

University team works on extract system to help keep tabs on civilians killed by police

September 25 2017, by Nancy Owano



Credit: CC0 Public Domain

(Tech Xplore)—Incidents of people who died during police encounters from news reports can be read, so to speak, by an AI system.

Shubham Sharma in *International Business Times* and several other sites have reported on the system developed by University of Massachusetts Amherst researchers, and the system can curate a database of police encounters by reading from [news](#) reports.

Matt Reynolds in *New Scientist* aptly said it was "a system that automatically scrapes news reports for mentions of police shootings."

The team wrote a paper (*Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*) about their work, "Identifying civilians killed by police with distantly supervised entity-event extraction."

Authors are Katherine Keith, Abram Handler, Michael Pinkham, Cara Magliozzi, Joshua McDuffie and Brendan O'Connor.

How they developed the system: A corpus of [news reports](#) created the database. Their work involved first using keywords like 'officer', 'cop', 'shot', 'died' to gather Google News articles.

"We download a collection of web [news articles](#) by continually querying Google News throughout 2016 with lists of police keywords (i.e police, officer, cop etc.) and fatality-related keywords (i.e. kill, shot, murder etc.)."

The information was processed to avoid duplication and general mistakes, added Sharma.

Results?

The results, said the authors, indicated their model was better than existing computational methods to extract names of people killed by police.

(Sharma remarked that while not extremely accurate, good enough "to prove the viability of the prospect where an AI could offer a faster [approach](#) to create a massive, unambiguous database.")

The authors stated that while they made progress on the application, more work was required "for accuracy to be high enough to be useful for practitioners."

(Obviously, compiling lists out of manually performed news analysis would be laborious.)

"We propose to help automate this process by extracting the names of persons killed by police from event descriptions in news articles."

This is one of three projects under a lab umbrella. The SLANG Lab, College of Information and Computer Sciences, University of Massachusetts Amherst, takes up the challenge of a question, What can statistical text analysis tell us about society? The SLANG Lab (Statistical Social Language Analysis) is directed by Prof. O'Connor.

Their focus is on developing "natural language processing, machine learning, and data analysis tools to improve scientific investigation about political and social phenomena. For example, we [analyze](#) political events in news articles, and sentiment in social media."

The authors wrote on the SLANG Lab site about this particular effort:

"In this project, we propose a new, socially-impactful task for [natural language processing](#): from a [news](#) corpus, extract names of persons who have been killed by [police](#)."

Reynolds reported that O'Connor was interested in pulling in more data from more sources. "O'Connor is hoping to improve the results by

feeding the algorithm a greater range of news sites and perhaps even social media [data](#)."

More information: Identifying civilians killed by police with distantly supervised entity-event extraction, (PDF)
aclweb.org/anthology/D17-1163

© 2017 Tech Xplore

Citation: University team works on extract system to help keep tabs on civilians killed by police (2017, September 25) retrieved 25 April 2024 from
<https://techxplore.com/news/2017-09-university-team-tabs-civilians-police.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.