

SHOT: Developing a Database for Police Shooting Incidents in the United States

Hasan Arslan, PhD
Criminal Justice and Security Department
Pace University
Pleasantville, NY 10570
harslan@pace.edu

Dan Farkas, PhD
Information Technology Department
Pace University
Pleasantville, NY 10570
dfarkas@pace.edu

Abstract--Gun violence and in particular police shootings have become prominent in the news media in the United States. However, there is currently no national database of police incidents. This paper describes the first phase of a project to develop a database to collect, edit, query and analyze police shootings. The SHOT Project (Statistics Help Officer Tactics), an academic collaboration between Criminal Justice and Information Technology departments, has developed a protocol for collecting data from open sources using content analysis. The database will enable the research community and law enforcement the ability to study the dynamics of shooting incidents, the ability to analyze this information and be better prepared to understand trends and patterns. The second phase of the project will be GIS Visualization and an Internet-based enterprise system for update, query and analysis.

Keywords: officer-involved shooting, police, database, use of deadly force

I. INTRODUCTION

Few studies have analyzed officer-involved shootings from a use of deadly force policy perspective. More specifically, little scholarly research on police shootings focusing on precursor variables is available because most research has been done by the newspapers mainly on city or local level scope. Others examine the consequences of involvement in shootings for officers. They focus on aftermath of a shooting and the officers who pull the trigger during the deadly confrontation [1, 2].

David Klinger's famous research on developing an understanding of officers' involvement in shootings is one of the few to mention here [3]. Klinger mainly addressed the issue of an officer's experience during and after shootings. The researcher interviewed eighty municipal and county police officers, who reported on 113 different cases where they shot citizens during their careers. One of the primary findings of the study was that "nearly half of the shootings occurred whilst the officers involved were working general patrol assignments, and a minority of the shootings occurred during tactical operations".

Similar to Klinger's study, McElvain & Kposowa examined the characteristics of officers involved in shootings and found that male officers are more likely than female officers to fire their weapons, and college-educated officers

are less likely to shoot than officers with no college education [4]. Furthermore, the risk of an officer being involved in a shooting decreases with age.

One of the earliest studies is a book titled "*Police Defensive Handgun Use and Encounter Tactics*" by Brian Felter in 1988 [5]. According to Felter, "the use of deadly force by police often occurs (80% of the time) with seven yards or less separating the officer and suspect; the incident is over within 3 seconds; and the average number of shots fired, by all those involved in the incident, is less than three".

As indicated above, most studies examine the issue in more local contexts. The following studies present findings from such research. In Boston, researchers attempted to uncover distinctive developmental trends in gun assault incidents over a 29-year period and found that "gun violence is intensely concentrated at a small number of street segments and intersections rather than spread evenly across the urban landscape between 1980 and 2008" [6]. Michael White analyzed the role and potential influence of external environmental factors on police shooting behavior in Philadelphia [7]. Jeffrey S. Adler approached the issue a bit differently by discussing the history of police brutality in Chicago from 1875 to 1920 [8].

The San Diego County District Attorney's office published a study that they conducted themselves over a span of ten years [9]. This seems to be one of the most inclusive studies that was conducted locally, which investigated 200 officer-involved shootings from 1996 to 2006. The study included how many fatalities and non-fatalities there were each year, the number of cases by month, and the percentage of cases by day of the week, etc. The San Diego study also revealed that most of their crimes occur at night between 12:00 PM and 1:00 AM.

Jimmy Fisher, a former FBI agent and true crime author of *The Fisher Report*, collected a dataset from open sources on police shootings in single year alone, and reported that police officers shot 1,146 people, killing 607 between January 1, 2011 and January 1, 2012 [10].

Perkins and Bourgeois investigated perceptions of police misuse of deadly force by looking at two studies [11]:

“Study 1 showed that as number of officers decreased and number of shots increased, perceptions of misuse of force were augmented. Number of shots per officer significantly predicted perceptions of misuse of force. Study 2 investigated the effects of social dominance orientation, blind patriotism, and right-wing authoritarianism. Results showed a significant interaction between number of officers, number of shots fired, and social dominance orientation.” (2006).

One of the best web-based police shootings database is managed by the Las Vegas Review Journal newspaper website in terms of number of variables and search functions [12]. Despite its user-friendly format, it only captures and records the shootings that took place within the Las Vegas metropolitan area. In sum, none of the studies directly address the issue by using a nationwide police shootings data. However, it should be noted here that there has been an influx of information in open sources, mainly newspapers, after the Michael Brown shooting in Ferguson, Missouri since August 2014. Therefore, many attempts were made to develop a primitive form of police shootings database nationwide by some non-profit organizations. Websites like fatalencounters.org, and <http://regressing.deadspin.com> are one of the few worth to mention [13, 14]. The SHOT project, however, brings a different approach by gathering data from open sources and analyzing and recording the data in a specific set of database variables. Not only, the database expands the number of variables that examine the police decision-making behavior during a deadly encounter, but also collects shooting incidents nationwide.

There is precedent to performing content analysis on open sources and storing incident data in a relational database. Hale did a similar project for worldwide terror incidents [15]. While the attributes were different, they were collected using content analysis of open sources and proposed using a relational database to study trends and patterns. Hale points out that “most of the research lacks a comprehensive database approach,” and that this approach offers to link a large number of variables and provide the organization and structure required to investigate valuable patterns and trends”.

II. METHODS: DATABASE DESIGN AND IMPLEMENTATION

The relational database is a type of database management system that stores data items in the form of organized, formally described tables enabling one to link related incidents with groups, individuals, places, properties, etc. It can also store large collections of data in a single location, support multiple users, indicate valuable relationships, and can be organized and updated with ease. Figure 1 shows the logical organization of the SHOT database. Incidents and associated attributes are populated from an initial source document. The major tables associated with an *incident* are the *source* (there can be more than one source contributing to an incident), the *officer* or multiple officers, the *suspect* or multiple suspects.

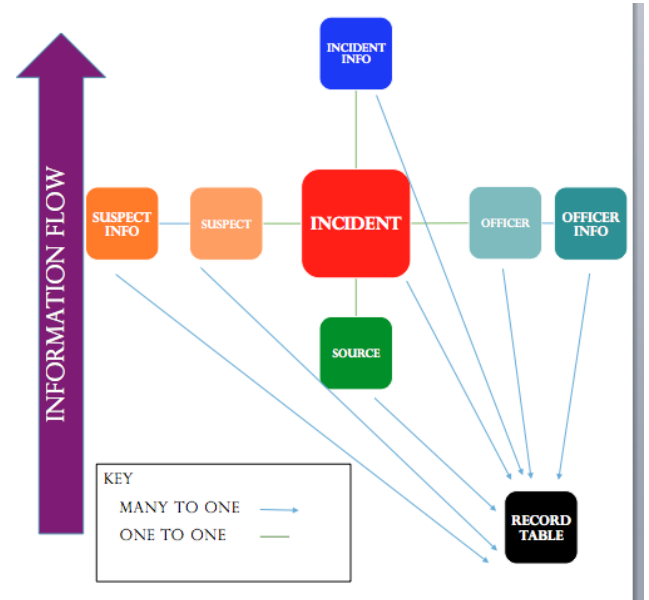


Fig. 1. Relational Diagram of SHOT database.

The development of SHOT used *Agile* (in the software development sense) methods. Agile development is iterative and incremental, and our understanding of both the data requirements and technical requirements has solidified during development. Weekly staff meetings (faculty and student developers) contributed to both the development and design.

III. DATA COLLECTION METHODS

The SHOT project uses an *exploratory research* method to analyze police shootings from a gun violence perspective. Content analysis is the primary data collection method to gather “officer-involved shooting” incidents from open sources. For example, using the NewsBank search engine as primary data collection instrument (Fig. 2), data entry personnel follow a three step process: (1) selecting a *standardized* search term; (2) setting up the timeline for the scope of the search (year by year); (3) determining the geo-location of the search (state by state search). For example, a simple query with a search term “*officer-involved shooting*” in California between January 1, 2000 and December 31, 2013 at infoweb.NewsBank.com search engine results 7,661 articles. Same query results 1,211 articles for Texas and Nevada with 2,581 hits. If the search term is replaced with “*police shooting*,” the results also vary significantly: California with 5,979, Texas with 1,713 and Nevada 1,019 news articles. SHOT collects data from twelve different types of sources (e.g., newspapers articles, websites, books, etc.)

Currently, more than 1,600 incidents from 2000 and 2014 have been collected. In addition to attributes for each incident (e.g., location, date, etc.) data is collected on the source (e.g. newspaper), suspect (e.g., gender), and police officer(s). There are more than 40 attributes for incident based categories (Table 1). All variables are retrieved from non-governmental sources, a.k.a. ‘*open sources*,’ which describes information

that is not deemed classified, secret, top secret, etc. and is readily accessible without security clearances.

TABLE 1. DATABASE COMPONENT WITH ASSOCIATED VARIABLES.

Category	Attributes
Source	Title, date, author, publication, URL, full-text.
Incident	Date (day, month, season, year), approximate time (early morning, morning, noon, afternoon, evening, night and midnight), number of officers on the scene, number of officers, who fired their guns, number of rounds hit the target, part of target being hit (head/neck, torso, limbs) and geographical information (region, state, city, address, location type).
Suspect	Age, race, gender, fatality, mental status, weapon possession, type of aggression, use of vehicle, occurrence of any foot or car chase, gang affiliation, nationality, fatality/injury and whether a lawsuit was filed.
Officer	Race, gender, experience, affiliation, department type, assignment, status (on/off duty) and type of police call.

All variables are retrieved from non-governmental sources, a.k.a. ‘open sources,’ which describes information that is not deemed classified, secret, top secret, etc. and is readily accessible without security clearances.

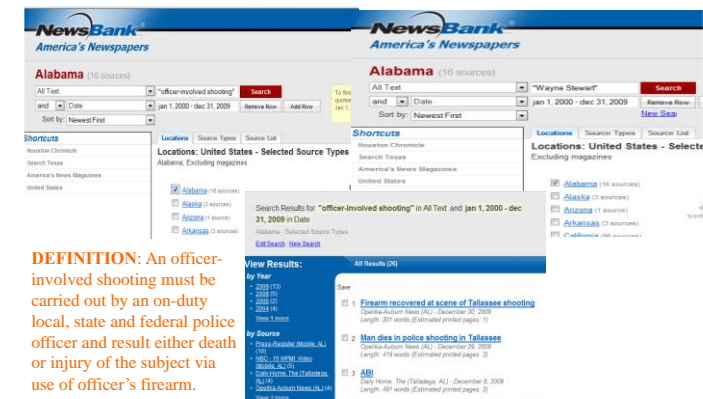


Fig. 2. Open Source Data Collection Methodology from Newspapers

The process involves simply establishing web-based open news sources and then using content analysis, collecting and recording police shooting incidents based on the variables to

populate. There are four main elements that will build the core of an open source data collection for this project (Fig. 3):

- 1) *Incident variables* answer the basic questions of what, when, where, and how. They provide factual information about shooting patterns of police officers as well.
- 2) *Subject variables* refer to information related to shooter or armed assailant. It provides vital information regarding to moments prior to use of deadly force policy by a police officer.
- 3) *Officer variables* mostly include demographic information about police officers, who are involved in a shooting. To be qualified for data entry, a shooting incident must end either with a fatality or injury of the shooter.
- 4) *Source variables* play a key role in SHOT database collection strategy. Every piece of information entered into the database must have a valid source. Data entry person will try to confirm the facts of the shooting with either the information on Police Department websites or at least with another source (*double-check rule*).

IV. DATABASE STRUCTURE

The database schema consists of seven tables that store different aspects of information about an incident and its corresponding source material. Incidents can have multiple sources from which data is extracted, there can be multiple officers per incident and there can be multiple suspects for an incident. The relationship schema is shown in Fig. 3. The tables include:

- **Incident:** This table is the central table to this Database, when a new Incident is created this table generates an Incident ID auto number that connects the other tables together. As the main connecting table for this database it does not store any information other than a user defined name of the incident and the corresponding Incident ID.
- **Source:** This table is directly connected to the Incident table via a many to one relationship. Its primary function is to store all of the information regarding the source materials.
- **Incident Info:** This table is directly connected to the Incident table via a many to one relationship. Its primary function is to store information regarding details of an Incident.
- **Incident Officer:** This table is directly connected to the Incident table via a many to one relationship. Its primary function is to store information regarding the involvement of the officer(s) involved in an incident but not personal information about the officer(s).
- **Subject Incident:** This table is directly connected to the Incident table via a many to one relationship. Its primary function is to store information regarding the

involvement of the Subject(s) involved in an incident but not personal information about the Subject(s).

- **Officer:** This table is not directly connected to the incident table but is connected to the Incident Officer table in a many to many relationship via a pointer table. This was done to prevent duplicate entries in the event an officer(s) are involved in more than one Incident. Its primary function is to store personal information about Officers.
- **Subject:** This table is not directly connected to the incident table but is connected to the Subject Incident table in a many to many relationship via a pointer table. This was done to prevent duplicate entries in the event a Subject(s) are involved in more than one Incident. Its primary function is to store personal information about Suspects.
- **Incident Shot:** This table is not directly connected to the incident table but is connected to the Suspect Incident table in a one to many relationships. This table details the location of each bullet that hit the suspect. This information is stored separately so in the event that a suspect is shot multiple times each bullet can be accounted for.

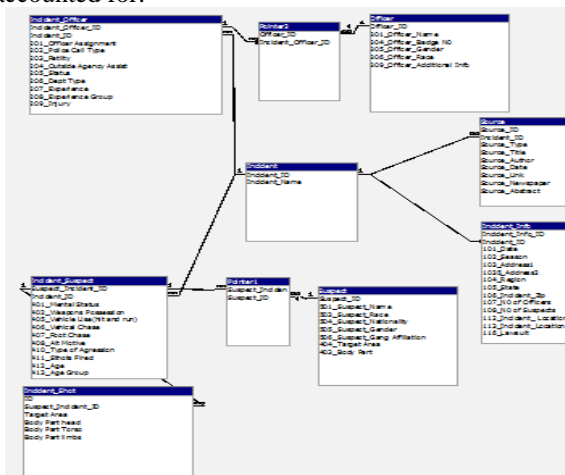


Fig. 3. SHOT Database Schema Relationships

V. DATA FLOW

Due to the flexible nature of the database the only requirement is that an incident ID is created to begin data entry. The user can create and modify any record or field at any time once an incident ID is selected. However it is assumed that the user will start with a new source and fill in the records that are detailed in that source. Thus the ideal data entry flow is as follows (Fig. 4):

1. Create a new Incident by entering an incident name thus creating a new incident ID
2. Enter source Data
3. Enter Incident info, suspect info and officer info as found in the source article
4. Repeat steps 2 and 3 until all sources are exhausted



Fig. 4. Data Entry Flow

VI. USER INTERFACE

In the pilot database, the interface was developed in Microsoft Access using “forms” and “sub-forms”. When a starts SHOT, they enter a new Incident Name, and by navigating tabs entering one or more source, suspect and officer information. Tabs open different Sub-forms that are attached to the current Incident and can be edited in any order at any time.

Fig. 5. Data Entry Form, 1st Page for a New Incident.

The only difference between the Suspect sub-form and the rest of the sub-forms in this database is the inclusion of the Shot sub-form. Like all the other sub-forms the Shot sub-form can be user modified at any time. However it is tied to the Suspect sub-form, meaning that the records that are displayed correspond to the suspect that is being modified.

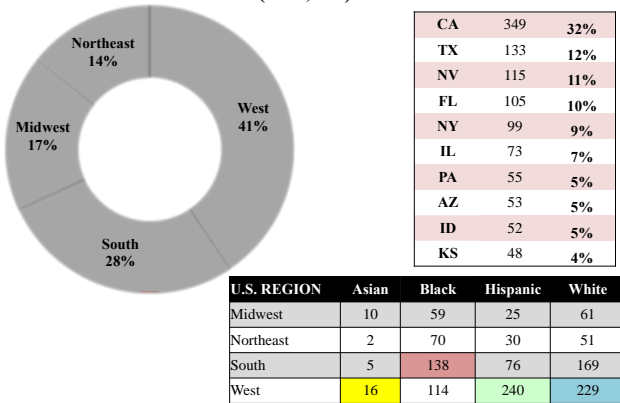
VII. SOME DESCRIPTIVE RESULTS

Based on a snapshot of the data (n~1800 incidents), there are some interesting descriptive statistics:

- 95% of the subjects are male
- 67% subjects were killed
- ¼ of the fatal shootings took place during *Domestic Calls, Warrant Services* and *Traffic Stops*.
- Out of 1806 subjects involved only 24 were not U.S. citizen
- 38% subjects in the 20-29 age group; 24% in the 30-39 age group
- 29% Black; 29% Hispanic; 39% White
- 80% possessed firearms and some types of weapon
- 69% mentally stable; 10% mentally ill; 21% suicidal
- 46% residential; 19% street

Fig. 6 shows the national distribution.

U.S. Officer-Involved Shootings 2000-2014 per Region & State (N=1,806)



NOTE: RI, ND, SC, and DE has the least number of shootings...

Fig. 6. Officer Involved Shootings Correlations by Region & Race.

VIII. NEXT PHASES AND FUTURE RESEARCH

The Pilot SHOT Database was successful in showing feasibility. The next steps are to build an enterprise web-based system with SQL Server as the back end. The data has promise for interesting results on this timely issue. For example one of our goal is to inform police training to reduce gun usage (which is sometimes necessary). By understanding interactions it may be possible to modify protocols for officers in potentially violent situations. Some potential research questions are:

- Is officer experience a predictor of shooting incidents?
- Is the time of day (day/night) a predictor of shooting incidents?
- Is suspect race, age gender predictors of shooting incidents (and in combination)?
- Does location have an impact (indoors or outdoors)?

- Are there spatial aspects to police shootings? Not only where they occur but also questions relating to proximity (e.g., school, restaurant, hospital). For example, Fig. 7 shows the distribution of incidents by weapon type. A specific incident and all its attributes can be displayed.

As the database grows, it will be possible to look at the jurisdiction level, zip code, etc.

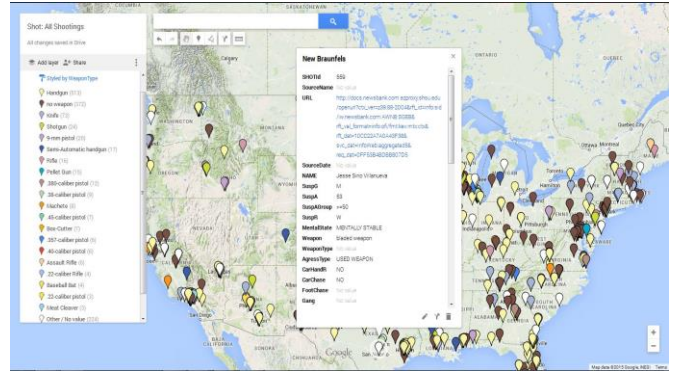


Fig. 7. Distribution of Shootings by Weapon Type

IX. CONCLUSION

The project is at the end of the first phase. The media, government agencies and law enforcement have become very interested in the nature of police shooting incidents, what causes them and how they can be prevented. Unfortunately there is no national database of incidents and the SHOT project will in an innovative way try to address the problem. As Phase I of the project completes, the goal is to establish an organization that, involving students and faculty, that will build the Enterprise Database, enhance data collection, and make its resources available to the larger community.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from the National Science Foundation under Grant No. 1027400. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the US government. The authors would also like to acknowledge the technical support from Luke Babak who contributed to the design and development of the SHOT prototype database. Luke was supported by the CyberCorps program under the same grant during the time of the research.

REFERENCES

- [1] Fooksman, L. (2006, December 4). Repeat Shootings Haunt Officers - Effects Can Be Traumatic, Lasting For Police Who Open Fire. *South Florida Sun-Sentinel* (Fort Lauderdale, FL). 1A. Retrieved on January 8, 2013 from http://articles.sun-sentinel.com/2006-12-04/news/0612040158_1_police-officer-shootings-two-officers

- [2] Xiong, C., (2013, January 8). Shootings often haunt the police officers involved. *Star Tribune (St. Paul, MN)*. Retrieved on May 29, 2014 from <http://www.startribune.com/local/stpaul/186108491.html>
- [3] Klinger, D., (February 01, 2002). *Police Responses to Officer-Involved Shootings*. National Institute of Justice Research Department of Justice. Retrieved on May 18, 2014 from <http://www.ncjrs.gov/pdffiles1/nij/grants/192285.pdf>
- [4] McElvain, J. P., & Kposowa, A. J., (2008). Police officer characteristics and the likelihood of using deadly force. *Criminal Justice and Behavior*, 35(4), 505-521.
- [5] Felter, B. (1988). *Police defensive handgun use and encounter tactics*. Englewood Cliffs, NJ: Prentice-Hall.
- [6] Braga, A., Papachristos, A., Hureau, D., (Mar 2010). The Concentration and Stability of Gun Violence at Micro places in Boston, 1980-2008. *Journal of Quantitative Criminology*, 26 (1), 33-53.
- [7] White, M. D., (2003). Examining the Impact of External Influences on Police Use of Deadly Force over Time. *Evaluation Review*, 27(1), 50.
- [8] Adler, J. S., (2007). Shoot to Kill: The Use of Deadly Force by the Chicago Police, 1875-1920. *Journal of Interdisciplinary History*, 38(2), 233-254.
- [9] Anonymous. (2007, December 4). Analysis of Cases Investigated by the San Diego County District Attorney's Office 1996 – 2006. Retrieved on January 8, 2013 from http://www.sdcca.org/office/ois_review_rpt.pdf
- [10] Fisher, J. (2013, December 25). Police Involved Shooting Statistics: A National One-Year Summary. *Jim Fisher True Crime Blog*. Retrieved on May 25, 2014 from <http://jimfishertruecrime.blogspot.com/2012/01/polic-e-involved-shootings-2011-annual.html>
- [11] Perkins, J. E., & Bourgeois, M. J. (2006). Perceptions of Police Use of Deadly Force. *Journal of Applied Social Psychology*, 36(1), 161-177. doi:10.1111/j.0021-9029.2006.00056.x
- [12] LVRJ, 2015. "Deadly Force: When Las Vegas Police Shout, and Kill," *Las Vegas Review-Journal*. Retrieved on February 1, 2015 from <http://www.reviewjournal.com/news/deadly-force>.
- [13] Burghart D.B., 2015. "Fatal Encounters". Retrieved on February 2, 2015 from <http://www.fatalencounters.org/>.
- [14] Wagner, K., 2014. "Deadspin Police-Shooting Database Update: We're Still Going Strong", *Regressing*. Retrieved on February 1, 2015 from <http://regressing.deadspin.com/deadspin-police-shooting-database-update-were-still-go-1627414202>.
- [15] Hale, W. C., (2006). Information versus intelligence: construction and analysis of an open source relational database of worldwide extremist activity. *International Journal of Emergency Management*, 3(4), 280-297.