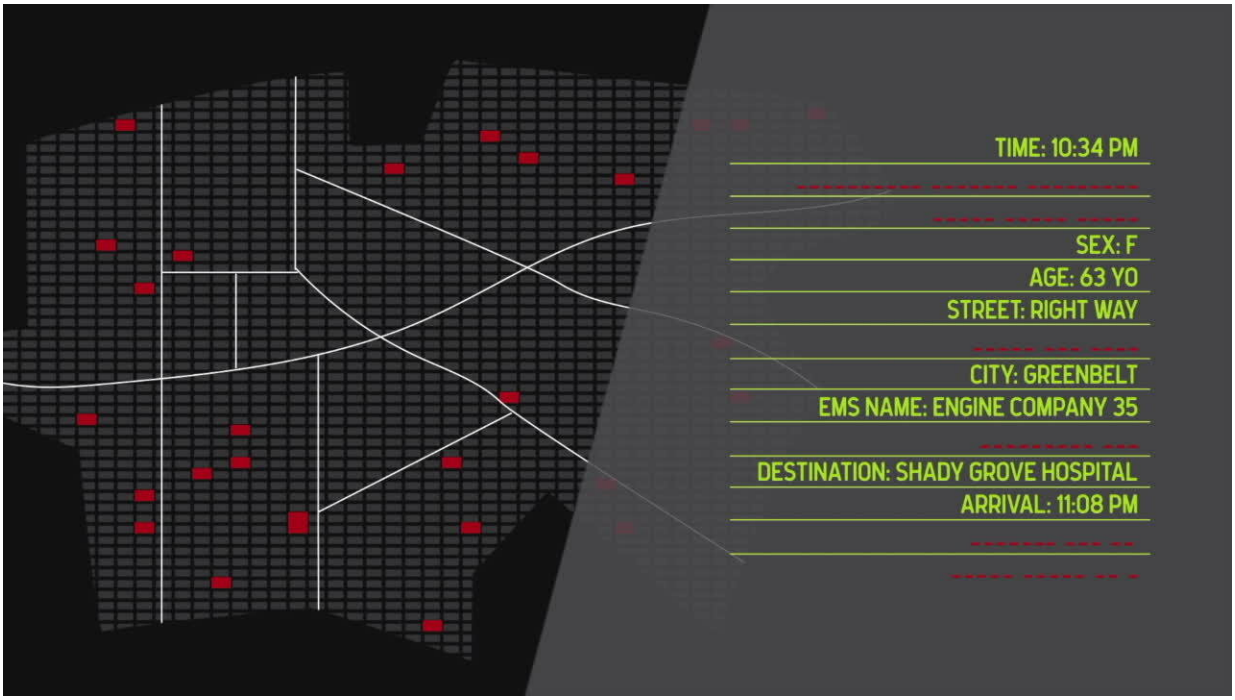
Crowdsourcing challenge to de-identify public safety data sets

October 2 2020, by Jennifer Huergo



First responders and other public safety workers may need access to sensitive data, but sharing that data with external analysts can compromise individual privacy. Credit: NIST

The U.S. Department of Commerce's National Institute of Standards and Technology (NIST) has launched a crowdsourcing challenge to spur new methods to ensure that important public safety data sets can be de-identified to protect individual privacy. The Differential Privacy

Temporal Map Challenge includes a series of contests that will award a total of up to $276,000 for differential privacy solutions for complex data sets that include information on both time and location.

For critical applications such as emergency planning and epidemiology, public safety responders may need access to sensitive data, but sharing that data with external analysts can compromise individual privacy. Even if data is anonymized, malicious parties may be able to link the anonymized records with third-party data and re-identify individuals. And, when data has both geographical and time information, the risk of re-identification increases significantly.

"Temporal map data, with its ability to track a person's location over a period of time, is particularly helpful to public safety agencies when preparing for disaster response, firefighting and law enforcement tactics," said Gary Howarth, NIST prize challenge manager. "The goal of this challenge is to develop solutions that can protect the privacy of individual citizens and first responders when agencies need to share data."

Differential privacy provides much stronger data protection than anonymity; it's a provable mathematical guarantee that protects personally identifiable information (PII). By fully de-identifying data sets containing PII, researchers can ensure data remains useful while limiting what can be learned about any individual in the data regardless of what third-party information is available.

The individual contests that make up the challenge will include a series of three "sprints" in which participants develop privacy algorithms and compete for prizes, as well as a scoring metrics development contest (A Better Meter Stick for Differential Privacy Contest) and a contest designed to improve the usability of the solvers' source code (The Open Source and Development Contest). The challenge is being hosted by

NIST's Public Safety Communications Research (PSCR) division and managed by DrivenData and HeroX.

The Better Meter Stick for Differential Privacy Contest will award a total prize purse of $29,000 for winning submissions that propose novel scoring metrics by which to assess the quality of differentially private algorithms on temporal map data. The three Temporal Map Algorithms sprints will award a total prize purse of $147,000 over a series of three sprints to develop algorithms that preserve data utility of temporal and spatial map data sets while guaranteeing privacy. The Open Source and Development Contest will award a total prize purse of $100,000 to teams leading in the sprints to increase their algorithm's utility and usability for open source audiences.

More information: To learn about eligibility requirements, visit challenge.gov, and for additional information about the challenge, visit DrivenData.org.

This story is republished courtesy of NIST. Read the original story here.

Provided by National Institute of Standards and Technology

Citation: Crowdsourcing challenge to de-identify public safety data sets (2020, October 2) retrieved 19 April 2024 from https://techxplore.com/news/2020-10-crowdsourcing-de-identify-safety.html