

WMP Briefing Paper

Analysis of Stalking and Harassment Offenders to Predict Offending Patterns

Ethics Committee (06 July 2022)

This project is at the proposal stage and is presented to the committee 'in principle' so that any immediate concerns can be raised.

The finer details of the methodology, exact data to be used and mode of communicating the results will not be determined until after the exploratory data analysis (EDA) phase has been undertaken.

Once the analyses have been completed the project will be presented to the Committee again so that the data used, methodology, findings, intention for deployment and communication plans can be examined in detail.

A Stage 1 DPIA is currently being reviewed by the Force Data Protection Officer (DPO) and a Stage 2 will be completed if advised. The Stage 1 is submitted to the Committee.

Tasking

This project was requested by Inspector Mark Cooper on behalf of the Crime Portfolio in February 2022. The request is to use the learning from academic research to underpin a systematic analysis of offenders identified for Stalking and Harassment offences and to assess the risk they pose.

The project is sponsored by Assistant Chief Constable (ACC) Claire Bell, the Force lead for delivering the Home Office strategy to tackle *Violence against Women and Girls* (VAWG).

Purpose

The purpose of this project is to assist the Senior Leadership Team (SLT) of investigative departments to prioritise their resources so that they focus on the known stalking offenders who are most at risk of escalating their offending and causing the greatest harm.

The request is to build on the findings of academic research and undertake a systematic analysis of stalking offenders using data available in a number of West Midlands Police (WMP) systems. This will entail identifying the nature of their stalking and the type of offender they are. It will also involve an examination of data relating to victims, including events which may not have been recorded as stalking. This wide-ranging analysis of data recorded for a number of purposes (calls for service, incidents, crimes, domestic abuse risk assessments, custody records and intelligence reports) relating to both victims and offenders is necessary in order to understand whether a pattern of behaviour which escalates to become very / extremely harmful can be identified early within police records. This will need to include the nature of the interaction between an offender and victim.

The intention is to develop a predictive model so that when new stalking offenders are identified their potential to escalate to more high risk offending or not can be assessed and the appropriate interventions offered (see later).

This analysis supports the Force’s Precision Policing Doctrine by enabling investigators to prioritise resources on the most harmful stalking offenders and reduce the risk they pose to victims. Whilst this will benefit all types of victims, it will particularly support the work of the VAWG reduction strategy.

Context

Figure 1 below shows the volume of Stalking and Harassment offences recorded by the Force each month since April 2019. In the 12 months from April 2021, the monthly average was 4,338 offences.

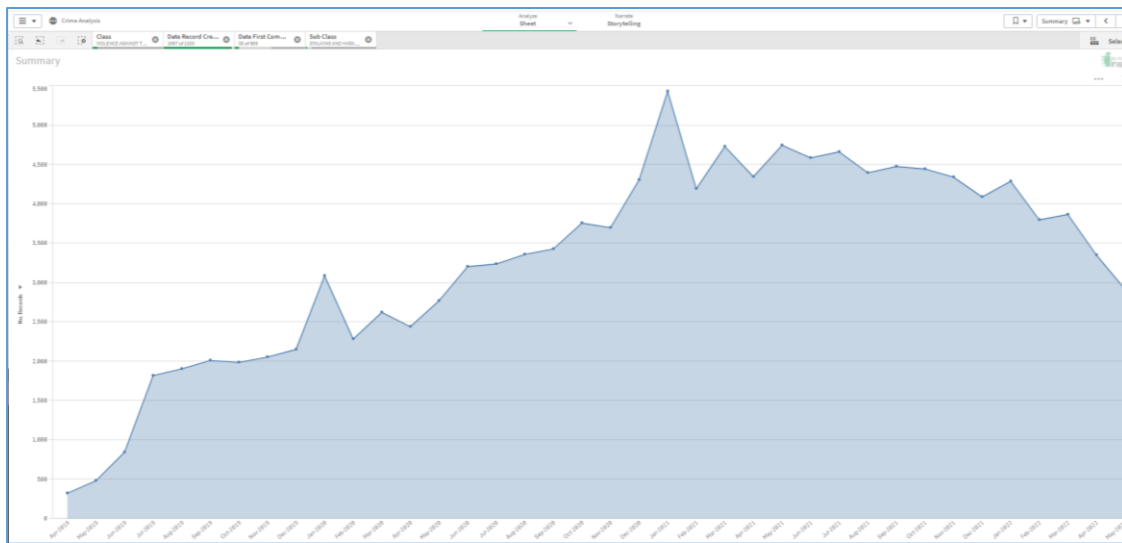


Figure 1: Monthly volume of Stalking and Harassment and Offences since April 2019

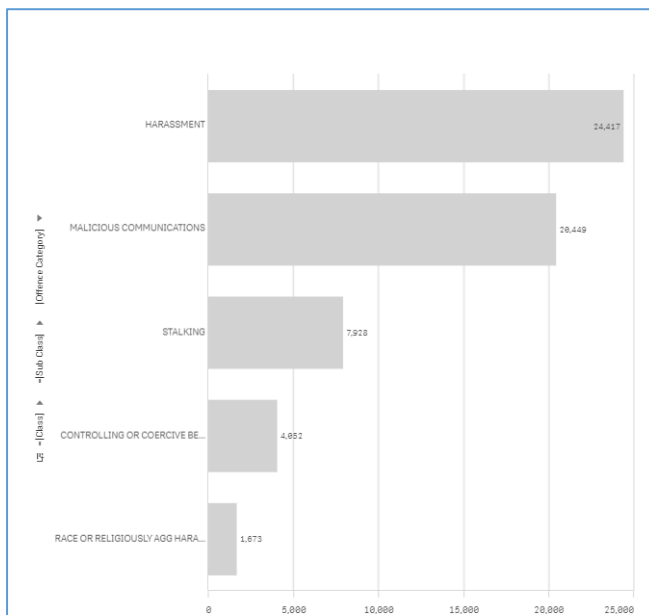


Figure 2: Breakdown of Stalking and Harassment offences recorded by WMP in the 12 months from April 2021

Stalking and Harassment sits within the Home Office offence classification of Violence against the Person. This sub-class includes the offences of Harassment, Malicious Communications, Stalking, Controlling or Coercive Behaviour and Racially or Religiously aggravated Harassment (Figure 2).

Whilst the offences of stalking and harassment are similar in nature and the volume of stalking offences is about a third of harassment, stalking is the more

concerning offence with the potential to escalate to more harmful offending.

A person's course of conduct amounts to stalking¹ of another person if:

- it amounts to harassment of that person,
- the acts or omissions involved are ones associated with stalking, and
- the person whose course of conduct it is, knows or ought to know that the course of conduct amounts to harassment of the other person.

The following behaviours are associated with stalking; following a person, contacting, or attempting to contact, a person by any means, publishing any statement or other material relating to a person, monitoring the use of a person on the internet, email or any other form of electronic communication, loitering in any place, interfering with any property in the possession of a person or watching or spying on a person. Stalking, in particular, is characterised by a person's fixation and obsession with their victim, as opposed to harassment where these factors do not feature.

In the 12 months from April 2021, WMP recorded 7,928 stalking offences, an average of 661 offences a month. The volume of stalking offences has increased and is projected² to remain at elevated levels in the next year. In part, this increase reflects changes in recording practices implemented by the Home Office, in order to better understand these cases, in particular where harassment escalates to stalking:

- April 2018 - stalking / harassment should be recorded in addition to the most serious additional victim-based offence involving the same victim-offender relationship. The old rules meant that the substantive offence was recorded instead of harassment.
- April 2020 - every domestic abuse (DA) related offence of harassment should instead be recorded as stalking unless the Force has a clear rationale for the lesser offence.

Figure 3 shows the counts over time for Harassment offences and Stalking offences which were related to Domestic Abuse and illustrate that these changes in recording practices have had an impact, with the reduction in Harassment offences mirrored by an increase in Stalking offences.

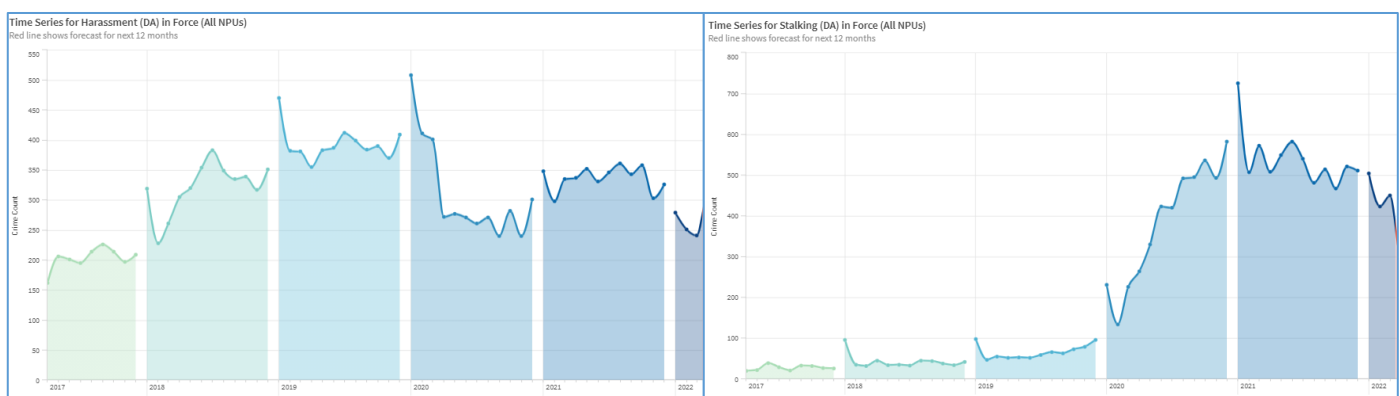


Figure 3: Harassment offences and Stalking offences relating to Domestic Abuse since 2017. Note: The repeated January peak reflects offences which have been recorded where the victim is unable to provide an exact date and so a default date of 01 January is entered.

¹ Home Office Counting Rules (HOCR) for Recorded Crime

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/992839/count-violence-jun-2021.pdf

² DAL Seasonality Planner due to be presented to Committee in July 2022

The Home Office and National Police Chiefs' Council (NPCC) lead for Stalking and Harassment have particularly focused attention on perpetrators as part of the wider strategy to reduce VAWG.³ In January 2020, Stalking Protection Orders (SPO), came into force. These allow magistrates' courts to impose restrictions and positive requirements on perpetrators such as perpetrator programmes to encourage a change in behaviour.

Research suggests that stalking may be a key indicator for future potential serious harm, including homicide⁴. In many cases of domestic homicide, there has been a footprint in police data. In the recent analysis of 124 Domestic Homicide Reviews (DHR) covering a 12 month period from October 2019, approximately 60% of perpetrators were indicated to have a previous offending history.⁵ Similarly, a review of domestic homicides and suspected victim suicides with a known history of domestic abuse covering 12 month period from April 2020⁶ found that many of the key characteristics and risk factors were the same across the cases, with almost half of all suspects previously known to police for perpetrating domestic abuse and some were serial and/or repeat abusers.

The College of Policing (CoP) Authorised Professional Practice (APP) for Investigations and Public Protection⁷ advises that police forces should create systems that aim to:

- identify stalking offences
- support completion of effective risk assessment
- monitor risk management activity
- empower officers to use protective measures to keep victims safe – for example, orders and target hardening
- draw from information about stalking to inform force strategic and operational plans

and ensure that stalking or harassment is recognised within the force strategic threat, risk and harm assessments, to drive resources,

In order to understand which offenders should be subject of SPOs and other intervention tools, the Force must first identify who pose the greatest risk of causing harm. Understanding and tackling the offence of stalking is complex and places an increasing demand on investigative teams.

³ <https://www.gov.uk/government/news/police-awarded-113m-for-programmes-to-prevent-domestic-abuse-crimes>

⁴ Monckton Smith, J., Szymanska, K., & Haile, S. (2017). Exploring the relationship between stalking and homicide. Suzy Lamplugh Trust. Retrieved from <http://eprints.glos.ac.uk/4553/> and Monckton Smith, J. (2019). Intimate Partner Femicide: Using Foucauldian Analysis to Track an Eight Stage Progression to Homicide. Retrieved from <https://asset-pdf.scinapse.io/prod/2966801301/2966801301.pdf> and <https://www.glos.ac.uk/content/the-homicide-timeline/>

⁵ Home Office Report, March 2022. Key findings from analysis of domestic homicide reviews <https://www.gov.uk/government/publications/key-findings-from-analysis-of-domestic-homicide-reviews/key-findings-from-analysis-of-domestic-homicide-reviews>

⁶ Bates, L., Hoeger, K., Stoneman, M., & Whitaker, A. (2021). Domestic Homicides and Suspected Victim Suicides During the Covid-19 Pandemic 2020-2021. Vulnerability Knowledge and Practice Programme (VKPP). <https://www.gov.uk/government/publications/domestic-homicides-and-suspected-victim-suicides-during-the-pandemic>

⁷ <https://www.college.police.uk/app/major-investigation-and-public-protection/stalking-or-harassment> and <https://library.college.police.uk/docs/appref/Advice-supervisors-managers-senior-leaders-stalking-harassment-offences.pdf>

Proposed data sources

Many of the risk factors identified in the research and DHRs can be found in WMP data sets and therefore the DAL has been tasked to develop a model, encompassing all stalking typologies (not just domestic abuse related) to understand if an escalation of offending behaviour can be predicted using the information held within police systems. Since the changes to recording practices for stalking and coercive and controlling behaviour, this type of offending should be more visible within police data.

In order to build a model to predict escalating offender behaviour and the potential for associated victim vulnerability, the analysis would take account of the following elements:

- All offences which have been recorded as committed by a stalking offender (not just stalking or domestic abuse offending).
- Contact data to understand when and how often victims or third parties have called WMP, including phone calls, LiveChat (online web chat service) and online reporting.
- ControlWorks logs – which contain free text information about the nature of any incident where a police response is provided. This may not necessarily result in a crime being recorded if no offence has occurred but could add to an overall picture of a dysfunctional relationship – in particular where a victim might be too intimidated to continue with a complaint.
- Non-crime records where details of domestic abuse incidents are reported but when no offence was recorded.
- Domestic Abuse Risk Assessment (DARA) forms (previously known as DASH) which are completed by officers when they respond to a domestic abuse incident.
- CASE records – investigation notes which record any additional contact with the victim. Analysis of this free text is likely to surface additional knowledge about the nature of the relationship between an offender and victim and may help to populate some of the eight stages of the Homicide Timeline identified by Monckton Smith.
- Police Intelligence Reports (PIRs) about victims and offenders, which hold free text data.
- Custody records which will include flags for existing mental health conditions, alcohol, and drug (mis)use which have been found to be exacerbating factors⁸.

It is important to gather data from a wide range of sources because potential victims may not be aware that a crime they are reporting in fact relates to stalking in the early stages. Equally, as an organisation we record each offence in isolation and on its own merits. This aim of this project is to identify if there are any patterns in the early stages of stalking behaviour. The analysis required goes beyond the complexity of current dashboards, due to the necessity to mine the free text data fields which may contain valuable information about stalking characteristics which are not easily identified in the standard coded fields.

In order to achieve the aim of the request, there is no less intrusive way to understand the history of offending behaviour nor the interaction between an offender and victim and to identify whether a common pattern exists which can be used to build a predictive model. The alternative is to continue with current practices to prioritise the most prolific offenders. This will not necessarily focus resources on those whose offending could escalate to cause the most serious harm.

⁸ Bates. L., Hoeger. K., Stoneman. M., & Whitaker, A. (2021). Domestic Homicides and Suspected Victim Suicides During the Covid-19 Pandemic 2020-2021. Vulnerability Knowledge and Practice Programme (VKPP). <https://www.gov.uk/government/publications/domestic-homicides-and-suspected-victim-suicides-during-the-pandemic>

Intended activity resulting from the project

This project supports the Force priorities of protecting vulnerable people and reducing serious harm, especially to women and girls. From the growing rich source of data, this will potentially support WMP in preventing loss of life of vulnerable victims.

The desired output is a Business Insight (Qlik) dashboard to assist in the prioritisation of the most dangerous offenders.

There are two distinct strands of activity that this work may inform. These strands are related to the offender and the victim. There is clear overlap within these strands.

From the victim perspective there is the essential safeguarding activity. The stalking behaviours are complex and the risks are difficult to extract easily from the broad data. This work will underpin the level of safeguarding that will take place.

This safeguarding can take many forms. This work is intended to support the use of Clare's law where appropriate and provide the legitimate capability to be proactive in our disclosures.

The Force continues to move towards an increasing use of civil interventions. The extracting and enhanced visibility of the risks will further inform where these interventions can be applied.

This increased identification/knowledge and visibility of behaviour will allow for more appropriate advice and guidance to the victims within the safeguarding that is tailored to the offending they are experiencing, including the sign-posting to 3rd party support.

From a suspect's perspective there are many tiers of interventions that this work will more accurately inform.

On an initial level it will enhance the prioritisation and mitigation of risk when offences are reported and suspects become known. This will be from the allocation of resources to progress these investigations/arrests expeditiously to the supervisory and governance oversight.

Disruption and prevention activity can be enhanced via the usage of prosecution, bail, out of court disposals, civil interventions and subsequent offender management. Further activity can be stimulated by appropriately accessing pre-existing activity such as the Early Awareness Stalking Intervention (EASI) project.

Additionally, the full work may be able to identify potential stalking behaviour that can be missed, such as the victim being subject of a burglary or a criminal damage with an unknown offender, but is in fact part of the stalking behaviour.

Academic research is increasingly identifying other criminal behaviours when researching specific questions, such as serial domestic abuse. This project may provide a capability to understand suspects wider offending and therefore create additional opportunities to intervene.

Data

Data to be used:

Only WMP data will be included in this project.

- Recorded Crimes of Stalking and of other historic offending (all offence types)
- Domestic Abuse Risk Assessment (DARA) forms (previously known as DASH)
- Custody Records
- Call data
- ControlWorks
- Case records
- Intelligence
- Non-crimes domestic abuse incident records

Level of analysis:

Individual

Individuals aggregated?

Yes

No

Specific Area:

Output Areas

Super Output Areas - Lower

Super Output Areas - Mid

Wards

Districts

West Midlands

Other

Reliability of data:

An extensive exploratory data analysis (EDA) phase will be undertaken to examine the extent of any data quality issues.

Sample or entirety: Entirety

Type of analysis:

Exploratory

Explanatory

Predictive

Optimisation

Proposed methodology:

This is to be investigated and developed as the project progresses.

Will the project eventually be automated:

Yes

No

Means of evaluation:

Any models built will include train and test datasets so the accuracy of initial models can be assessed. If progressed through to productionisation accuracy will be checked on an on-going basis (potentially in the future including any effects of WMP responses to those estimated to have a high enough probability of going on to commit high harm crimes).

ALGO-CARE considerations

As this project is at the proposal stage and is presented to the committee 'in principle' in order that any immediate concerns can be raised, the finer details of the methodology will not be determined until after the EDA. Once the analyses have been completed the projects will be presented to the Committee again so that findings and methodology can be examined in more detail.

| Advisory | |
|--|---|
| If applicable, are the outputs from the algorithm to be used in an advisory capacity? | The output will alert officers to the potential risk a stalking offender may pose, given their previous offending, relationships to victims, etc. This alert will be advisory only. |
| Does a human officer retain decision-making discretion? | The investigating officer will retain decision-making discretion once they have seen the output of the prediction. They will decide how to progress the investigation, manage the offender and support the victim as appropriate. |
| Lawful | |
| What is the policing purpose justifying the use of the algorithm (means and ends)? | The academic research demonstrates that some stalking offenders will go on to commit the most harmful offences, including homicide. Given the increase in the volume of stalking offences, providing an evidence-based tool which helps to identify those offenders who should be managed more closely will support the national agenda of reducing violence against women and girls and reducing violence in general. This is in line with the policing purpose. |
| Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes? | All data used in the model will have been collected for legitimate policing purposes – all data held about victims will have been generated by them contacting the police in the first instance. Offender data will have been generated during the course of previous investigations. |
| In what way will the tool improve the current system and is this demonstrable? | Currently, if an offence of stalking is recorded the Force has no way to determine the potential risk of that offender causing serious harm in the future. This project will build a statistical model which estimates the probability of an escalating offending pattern. This tool will provide an evidence base on which investigators can make their decisions. |
| Are the data processed by the algorithm lawfully obtained, processed and retained, | The data are from WMP systems and collected to enable their normal day-to-day operations. |

| | |
|--|---|
| according to a genuine necessity with a rational connection to a policing aim? | |
| Is the operation of the tool compliant with national guidance? | The analyses proposed would accord with the Government Digital Service Data Ethics Framework 2020. ⁹ |
| Granularity | |
| Does the algorithm make suggestions at a sufficient level of detail given its purpose and the nature of the data processed? | The prediction will be made at the level of the individual to assist with offender management decisions. |
| Are data categorised to avoid broad-brush grouping and results and therefore issues of potential bias? | Some data may be categorised (via feature engineering), but it is envisaged that any such categorisation would relate to groupings of individual level data as opposed to categorisations across nominals. |
| Do the potential benefits outweigh any data quality uncertainties or gaps? | The project will include an extensive EDA element and this should highlight areas of heightened uncertainty in the data or where particular gaps exist. Should any such issues be identified, these would be addressed as a part of the project. Given the benefits of reducing violent crime and protecting the vulnerable it is not expected that any data quality issues would be of such a magnitude as to warrant not undertaking the project. |
| Is the provenance and quality of the data sufficiently sound? | The data will be those that are gathered during the day-to-day work of WMP and will enable analyses of the type envisioned for this project. |
| If applicable, how often are the data to be refreshed? | To be determined once we understand the output |
| If the tool takes a precautionary approach in setting trade-offs, what are the justifications for the approach taken? | This will not be known to be applicable until the project has progressed. |

⁹ <https://www.gov.uk/government/publications/data-ethics-framework>

| Ownership | |
|---|---|
| Who owns the algorithm and the data analysed? | WMP would own the analyses and data. |
| Does WMP need rights to access, use and amend the source code and data? | No |
| Are there any contractual or other restrictions which might limit accountability or evaluation? | No |
| How is the operation of the algorithm kept secure? | The analyses developed would be kept wholly within the secure WMP computing environment. |
| Challenge | |
| What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias? | The presence of any biases would be investigated as part of the EDA and mitigated once they are known. If productionised, the model(s) outputs would be monitored on a regular basis. |
| If the algorithm is to inform criminal justice disposals, how are individuals notified of its use? | Not applicable. The tool will only be used in an advisory capacity in conjunction with professional judgment of investigators and offender managers. |
| Accuracy | |
| Does the specification of the algorithm match the policing aim and decision policy? | Not known until project progressed. |
| Can the accuracy of the algorithm be validated periodically? | Not known until project progressed. |
| Can the percentage of false positives / negatives be justified? | Not known until project progressed. |

| | |
|---|--|
| How was the method chosen as opposed to other available methods? | Not known until project progressed. |
| What are the (potential) consequences of inaccurate forecasts? | Inaccurate forecasts could potentially lead to extremely high harm crimes occurring which may have otherwise been avoided or too many people brought into the net of potential investigation. Because of the high harm crimes that can result from stalking it will be important to take the unbalanced costs of potential decisions into account (during model build and assessment). |
| Does this represent an acceptable risk? | Not known until project progressed. |
| How are the results checked for accuracy and how is historic accuracy fed back into the algorithm for the future? | Not known until project progressed. |
| How would inaccurate or out-of-date data affect the result? | Inaccurate data could lead to levels of inaccuracy that are unacceptable. Extensive assessment of the data and any issues will be undertaken as part of the EDA. |
| Responsible | |
| Would the operation of the algorithm be considered fair? | During the development of any analyses, the presence of any biases in the underlying data that could feed through to predictions would be fully examined and mitigated if the potential was present. |
| Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review? | The details of any analyses arising from this project would be provided. |
| Would it be considered to be used in the public interest and to be ethical? | Enabling the prevention of serious harm occurring by identifying offenders whose offending behaviour is more likely to escalate would be in the public interest and ethical. |
| Explainable | |
| Is information available about the algorithm / decision-making rules and the impact of each feature? | A technical report will provided detailing the analyses developed. |

Appendix 1: Glossary of Terms

| WMP / Law Enforcement Terminology | |
|--|---|
| ACC | Assistant Chief Constable |
| APP | Authorised Professional Practice |
| CoP | College of Policing |
| DAL | Data Analytics Lab |
| DARA | Domestic Abuse Risk Assessment |
| EASI | Early Awareness Stalking Intervention project |
| EBP | Evidence Based Practice |
| HOCR | Home Office Counting Rules |
| NPCC | National Police Chiefs' Council |
| SLT | Senior Leadership Team |
| SPO | Stalking Protection Orders |
| VAWG | Violence against Women and Girls |
| WMP | West Midlands Police |

| Data Science Terminology | |
|---------------------------------|--|
| ALGO-CARE | All projects have used the ALGO-CARE to consider ethical implications: Advisory, Lawful, Granularity, Ownership, Challenge, Accuracy, Responsible, Explainable |
| EDA | Exploratory Data Analysis |
| Productionise | To 'productionise' means that once we are satisfied that the model works well, we would automate the process of providing predictions every 4 weeks. |