Testing the use of advanced data analytics on large-scale police data to identify high risk VAWG perpetrators

Project carrying out analysis using natural language processing on thousands of incident reports to identify high-harm high-frequency perpetrators, which would provide a tool to predict risk.

Key details

Lead institution	University of Liverpool
Principal researcher(s)	Prof Barry Godfrey, Prof Simeon Yates and Dr Tom Nicholls barry.godfrey@liverpool.ac.uk
Police region	North West
Collaboration and partnership	• Cheshire Police • Merseyside Police • Dell Computing
Level of research	Professional/work based
Project start date	July 2026
Date due for completion	July 2027

Research context

High-harm-high-frequency perpetrators are responsible for a significant amount of harm, with the average offender victimising women and girls for an average of 15 to 20 years, with 30 to 40 incidents requiring police service per year. They frequently cause serious physical and long-lasting mental injury to their victims and consume considerable police resources. Prioritising those that offer the most harm is one of the highest police priorities. This study will attempt to assess and

analyse qualitative officer notes within risk assessment processes, which has never been attempted on this scale.

Research methodology

Stage one: Research and text analysis of police data (months 1–8)

- Fully review prior studies for best methods and approaches to similar data sets. Select modelling approaches and train on several years' data, testing those models on a separate set of historical data to assess performance on predicting future offending (pre-project and months 1–3).
- Train machine learning models to predict (historical) risk: Model on the basis both of existing quantitative metrics (including those currently used) and the texts associated with perpetrators (months 3–8).
- Evaluate classical machine learning: Classical machine learning models, such as Support Vector Machines, and transformer-based models, such as BERT and Llama, will be evaluated. These will be locally trained and hosted to ensure reproducibility and the security of data (months 3–8).
- Models will be calibrated and performance tested on unseen data, with false positive/negative rates identified to enable safe operational deployment (months 5–8).
- Provide interpretable models or explanations: Explanations for complex modelling will be given, using interpretable techniques where appropriate (months 6–8).
- Retrain and validate locally and regularly (months 7–9).

Stage two: Testing/validating AI predictions (months 9–12)

- Compare the models to the existing approach to assess which are more useable and more effective. Detect and address credibility signalling and bias-inducing language (months 9–10).
- Implement fairness and bias mitigation safeguards: Audit subgroup performance, detect biased linguistic indicators and use fairness aware AI/machine learning methods (months 9–10).
- Validate the model with policing and domain experts through workshops involving police (months 9–12).
- Disseminate findings widely through academic publications (month 12).