Sub-Category: Artificial Intelligence

COMPENDIUM OF RESEARCH ON VIOLENCE AGAINST WOMEN

1993-2020

Table of Contents

Category A: JUSTICE & RELATED SYSTEMS

x. Artificial Intelligence

2017-VA-CX-0033:	**Using Machine Learning to Identify High Risk Domestic Violence Offenders in New York City**
Amount:	**$452,553**
PI:	**Jens Ludwig**
Status:	**Ongoing**

The researchers propose to use machine learning (ML) techniques to develop a tool that predicts risk of domestic violence (DV) victimization and to work with NYPD to test the efficacy of better targeting of high risk IPV victims through a large-scale randomized controlled trial. The ML tool will combine administrative NYPD data with data from sources (e.g., 911 calls). The team will also use techniques to extract signals from text in officer reports, victim statements, and 911 calls. Preliminary results suggest that an early version of the algorithm considerably outperforms both existing risk tools and business-as-usual decision-making by police officers. In partnership with NYPD, the team will formally test this via RCT of the tool to compare the effects of targeting officer home visits to victims—an intervention that has been found promising in quasi-experimental evaluations—using ML relative to status quo. Key outcomes include rates of repeat victimization and serious injury, as measured by NYPD crime complaints and 911 calls (to capture calls for ambulance service, even if police are not called).

2018-VA-CX-0002:	**Using Sentiment Analysis and Topic Modeling in Assessing the Impact of Police Signaling on Investigative and Prosecutorial Outcomes in Sexual Assault Reports**
Amount:	**$714,199**
PI:	**Rachel Lovell**
Status:	**Ongoing**

The purpose of this study is to identify signaling in narratives of police officers' sexual assault reports that influence subsequent decision-making, case flow, and attrition. Signaling is defined as information conveyed by responding officers in the narratives of police reports regarding a victim's credibility and rape-myth adherence. The goal is to better understand whether and how responding officers' written reports in sexual assault cases impact investigating officers' decision-making and how cases proceed, or fail to proceed, in the criminal justice process. The objective of the proposed study is to explore the first step in the investigative process to reveal facilitators and barriers to sexual assault cases reaching a successful disposition. The team will use text mining and machine-learning methods using state-of-the-art software for natural-language processing and advanced statistical analyses to evaluate the narratives of over 6,300 police reports of sexual assaults where victims had sexual assault kits collected in one large, urban jurisdiction over the span of nearly two decades (1993 through 2009). These reports will be analyzed using two computational methods: (1) topic modeling, a statistical approach to identifying and coding latent topics; and (2) sentiment analysis, which involves identifying positive and negative opinions, emotions, and attitudes. The three research questions driving this study include: (a) to what extent do sentiments in the responding officers' narratives reveal positive or negative signaling of victims' credibility? (b) to what extent are the topics and sentiments in the responding officers' narratives different in cases with increased investigative activity compared to those with less and (c) to what extent do both the topics and sentiments in the responding officers' narratives differ for cases that were successfully investigated and prosecuted, compared to those that were not?

2019-NE-BX-0001:	**Assessment of Sexual Assault Kit (SAK) Evidence Selection Leading to Development of SAK Evidence Machine-Learning Model (SAK-ML Model)**
Amount:	**$250,000**
PI:	**Chad Grundy**
Status:	**Open**

The purpose of this study is two-fold. The first purpose is to evaluate crime laboratories' decision-making protocols on the selection of sexual assault kit DNA evidence to develop research-based guidelines on how many and which

swabs should have completed DNA analysis. To achieve the first purpose of the study, data will be collected from three publicly funded forensic laboratories on DNA analysis findings, and from corresponding forensic nursing teams on variables found to be predictive of developing CODIS eligible profiles. The second purpose of the study is to create and implement a software program, Sexual Assault Kit evidence Machine Learning Model (SAK-ML Model), to generate probabilities of developing CODIS-eligible DNA profiles from sexual assault kit swabs. The SAK-ML Model will guide forensic scientists in the selection of the most probative swabs to analyze. Development of the SAK-ML Model is possible, as the Utah Bureau of Forensic Services and the partnering Brigham Young University research team (Julie Valentine, PI) already have compiled a large dataset on DNA findings from sexual assault kits. Utah will enact statewide forensic electronic medical records beginning in July 2019, which will allow for automatic generation of the SAK-ML Model. Implementation of the SAK-ML Model will enhance crime laboratory efficiency and improve DNA analysis outcomes from sexual assault kits.

For an index of all grants, go to ojp.gov/sites/g/files/xyckuh241/files/archives/ncjrs/223572-grants-index.pdf.