could follow in developing its own Business Plan.

Follow-up conversations with CTDOT also resulted in the realization that the Business Plan should be more fully described in this report. In particular, CTDOT should seek to develop an early version of the Business Plan that describes the vision for the crash records system and sets some overall goals for initiation and completion of the various projects that are described (the five critical projects listed above). As the State progresses toward actual implementation of projects under the Business Plan, the plan should be updated to include detailed descriptions of tasks and activities required to complete each of the critical projects. The Business Plan should be viewed as a "living" document – one that the CTDOT will update and refresh periodically as projects move forward. The level of detail required at the project level need not be overwhelming. The role of the Business Plan is to set down in writing the deliverables and expected due dates for each of the key projects so that managers and partners (such as represented by the TRCC) can track progress and, most importantly, react quickly to events that threaten to impede progress.

- The TRCC should approve the Business Plan and monitor completion. This recommendation has changed slightly since the original Draft Matrix report submitted to CTDOT at the end of the TAT's site visit. Some elements of the original recommendation were rolled into the preceding recommendation describing the desired contents of the plan. However, the recommendation that the TRCC should be given the role of monitoring completion of the Business Plan is retained as a separate item. This is not to imply that the TRCC should become the de facto project manager—project management is the responsibility of the individual agencies responsible for each project. CTDOT has overall responsibility for those projects which it is funding (either directly or with grant funds administered by that Department). Further, we recommend elsewhere in this report that the State hire a full-time project manager responsible for the crash system's improvements. The TRCC does have an oversight role and, because of its broader representation of stakeholders, is an ideal venue for discussions related to users' needs for access to the crash data.

- Encourage local LEAs that have already implemented electronic data collection systems to share their data with the UConn data repository directly. UConn has the technical expertise to develop necessary interfaces with the variety of systems in use by LEAs. If UConn can obtain the data electronically, they would then be able to run edit checks and create a database that could be shared with CTDOT when completed. This should result in savings for CTDOT in that it would not be required to develop the multiple interfaces itself, and because it would receive data that have already been validated. UConn, in turn, should be encouraged to develop a data transfer protocol and help LEAs to implement a standard data transfer.

- Ensure that all LEAs' electronic data collection systems incorporate a standard set of edit checks so that data submitted to the centralized crash records system meet a minimum quality standard. Another benefit of establishing this standard is that it will facilitate electronic data transfer from the LEAs to CTDOT. The edit check standard set should include, but not be limited to, those already implemented within CTDOT's crash records system. We recommend that a subcommittee of the TRCC be established to define the edit checks to cover the complete range of data elements on the crash report form. This should be a separate subcommittee from that tasked to advise on user access issues. The task of this subcommittee would be to review all of the existing edit checks and develop new ones as needed. The edit check definitions should be shared with all LEAs and their crash data collection system vendors.

- As the proportion of crash data received electronically increases, plan for transitioning the CTDOT crash data entry staff to a quality assurance role. This recommendation stems from experience in other States. States that are able to retain their data entry clerks and subsequently transition some or all of this expert-level staff to data quality management are better off than those that do not do so. The reason is that with extensive automated edit checks, there is no good way to ensure the quality of data in terms of its internal consistency and logical agreement among multiple variables. The computer cannot "read" the narrative and diagram, nor can software effectively compare the "picture" of the crash that emerges from the officer's narrative and diagram to the data that are coded in the remainder of the form. Frequent users of crash reports, especially for engineering purposes, will note that only by reviewing the narrative and diagram can they be sure that critical information such as the pre-crash movements of vehicles and the type of crash be reliably understood (and often corrected) in the crash report. The TAT strongly recommends that the staff in the current crash reporting operation be gradually transitioned to a higher-level position and tasked with reviewing and improving the quality of data in the crash reports. Such data quality management is critical to achieving a reliable dataset for network screening. Because we also envision that the crash data will be nearly day-current, the efforts of these data quality managers need not focus on correcting errors themselves. Rather, they should be able to return erroneous crash reports to the LEAs with a note explaining the error detected and asking for the officer and agency to resubmit a corrected copy. Ideally, the resubmission of such reports would be tracked as well so feedback can be provided to LEAs on both their overall data quality and the timeliness of their efforts at correcting reports.

2 Crash Location Process

CTDOT maintains a route/milepost LRS for State routes. CTDOT maintains a separate town/road number/milepoint system for local routes. The CTDOT Geographic Information System/Computer Systems Section in the Office of Systems Modeling and Forecasting is developing a statewide map that incorporates the LRS for all public roads. This is a multi-year project that is viewed as critical for the ability to standardize location referencing for crashes in a way that is compatible with the State's roadway inventory and traffic volume data. The system development has suffered from delays, interagency coordination, and procurement problems. There is no definite timeline for completion of the digitized roadway network (the base map).

CTDOT crash data entry clerks manually review the location information on every crash report and assign a location code. For CSP-provided crash reports, the location data are supplemented by global positioning system (GPS) coordinates collected by the officer in the field.

The CSP uses NexGen ® software for field reporting. Their implementation includes a map-based location tool which can aid the officer in completion of location fields on the crash report and other forms. This tool does not qualify as a "smart map" because it does not automatically complete all relevant location fields on the form, nor does it update the GPS coordinates based on the mapped location selected by the officer.

2.1 Crash Location Process Recommendations

- Implement the Business Plan components related to CTDOT's improved methods for location coding as planned for the digital roadway network/base map project. This project's plan should include definite milestones and deliverables so that its completion can be tracked by the project managers and upper level management within CTDOT. The TRCC should also maintain an awareness of progress on this project through periodic briefings. This should elevate the priority of this project and help to ensure that it is completed in a timely fashion. Much of the State's future ability to locate crashes in an automated manner and link crash data to roadway inventory records depends on the completion and maintenance of the digital roadway network. The TAT recommends that this project be considered a top priority within CTDOT's safety-related and asset management functional areas.

- Develop automated methods of location coding for use during crash data entry. Track the success of automated processes separately from the overall measure of successful linkage described in the data integration portion of this report. The current processes for coding locations of crashes is manual and labor intensive. It adds significant time to the delay in completing a crash record in the CTDOT database and contributes to the overall backlog of crash reports. Even with electronic submission of crash reports, if the location coding

process remains largely manual in nature, the staff in the CTDOT crash section will be spending too much of their time adding this information to the record in lieu of managing data quality. It is also clear that automated processes can only go so far. If the LEAs are not submitting valid location information in the expected format, no automated process will be able to overcome the errors and produce a valid location code. For this reason, the TAT also recommends (in the following recommendation) that tools be developed for LEAs use in the field to ensure accurate location information is gathered from the onset. Again, monitoring the success of the automated process for assigning location codes will be important because it will be the primary indicator of success both for LEAs' data collection and for reduction in staff time devoted to the manual processes of location coding.

- Implement a full "smart map" capability for CSP and other LEAs' crash data collection systems. Incorporate the to-be-developed CTDOT base map as available. While the standardized base map is an important component of the long-term vision for this tool, we recommend that the smart map be deployed in advance of having the complete base map if the latter project has a timeline for completion that is longer than about one year. This is because the smart map technology will help improve location data even without a CTDOT-compliant base map. Other sources of digital base maps will be "close" to what the eventual CTDOT base map shows and will help the LEAs and CTDOT produce a crash database that is more accurate and less costly to collect and create. Eventually the larger benefit of collecting records in the field that automatically match locations to the CTDOT base map is a goal worth pursuing. In this scenario, an officer will point-and-click at the precise location of the crash – zooming in on a map to show the location down to a level of detail not currently possible with a system based on street names or even the CTDOT's LRS. The smart map can then complete all location-relevant fields on the crash report form and provide spatial coordinates and the corresponding LRS value of the location. This type of system eliminates problems with typographical errors, alias names for roadways, and use of uncommon variants (such as "AV" for "avenue") that can make it difficult to automate the process of location coding. Once the smart map incorporates the CTDOT's own base map and LRS, CTDOT should see a drastic reduction in the amount of time and effort it spends on location coding crashes. At the same time, CTDOT should see a marked increase in the success of assigning location codes. Again, we recommend that the savings in staff time be devoted to quality assurance functions performed by the same staff in CTDOT's crash section.

3 Data Quality Assessment

There are numerous business edits and cross-field checks for logical agreement built into the CTDOT crash data management processes. These are run as part of the data entry/data acceptance processes for creation of the CTDOT crash database. Only the subset of data elements captured by CTDOT are checked in this manner. LEAs using field data collection software are provided with the list of CTDOT edits, but not all have implemented or maintained the edits as part of their field data collection system.

CTDOT Bureau of Policy and Planning produces a biennial summary of the latest year's and ten-year trend in crash experience—the *Traffic Accident Facts* report. The most recent report is for 2008 (http://www.ct.gov/dot/lib/dot/documents/dpolicy/ctaf/ctaf.pdf) with trend data for 1999 through 2008. Through the process of producing this report and generating other summary analyses, comparisons are produced that can identify large year-to-year changes in the data. There were no other periodic analyses discussed that would provide an indication of data quality.

3.1 Data Quality Assessment Recommendations

- Establish a comprehensive, formal data quality management process for crash data. The TAT recognizes that this recommendation may not be fully implementable given the current data entry backlog. Therefore, we recommend the TRCC begin now to work with CTDOT and UConn to develop a plan for data quality management that recognizes the migration to fully electronic data collection and data submission. A comprehensive, formal data quality management program is typically described as including the following functions and activities:

    - Data quality metrics for overall timeliness, accuracy, completeness, consistency, integration, and accessibility.
    - Daily quality monitoring reports for use by the crash data managers.
    - Monthly summary data quality reports for use by upper level managers and reporting to users and oversight (including the TRCC).
    - Frequent reviews of data quality by a subcommittee of the TRCC with strong user representation.
    - Procedures for returning erroneous reports to the originating agency and officer.
    - Tracking of returned reports to ensure that they are corrected and returned in a timely fashion.
    - Continuous auditing of data quality.

  - Periodic review by independent auditors.
  - Feedback mechanisms that report agency-level data on performance indicators of timeliness, accuracy, completeness, and consistency. These data should be reported directly to the individual LEAs.
  - Mechanisms for routinely updating training for law enforcement officers based on the data quality audit results.

- Design an expanded *Accident Facts* report that meets the needs of users beyond engineering areas to include behavioral safety program areas and other key safety partner's uses for crash information. This recommendation requires an expanded dataset (beyond the current CTDOT subset of data elements) to be made available to analysts. Examples of excellent Crash Facts reports may be obtained from several States through the NHTSA Regional Office. In particular, the goal of the redesigned report should include detailed examination of crash experience in each of the major program areas recommended by NHTSA and as represented in the State highway safety office functional area. These areas focus primarily on the human contribution to crashes and give rise to the need for detailed data on occupant protection (belts, helmets, and child protection devices), impaired driving, young drivers, older drivers, motorcycle safety, large truck and bus safety, speeding, bicycle/pedestrian safety, aggressive driving, distracted driving, and a number of other issues. It is also reasonable to examine the report's contents within the TRCC. There are likely to be users whose needs could be met easily with an annual summary of statistics related to the driving population and crash involvement or outcomes.

4 Timeliness

Overall timeliness is expressed as the number of months from the end of a calendar year to the date when those data are available for analysis. The original PR-1 and PR-2 – the fatal crash supplement –are required to be submitted within five working days, if applicable.

Currently the backlog is 14 months. This fact alone is responsible for the most significant barriers to data quality improvement in Connecticut because feedback to law enforcement is ineffective with such a significant time lag.

There are no measures of the timeliness of crash report submissions from LEAs. CTDOT accepts paper reports and does not record the receipt date. CSP submits all their crash reports electronically (approximately 35 percent of statewide crashes). In addition, CTDOT does not have a unique identifier for all LEAs.

CTDOT does not maintain a process flow chart for the overall crash processing system and each of its component processes. It would be worthwhile to develop an annotated process flow diagram for the complete crash processing system (both paper and electronic) and the component processes.

4.1 Timeliness Recommendations

- Develop a metric calculating the time between the crash event and data availability on the statewide crash database (inclusive of all component processes). This should be recorded for each crash event, and the average value over all crashes should be reported as the average reporting delay. Report this value for the entire crash records system and individually for each LEA. This implies that the data set must include the LEA identifier. This report could take the form of a ranking report as shown in Table 1.

Table 1. Example Ranking Report for Timeliness of Law Enforcement Agencies' Crash Report Submissions.

Rank[1]	Agency Name	# Reports Submitted		Average Days to Submit	
		This Period	YTD	This Period	YTD
#	Ansonia PD	#	#	#	#
#	Avon PD	#	#	#	#
#	Berlin PD	#	#	#	#
#	…	#	#	#	#
#	…	#	#	#	#
#	Wolcott PD	#	#	#	#
#	Woodbridge PD	#	#	#	#

1. Rankings are based on the year-to-date average days to submit reports: 1 = best.

Note that the TAT was not able to compute values for this sample report because there are no LEA identifiers in the CTDOT database and because there are no time stamps that would support measurement of submission timeliness.

- Develop a unique identifier for each LEA. The unique identifiers should be added to the electronic reports as collected by LEAs and submitted to CTDOT. For paper reports, CTDOT will need to enter the identifiers manually, but the need for this process should disappear once the reports are received electronically.

- Develop a method to timestamp crash reports when received (on paper or electronically) at CTDOT. For paper reports, this requires that someone stamp each report as it is received. In addition, the CTDOT database must be modified to allow clerks to enter the date of receipt. For electronic reports, the date of submission should be recorded electronically as the data are received into a "pending" folder at either CTDOT or UConn (whichever is most efficient for receipt of electronic reports). The date the records are posted to the database (for both paper and electronic reports) should also be captured automatically by the system (either at CTDOT or at the UConn repository – whichever is most efficient). NOTE: Because of resource constraints in CTDOT, it may not be feasible to implement a timestamp process for all paper reports received by CTDOT today. We suggest deferring this recommendation until the staff levels available are able to handle the incoming workflow (e.g., as a result of a greater proportion of crash reports being received electronically so that the remaining paper reports can be managed using this recommended timestamping process).

- With the implementation of the crash data repository at UConn, develop a metric calculating the time between crash event and date of submission to CTDOT (or to the repository – whichever is most efficient to collect). The average reporting days over all crashes and for each LEA should be reported on a periodic basis. This implies that the data set must include the LEA identifier.

- Develop timeliness metrics for component processes, including error correction, location coding, and other key processes. These should be reported on a periodic basis to all stakeholders including the Connecticut TRCC. As with the date of submission, each of these additional timeliness metrics will require that the relevant date be captured in the crash record at CTDOT or UConn – whichever is most efficient. The system can then be programmed to automatically calculate overall timeliness and the timeliness of each component process from initial submission through data validation, location coding, and final posting of the record to the production database.

5 Accuracy

There are no measures of the accuracy of crash data. CTDOT does maintain a crash data edit table and requires the edits to be included in field electronic crash collection systems. These edits are focused on the 1/3 of crash data fields that CTDOT incorporates into their centralized database. There are no established edits for the other 2/3 of the fields on the form. Additionally, the edit tables have not been reviewed recently and the CSP, as well as other LEAs, indicated difficulty using them in their systems. As noted earlier, some LEAs have turned off the edits in their field data collection systems.

There were some indications that the edits specified by CTDOT and those implemented in the CSP's software may conflict. One concern is that there is a problem with LEAs having to choose which source documents to use when training their officers and establishing edit checks. The crash report data collectors' manual has not been updated in decades.

Figure 1 displays data showing that the edit checks that are in place appear to be effective. Compared to other States, Connecticut has a relatively low proportion of records in which there is an obvious disagreement between the time of day recorded for the crash event and the lighting condition coded. The figure shows that there are relatively more errors where daytime crashes are coded as happening in the dark than there are nighttime crashes coded as happening in daylight. However, the overall error rate is very low.

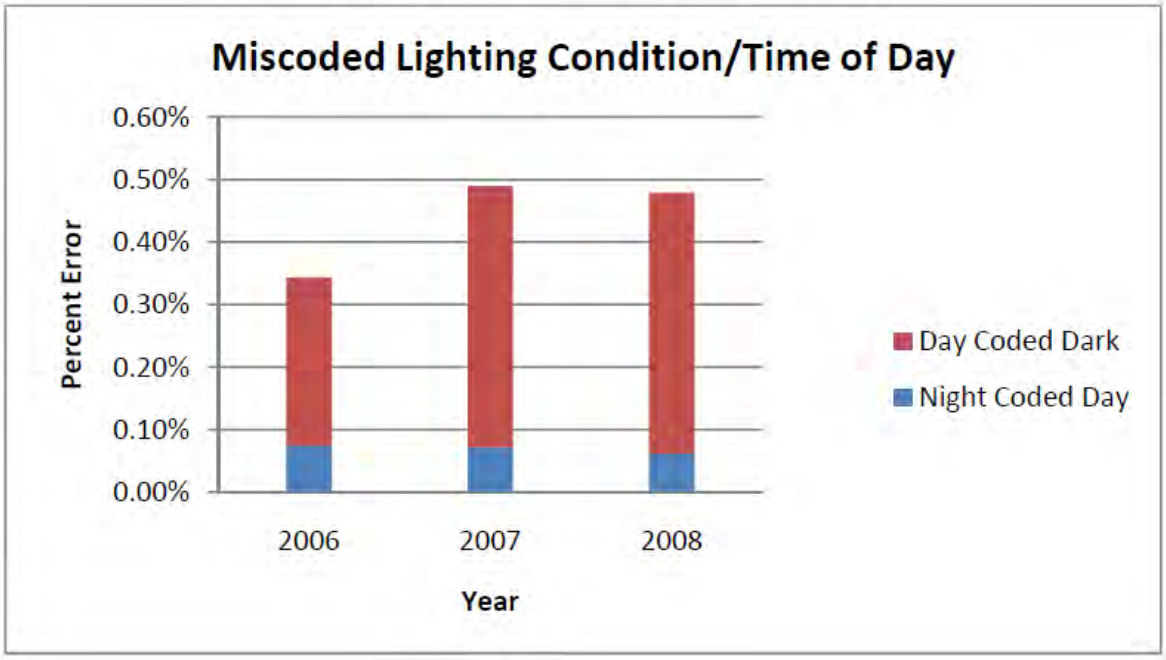

Figure 1. Comparison of Time-of-Day and Lighting-Condition.

5.1 Accuracy Recommendations

- Develop a working group including representatives from CTDOT and law enforcement to review the existing edit table. Update and expand the edits to assess crash data accuracy.

  We offer the following suggestions as a preliminary list of important cross field logical edits, but add the caveat that this set may not meet the most pressing needs of key users. There is no substitute for user-centered review and support.

  Suggested cross field logical edits

  - **Highest vehicle number does not equal number of vehicles reported.** This edit check compares a field on the form (i.e., the number of vehicles involved in the crash) against the highest vehicle number (i.e., the last vehicle reported in the crash). If these do not match, there is either missing data or the reported number of vehicles is incorrect.

  - **No units recorded as motor vehicles** (i.e., it is not a crash involving a qualifying motor vehicle). This is a cross field edit in that it requires examination of all unit-level records to ensure that each crash includes at least one motor vehicle, otherwise the crash is not reportable.

  - **Sequence of Events inconsistent with Unit Type and/or number of units recorded.** This edit checks for whether the unit types and number implied in the vehicle sequence of events codes are recorded in the crash report. For example, if one unit has coded a collision with a pedestrian as part of the sequence of events, then there should be a pedestrian recorded as a unit in the crash. Similarly, if there is one unit with a collision with a motor vehicle in transport coded, then there must be at least one other motor vehicle unit recorded in the crash.

  - **Incorrect crash severity.** The crash severity should match the code of the highest level injury/fatality to any involved person. This check may be skipped if crash severity is a calculated field. If crash severity is entered by the officer, a check of the injury codes for all person records should be run to capture the highest level for any involved party.

  - **Report date/time precedes crash date/time.** The crash date/time should always be earlier than the report date.

  - **Vehicle Body Type and Vehicle Configuration do not match.** For example, if the vehicle type is 08, the expected code for Commercial Motor Vehicle Configuration would be 10 (Bus/Large Van) or 11 (Bus (seats for more than 15)) or, perhaps 77 (other).

  - **A commercial vehicle type is coded, but commercial vehicle data fields are not completed.** If Vehicle_Body_Type code 20 (Medium/Heavy Trucks) is selected,

other fields should be completed including Cargo_Body_Type, Comm_GVWR/GCWR, and Commercial_Motor_Vehicle_Configuration. Likewise, if the Comm/Non-Commercial field is coded 1 (Interstate carrier) or 2 (Intrastate carrier) then there should be information entered in the Motor Carrier Name, Address, and USDOT Number fields.

- **Invalid Distance from Intersection.** If the "From the Intersection of" block is completed, the feet or miles fields should contain a non-zero number.

- **VIN and vehicle make, year, type do not match.** In post-processing, the vehicle identification number (VIN) decoding can return a value for vehicle make, vehicle year, and (within limits) vehicle type. These values can be cross checked to ensure that the VIN was recorded correctly and that the vehicle descriptors were also recorded correctly.

- **Vehicle Body Type and/or Commercial Motor Vehicle Configuration inconsistent with Trailer Type.** Some vehicle types would be unlikely to have a trailer (train, moped, low speed vehicle). Others should only have certain types of trailers (e.g., an automobile is unlikely to tow a single or tandem semi trailer, a pole trailer, etc.). Likewise, concrete mixers do not usually tow any trailing units.

- **Driver's Actions (multiple codes per driver) cannot be coded 01 and any other value.** This check looks across the three possible values of Drivers Actions to make sure that if more than one value is coded, none of the values is set to 01.

- **Contributing Circumstances: Road (multiple codes) inconsistent.** If code 1 (None) is used, no other codes should be reported.

- **Vehicle Maneuver Action inconsistent with Harmful Event/Sequence of Events.** The Vehicle Maneuver Action code 08 (Parked) is inconsistent with any of the event codes implying movement of the vehicle at the time of the crash.

- **First Harmful Event (crash level) inconsistent with Harmful Event and Sequence Events (vehicle level).** Whatever is coded for the First Harmful Event in the crash level record should appear as the first harmful event in the sequence of events for at least one of the involved vehicles involved in the crash.

- **Vehicle Defect (multiple codes per vehicle) cannot be coded 01 and any other value.** This check looks across the two possible values of vehicle defect to make sure that if more than one value is coded, neither value is set to 01.

- **Missing Driver Information.** If Vehicle in Transport and Not Hit & Run are both coded, driver information fields of the form should be completed.

- **First Harmful Events inconsistent with Vehicle Type.** If the Harmful Event code 4 (Jackknife) is used, then there must be at least one trailer coded for the vehicle.

- **Lighting Condition inconsistent with Time of Day.** This edit check is discussed in detail in the report. It is designed to check the recorded lighting condition against the time of the crash to ensure that when it is nighttime, crashes are not coded as 1 (Daylight), and when it is daytime, crashes are not coded as 4, 5, or 6 (the codes for Darkness with/without lighting, and unknown lighting).

- **Type of Intersection is inconsistent with location information boxes used.** If the code 01 (Not At Intersection) is used, the location box for "at the intersection of" should not be completed. If any other code is used, the location box for "at the intersection of" should be used.

- **Estimated Vehicle Damage inconsistent Area of Impact indications.** The estimated vehicle damage should not be $0.00 if there are indicators of damage.

- **Restraint Systems and/or Air Bag Deployed inconsistent with Vehicle Body Type.** It is unlikely (though not impossible) that code 7 (Child restraint) would be valid when paired with a vehicle type of 11 (motorcycle). Likewise, bicycles, motorcycles, and mopeds do not generally have air bags so any Air Bag Deployed code other than 1 (Not Applicable) should be flagged for review.

To further assist the working group, we have provided under separate cover the business rules (edit checks) defined for Michigan's crash data system.

- Develop measures of crash data accuracy based on the following:
  - Automated review of the accuracy of data elements using the improved edit tables. This check should include validation that the reported elements meet the data type and allowed range, as well as verifying the logical consistency among data fields.
  - Develop transaction tables to store results from running the edits against the as submitted data and as corrections are made through quality control processes.
  - Develop metrics to assess crash data accuracy by comparison of original (as submitted) data to the corrected data using the recommended transaction tables.

- Report these measures periodically to reflect overall accuracy and the accuracy of reporting by each LEA. Share the results with stakeholders, including the Connecticut TRCC. The accuracy metrics should be reported to LEAs in a similar ranking format as that shown in Table 1 for timeliness measurements.

- Update the crash data collectors' manual and training for law enforcement. This will be especially important as the new crash report is implemented; however, a new version of the manual to clarify fields on the current form is clearly needed.

6 Completeness

There are no measures of the completeness of crash data. CTDOT improved the completeness of their crash database by collecting property damage only (PDO) crash data on local roadways starting with 2007 information.

As noted earlier, CTDOT does not collect all information included on the PR-1 crash report form. Complete information is gathered from electronically submitted reports from the CSP.

6.1 Completeness Recommendations

- Capture all data on the crash report form in a centralized crash database. If the UConn system serves as that resource, ensure that the dataset can securely include personal identifiers.

- Develop measures of crash data completeness based on the following:
  - Automated review of missing and "unknown" values of data elements assessed by the improved edit table. These should be tracked separately from the accuracy metrics recommended above.
  - Comparison of current to prior years reporting levels statewide and for each LEA.
  - Comparison of current to prior years on proportion of (fatal + injury)/total crashes statewide and for each LEA.

- Report these measures periodically to reflect overall completeness and the completeness of reporting by each LEA. Share the results with stakeholders, including the Connecticut TRCC. An example of a report that shows the ratio measure for the top 25 towns (based on total crash reports submitted) is shown in Table 2. This example is not exactly as we describe in the recommendation because the CTDOT data do not include a LEA identifier. In the example, we ran the analysis based on town name rather than reporting agency. The form of the report recommended above would require that the State first establish unique identifiers for each LEA and then run the report shown in Table 2 based on LEA identifiers, not town names.

  This report can be used to identify LEAs that may have changed their reporting practices (typically under-reporting PDO crashes) as the ratio measure responds to changes in the proportion of serious crashes in the context of all crashes reported. Agencies with a relatively high ratio should be examined for potential under-reporting.

Table 2. Top 25 Towns for Crash Report Submissions
Showing the Ratio of (Injury + Fatal crashes)/Total Crashes

Township Code	Township Name	2006 Ratio Fatal+Injury/ Total	2007 Ratio Fatal+Injury/ Total	2008 Ratio Fatal+Injury/ Total	Total Crashes	Total Ratio Fatal+Injury/ Total
93	NEW HAVEN	0.47416	0.27098	0.25450	14,681	0.31326
15	BRIDGEPORT	0.51897	0.27171	0.28845	14,658	0.33095
64	HARTFORD	0.55814	0.27704	0.28168	14,206	0.32916
151	WATERBURY	0.49775	0.26624	0.25476	13,137	0.30897
135	STAMFORD	0.43824	0.21964	0.21679	10,552	0.26384
103	NORWALK	0.36781	0.22574	0.21116	9,458	0.25365
34	DANBURY	0.34713	0.19251	0.21936	8,328	0.24087
57	GREENWICH	0.37649	0.20415	0.19421	5,909	0.23727
156	WEST HAVEN	0.41619	0.26456	0.26782	5,414	0.30033
77	MANCHESTER	0.48061	0.28292	0.23988	5,231	0.30816
80	MERIDEN	0.41637	0.23008	0.25390	4,891	0.28195
17	BRISTOL	0.43288	0.23956	0.26399	4,881	0.28949
62	HAMDEN	0.34621	0.25271	0.24005	4,872	0.27360
138	STRATFORD	0.33567	0.23711	0.23343	4,773	0.26231
89	NEW BRITAIN	0.55348	0.27320	0.30452	4,650	0.34301
43	EAST HARTFORD	0.37728	0.28675	0.26914	4,452	0.30368
155	WEST HARTFORD	0.48908	0.29817	0.30256	4,354	0.34015
148	WALLINGFORD	0.34907	0.23810	0.22643	4,299	0.26285
104	NORWICH	0.30345	0.22092	0.20142	4,236	0.23466
84	MILFORD	0.45864	0.35393	0.37802	4,137	0.38990
158	WESTPORT	0.30241	0.24983	0.23823	3,991	0.26059
51	FAIRFIELD	0.36685	0.29816	0.32667	3,844	0.32700
101	NORTH HAVEN	0.33710	0.26083	0.24322	3,464	0.27569
131	SOUTHINGTON	0.37468	0.27009	0.27741	3,182	0.29887
144	TRUMBULL	0.33446	0.26807	0.25554	3,118	0.28287

7 Consistency/Uniformity

The only measure of consistency relates to MMUCC compliance of the current crash report form. A workgroup of the Connecticut TRCC is working on an updated electronic PR-1 which will improve MMUCC compliance. There are no measures of consistency of reporting from year-to-year statewide or for individual LEAs.

It was reported that the LEAs and CTDOT are not consistent in the way they prefer to code specific data elements or crash events. There were three examples provided: the CTDOT does not allow LEAs to code crashes as "same direction/sideswipe", the vehicle type codes are inconsistent among DMV and CTDOT, and a town changed its name and the town code numbers now differ between CTDOT and the LEAs.

7.1 Consistency/Uniformity Recommendations

- Use the recommended measures of completeness as measures of internal consistency among LEAs. In essence this recommendation recognizes that there is overlap between completeness and consistency as they relate to LEAs reporting of crashes. By comparing multiple years of data for each LEA and comparing among LEAs on the metric of (fatal + injury)/total crashes, the State also arrives at a measure of consistency.

- As planned, design the new electronic PR-1 crash report form to increase MMUCC compliance. Request a review of the form's compliance through NHTSA. The State may wish to consider additional data elements beyond those included in the MMUCC guideline (which is designed as a "minimum" dataset). It is important to ensure that not only are specific data elements and attributes collected (or available through linkage) but that the collection is at the correct "level" in the crash report. For example, harmful event codes should be collected at the vehicle level (so that each "unit" involved in the crash has codes for the harmful events) rather than only at the crash level (one set of codes that applies to all vehicles).

- Task the TRCC subcommittee to address the coding discrepancies among CTDOT, the DMV, and the LEAs. Establish a single coding standard and a data definition for each field on the crash report. These should be reflected in an updated manual for use by LEAs in training officers and in the edit checks defined for each of the data fields.

8 Integration

Examples of data integration include: use of the driver's license and vehicle registration to auto-populate fields on the CSP electronic crash collection system; linkage of crash and roadway inventory data; and a Crash Outcome Data Evaluation System (CODES) project listed in the Traffic Records Strategic Plan. No measures of integration were supplied.

The team was not presented with results of analyses arising from the CODES project. It was stated that the CODES project has successfully linked crash and injury surveillance data through 2007. It was not clear if the linked dataset has been used to produce analyses of crash outcomes to date.

8.1 Integration Recommendations

- Develop macro-level measures of data integration listing the data sets that can be linked and the numbers of years of data for each possible linkage. This type of measure of integration is simple to create and is relatively stable from year-to-year. Reporting these metrics will give users and oversight (including the TRCC) a rough idea of the availability of linked datasets capable of supporting safety analyses.

- Develop micro-level measures of data integration showing the strength or success level achieved in each linkage between crash and other data sources. These measures are crucial to understanding the State's ability to merge individual records among multiple datasets. The strength of linkage tells users how reliable are the processes involved in merging records and gives users another way to assess the overall quality of the data in each of the individual datasets. In essence, the ability to merge datasets is a demonstration of the underlying quality of the individual data sources. Where there are errors, inconsistencies, or gaps, the records will fail to find "partners" in the other databases. As an example, a performance measure could be created that would show the percentage of persons coded as injured in crash reports for whom no medical record (EMS, trauma registry, Emergency Department, or hospital discharge data) could be located. A companion metric could be produced that would show the proportion of individuals coded in the medical records where the cause of injury is "motor vehicle collision" for whom no corresponding person record can be located in the crash database. These two metrics define the "expected" matches based on the data from one dataset (crash or injury surveillance) and calculate the percentage matched in the other dataset.

- Develop measures of location coding success based on the ability to link crash and roadway inventory data. Include measures of the success of automated processes and the ultimate strength of the linkage based on the final result after automated and manual processes have been applied.

- Promote the use of linked data produced by the CODES project to enhance the understanding of crash consequences in the State.

9 Accessibility

There are few examples of crash data accessibility to stakeholders outside of CTDOT and law enforcement accessing their own data.

The number of requests for accident reports is maintained by the CSP Reports and Records Division. The *Accident Facts* report is a biennial product of CTDOT. CTDOT makes other resources available as well. Unfortunately, the Department's web site is difficult to navigate and the reports are not presented online in an obvious and easy-to-locate manner.

UConn is developing a web-based analysis system that is user ID and password protected. The web site allows any user (including the public) to specify a query, produce a list of crash reports meeting a selection criterion, download a data extract, and produce a summary table (cross-tabulation). The system is still under development.

9.1 Accessibility Recommendations

- Establish metrics of data accessibility to measure the number of requests for analysis (not just individual crash reports), downloads of standard reports (e.g., the *Accident Facts* report), and use of the online query tool available through UConn.

- Improve the web "presence" of safety-related reports and information. The CTDOT *Accident Facts* reports should be featured on the CTDOT web site and on the web sites for other safety partners. Another improvement could be to create a web-based "portal" for traffic safety information in general. One good example of this concept is the Safety Data Portal under development by the University of Wisconsin under contract to the Wisconsin Department of Transportation. It is possible that the UConn data repository could add this feature relatively easily.

- Measure user satisfaction with access to data through a periodic survey. The TRCC members would be an ideal group to survey, along with other safety stakeholders.

10 Summary

Connecticut is at a crossroads with respect to crash data processing. It is tempting to say that a crisis point has been reached; however, this is not strictly true at present. The system that exists today in CTDOT does not meet their own needs and clearly does not serve the broader safety community. It is possible that the system could be maintained in its present form for a few more years – thus it would be premature to raise the alarm and say that a crisis is upon the State. If no changes are made, however, it is equally clear that the current system is not sustainable in the long term. Even with Federal agency assistance, the crash data entry backlog is at 14 months. This compares poorly with the ability of many States to have data posted to a production crash database within a month of the crash event. Some States are going further and achieving "day-current" posting; making data available for analysis within 24 hours of the crash event.

The situation in Connecticut is much less bleak than the current backlog would imply. In fact, much of the data that are being entered manually today already exist in electronic files maintained by LEAs. Also, CTDOT has a modern database that is easily modified and expanded. It can hold the complete crash report record including narrative and diagram – it already does so for the crashes reported electronically by CSP. From an infrastructure point of view, Connecticut is poised to make a huge leap forward in electronic crash reporting and submission and, as a consequence, reap the benefits of reductions in manual data entry. The State has many good plans and projects already. Implementation of many of these is recommended without reservation in this report.

The most obvious need is for written plans – here we call for a Business Plan – to serve as tools to guide the completion of the various projects aimed at improving the crash data processes and system. One over-arching Business Plan that incorporates all of the various projects into one coherent and well sequenced set of timelines, milestones, and deliverables is the best way for CTDOT to ensure success. The Business Plan will help the State attract funding (from USDOT and other sources). It will also enable stakeholders throughout government to monitor and promote the key projects that are incorporated into the plan.

We also strongly recommend that CTDOT establish a position for a full time manager of the crash data improvement process. This could be a grant-funded position, or, if funds are available, a State-funded position. The individual selected should be empowered to implement the Business Plan – working with partners responsible for each of the projects and promoting the plan throughout the State and especially among law enforcement. The goal of 100 percent electronic reporting is central to everything that Connecticut hopes to accomplish with safety data analysis. Without it, there will not be a complete, centralized crash system. Without it, there will not be timely data available. Without timely data, the State will not be able to implement the most effective methods of data quality improvement – those relying on timely

feedback to law enforcement. The need to focus on electronic data capture and submission cannot be over-emphasized. It is the key to the future of crash data management in Connecticut (as it is elsewhere in the United States).

We also wish to stress that this improvement effort needs champions at the upper levels in CTDOT and throughout State and local government. The vision for the future of safety analysis should be shared among all stakeholders and all should be committed to making the various projects succeed. CTDOT must take the lead, however. There is no substitute for CTDOT as the custodian also producing the upper-level champions for improving the timeliness, accuracy, completeness, consistency, integration, and accessibility of crash data. With CTDOT's commitment, we believe that the State can have a state-of-the-art system in less than five years. Without that commitment, we do believe that the current system, unsustainable in the long term, could well collapse in that same five-year time frame. Resources for continued data entry are not going to be available in the future.

It is equally clear that UConn has become an important partner with a key role to play in the future of Connecticut's crash records system. While CTDOT must be the upper-level champion for improved data, UConn is slated to become the means by which many improvements will be achieved. They are poised to become a storage and access point for crash data that will serve the broadest possible user community. In addition, many of the recommendations in this report could be implemented efficiently through UConn by virtue of their expertise in data integration and electronic data management.