



Caltrans Division of Research,
Innovation and System Information

Research

Notes

Transportation
Safety and
Mobility

OCTOBER 2015

Project Title:
TASAS (Traffic Accident Surveillance
and Analysis System) and Injury
Data Base Development.

Task Number: 2906

Start Date: October 1, 2015

Completion Date: September 30, 2017

Task Manager:
Hamid Ikram
Transportation Engineer
hamid_ikram@dot.ca.gov

TASAS (Traffic Accident Surveillance and Analysis System) and Injury Data Base Development

Streamline and digitize the process of Caltrans collision post-mile coding, data extraction from TCRs (Traffic Collision Reports) for populating the TASAS database.

WHAT IS THE NEED?

The TASAS is an electronic database and data processing system that contains data for collisions that are state highway related. Each collision record in the database is referenced to a post mile address that ties to the highway database. The highway database contains data on 15,200 miles of highway, 20,000 intersections, and 16,000 ramps in California.

Caltrans Collision Post-mile Coding unit processes the Police TCR's and assigns specific location values. The collision detail information is then transferred to the Statewide Integrated Traffic Records System (SWITRS) and TASAS databases. SWITRS is a database that serves as a means to collect and process data gathered from a collision scene. There are variations in the police TCRs since these are provided by approximately 100 California Highway Patrol (CHP) area offices plus over 400 different local police departments making the coding and data extraction process based on non-uniformly prepared TCRs.

Since CHP is rolling out an electronic system, all the TCRs from the 100 CHP area offices will be prepared electronically. SWITRS data can then be generated directly from what the investigating police officer preparing the TCR inputs into the electronic TCR system. It will therefore be feasible to develop techniques and algorithms for automatically extracting data from such digitized system into the TASAS database.

In addition, there is a need for methods to automatically or at least semi-automatically extract similar data from non-digitally generated TCRs from other local police agencies who may not have digitized their process. Additionally, such automated or semi-automated techniques are needed for any legacy TCRs due to backlogs or existing reports that have not yet been coded or extracted into SWITRS and TASAS.



DRISI provides solutions and
knowledge that improves
California's transportation system

The research will address these issues and will build upon the demonstrated previous work of the Advanced Highway Maintenance and Construction Technology (AHMCT) Research Center at the University of California-Davis. It will develop methods and algorithms for a system that would push the process through digitization and automation with the potential need of coding units for quality assurance and handling of non-uniform TCRs.

WHAT ARE WE DOING?

We are performing research on methods of streamlining and digitizing the process of Caltrans Collision Post-mile Coding as well as Data Extraction from the Police Traffic Collision Report for populating the TASAS database as well as extending the work zone injury database to include some non-work zone highway accidents.

This work builds upon previous very successful work of AHMCT in developing a California work zone accident injury database that can semi-automatically extract some data from police TCRs. This research is extending such methods for data extraction, processing, and management at a much larger level for the TASAS database.

WHAT IS OUR GOAL?

The main goal of the research is to develop methods that would facilitate data extraction and processing for both the coding portion as well as for populating the TASAS database. This ongoing effort is also intended to be consistent and take advantage of the new efforts by the CHP to have all TCRs in electronic format.

The secondary objective of the study is to use some of the newly processed police TCRs to add to the existing California work zone injury database for potential evaluation of mitigation methods. The results of this research can provide data and answers to the following research questions:

- The extent to which data coding in terms of determination of district, county, route, post mile, travel direction and post mile markers can be digitized and streamlined.
- The extent to which data from electronic TCRs can be digitally extracted and automatically put into the TASAS database.
- The extent to which data from narrative portion the Police TCRs can be automatically extracted and codified and used to digitally populate the TASAS as well as the Work Zone Injury Database.

- How do the CHP efforts to provide the police TCRs in an electronic format impact Caltrans operation in populating the TASAS database as well as providing the coding to the CHP?

WHAT IS THE BENEFIT?

The following benefits are expected from this research effort:

1. An optimized method for optical character recognition to extract data from police TCRs appropriate for TASAS.
2. An improved process for linking collision data to Caltrans Geospatial Linear Reference System for TASAS.
3. A computerized system for automated post mile coding from CHP electronic TCRs.
4. A computerized system for automated or semi-automated extraction of TASAS Data from CHP electronic TCRs.
5. An experimental computerized system for automated or semi-automated data extraction from high quality PDF files of police TCRs.
6. A turnkey integrated approach for TASAS data coding and extraction.

WHAT IS THE PROGRESS TO DATE?

A research project kick-off meeting was held on October 23, 2015. A Technical Advisory Group (TAG) was formed for this research project. We started work with the TAG and all stake holders to understand all requirements for Post Mile Coding and TASAS database automation and "Card 8" process to develop a system requirement document.