Ethics Committee Briefing Note

Project Reference: DAL_2020_01_CSE Networks

Purpose of data analysis

The overall aim of this project is to identify organised criminal networks involved in the sexual exploitation of children across WMP.

The project seeks to:

- Identify key networks and nominals that are linked to child sexual exploitation (CSE).

- Identify and prioritise the greatest threat and risk posed by these individuals and networks from an exploitation perspective to direct further enhanced intelligence work.

This is linked to the project identifying networks of serious organised crime groups (OCGs) in the West Midlands (DAL_2019_0002_SOC_Networks).

Context

West Midlands Police (WMP) is the second largest police force in the country, covering an area of 348 square miles and serving a population of almost 2.8 million - it is the third most densely populated region in England. The region sits at the very heart of the country, with an average of 170,000 motorists travelling through the region on a daily basis making our motorways some of the busiest in Europe and making the West Midlands an attractive centre where organised criminal networks operate.

Policing the West Midlands area is complex and the communities of the West Midlands are particularly vulnerable to crime and exploitation by organised criminals. Over a quarter of our communities live in the top 10% of the most deprived areas in the country and over a third of the population (35%) is under 25 years old, with lower than average levels of qualifications on leaving school and higher rates of unemployment compared with other young people in England.

In 2018, the National Crime Agency assessed that the threat from serious and organised crime is increasing and that organised criminals are continually looking for ways to sexually exploit new victims, preying on the most vulnerable in society, including young children, with devastating and life-long consequences (Home Office Serious and Organised Crime Strategy 2018).

> *"Child sexual exploitation (CSE) is a form of child sexual abuse. It occurs where an individual or group takes advantage of an imbalance of power to coerce, manipulate or deceive a child or young person under the age of 18 into sexual activity (a) in exchange for something the victim needs or wants, and/or (b) for the financial advantage or increased status of the perpetrator or facilitator. The victim may have been sexually exploited even if the sexual activity appears consensual. Child sexual exploitation does not always involve physical contact; it can also occur through the use of technology."* *(HM Government: Statutory definition of child sexual exploitation - Feb 2017).*

The *WMP Strategy for Tackling CSE* (2017) states that CSE is a complicated issue with multiple components which makes resourcing decisions challenging when there are limited resources available. It highlighted the difficulty in developing a reliable picture of the demand, because of under-reporting and the fact that relevant data is held across multiple agencies. It notes that whilst harm through physical contact is greater, online grooming acts as a key gateway to physical exploitation and estimated that about 3,500 children aged 11-16 years within the West Midlands may have shared indecent images of themselves. Nationally, the number of referrals to the NCA relating to online child sexual exploitation has increased by 700% in 4 years *(Home Office Serious and Organised Crime Strategy 2018)*.

The WMP strategy also identified that the pathways to CSE which include organised exploitation by criminal networks may produce the highest levels of harm; yet existing risk frameworks tend to focus on victims rather than offenders or locations.
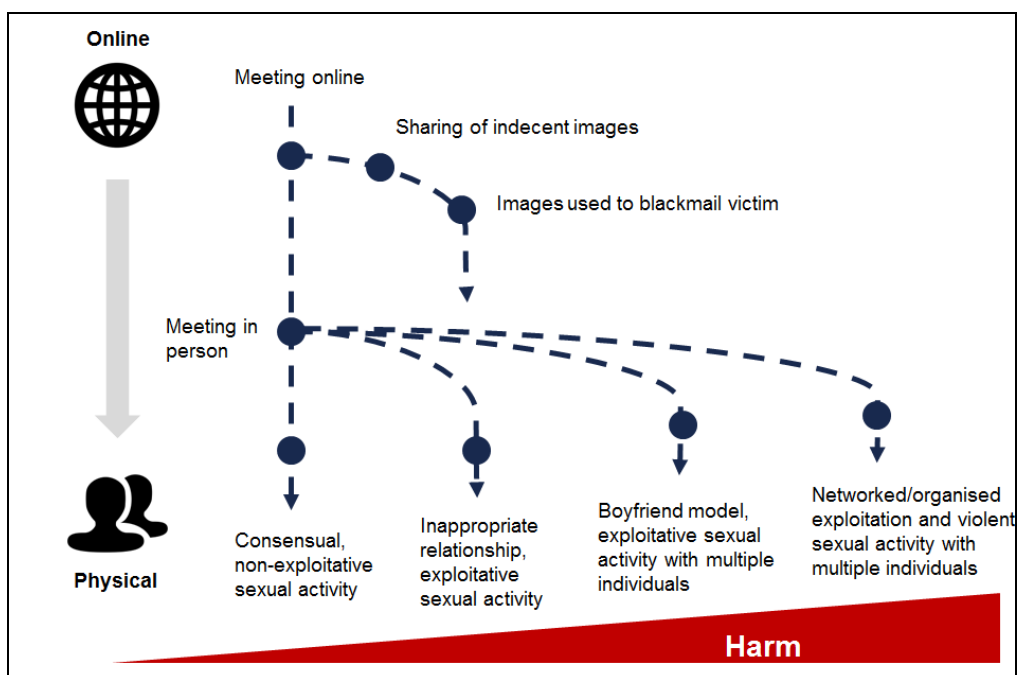
Our current understanding of the extent of such networks and the threat posed by them to exploit the most vulnerable in our communities is not fully understood. A key question raised by the WMP strategy was '*How do we assess the risk of the demand profile so that we might focus resources in the right places at the right times?*'

Therefore, in line with the Force's mission to protect the public, the identification and assessment of organised criminal networks and nominals associated with CSE is a priority.

Intended activity resulting from the project

Force Intelligence will use the evidence from this project to identify and prioritise those OCGs actively engaged in the sexual exploitation of children. This will enable the Force to maximise the impact of its limited resources by focusing on dismantling those OCGs which cause the greatest harm.

Force Intelligence will identify opportunities to work alongside partner agencies to understand the risk and also enable the sharing of information and joint agency working.

Source of analytical question / hypotheses to be examined:

The business question was posed by the Intelligence Department.

Data to be used:

Level of analysis:

☑ Individual
    Individuals aggregated?
    ☐ Yes
    ☑ No
☐ Specific Area:
    ☐ Output Areas
    ☐ Super Output Areas - Lower
    ☐ Super Output Areas - Mid
    ☐ Wards
    ☐ Districts
    ☐ Other

☐ West Midlands
☐ Other

Reliability of data:

The data are sourced from WMP systems. The assessment of the quality of the data undertaken for DAL_2019_0002_SOC_Networks will inform this project coupled with an assessment of data quality regarding CSE. Any data quality issues will be noted and where applicable will be incorporated into the project (e.g. by excluding some data from a system if it is felt to be unreliable). These data, as part of the analytical project life cycle, will also be assessed for missing values, etc.

These systems are those currently used by WMP in their day-to-day business. Specifically in the case of intelligence data, these will be examined as to their veracity, source, etc. prior to inclusion (i.e. only intelligence considered to be credible from credible sources will be used).

Sample or entirety:

If sample: N/A

Method of sampling: N/A

Method of choosing sample size: N/A

Sample size: N/A

Type of analysis:

☑ Exploratory
☐ Explanatory
☐ Predictive
☐ Optimisation

Proposed methodology:

The first element of the project creates the network based upon identifying individuals who have had an involvement with CSE. Crimes, intelligence and other systems are then used to identify if a (potential) relationship exists between individuals (e.g. have they ever been arrested together, etc.). Following this, a graph would be constructed of the linkages between individuals, including edge values and temporal information. This would allow for the identification of links between the various sub-graphs, calculation of centrality within the various networks, calculation of the levels of harm created by these networks, etc. This would also include examining the systems for victims. This would then also couple this network with the SOC network in order to identify links with organised crime groups.

Will the project eventually be automated:

☐ Yes
☑ No

At present it is envisaged that this would be built to investigate the extent of CSE related networks. It is possible that future ad hoc runs could also be undertaken as future requirements dictate.

Means of evaluation:

Not applicable.

ALGO-CARE considerations:

Advisory:

If applicable, are the outputs from the algorithm to be used in an advisory capacity?

The project would enable a delineation of CSE networks and any links through to SOC networks, the links within and between them and measures of centrality. Assessing the harm created by individuals through their criminal activities will also enable the extent of the harm created within the WMP area resulting from these networks to be assessed.

This information can then be used to focus further enhanced intelligence work.

Does a human officer retain decision-making discretion?

The project would aim to highlight the extent and membership of networks of individuals involved in CSE with the view to informing further intelligence related work. Other than direction of further intelligence work, no decisions would be based upon the findings from the project.

Lawful:

What is the policing purpose justifying the use of the algorithm (means and ends)?

CSE creates harm within the WMP area via the exploitation of vulnerable members of the community. As such the investigation of CSE activities is a high priority. At present, the extent and nature of CSE related activities is not known and therefore assessing the extent of CSE related networks and the harm generated by them is a necessary step to investigating and preventing such activities.

Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes?

The project only utilises data collected from WMP systems in their normal day to day activities. Assessing the priorities that should be tackled by the Police in their operations due to the risks and threats present is a legitimate policing purpose and this project would feed into that process.

In what way will the tool improve the current system and is this demonstrable?

There is currently no system or tool to undertake large scale network analysis or to delineate the harm created via any such networks.

Are the data processed by the algorithm lawfully obtained, processed and retained, according to a genuine necessity with a rational connection to a policing aim?

The data are gathered in the normal day to day operations of WMP in line with the aims of WMP.

Is the operation of the tool compliant with national guidance?

The analyses proposed would accord with DCMS Data Ethics Framework 2018.

Granularity:

Does the algorithm make suggestions at a sufficient level of detail given its purpose and the nature of the data processed?

The analyses would use information at the level of the individual to develop the networks and assess the overall levels of harm created by the networks.

Are data categorised to avoid broad-brush grouping and results and therefore issues of potential bias?

The underlying data are not categorised.

Do the potential benefits outweigh any data quality uncertainties or gaps?

The project could help focus resources of WMP intelligence and lead to enhanced assessments of the threat posed by networks towards vulnerable members of communities. Given the harm that CSE generates within the West Midlands uncertainties as to data quality (bearing in mind such issues

will be investigated in the preliminary stages) are acceptable.

Is the provenance and quality of the data sufficiently sound?

The data have been gathered during the day-to-day investigative work of WMP and enable analyses of the type envisioned for this project.

If applicable, how often are the data to be refreshed?

The underlying data are refreshed on a sub-daily basis as they are part of WMP's core systems. The analyses may be run intermittently on an as and when needed basis but do not require specific data extracts.

If the tool takes a precautionary approach in setting trade-offs, what are the justifications for the approach taken?

Not applicable.

Ownership:

Who owns the algorithm and the data analysed?

WMP once developed.

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 data and the analyses are contained wholly within the WMP Hadoop system and the security measures employed therein.

Challenge:

What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias?

Any findings relating to particular individuals that are deemed useful for further analysis would be subject to the normal intelligence processes of WMP and as such any incorrect findings could be notified to the Lab and incorporated into any future runs of the analyses.

If the algorithm is to inform criminal justice disposals, how are individuals notified of its use?

Not applicable.

Accuracy:

Does the specification of the algorithm match the policing aim and decision policy?

The nature of the analyses chosen have been determined to be the best means of addressing the research question.

Can the accuracy of the algorithm be validated periodically?

Not applicable.

Can the percentage of false positives / negatives be justified?

Not applicable.

How was the method chosen as opposed to other available methods?

Due to the nature of the research question (network analysis).

What are the (potential) consequences of inaccurate forecasts?

Not applicable.

Does this represent an acceptable risk?

Not applicable.

How are the results checked for accuracy and how is historic accuracy fed back into the algorithm for the future?

Not applicable (other than any intelligence analysis arising from findings as noted above).

How would inaccurate or out-of-date data affect the result?

If data were to be wholly inaccurate then the analyses would essentially provide inapplicable findings. The Lab will seek to minimise this potential through a thorough analysis of the data and their pitfalls, issues and overall nature as well as through discussions with SMEs.

Responsible:

Would the operation of the algorithm be considered fair?

The analyses will be fair in that each data point will be considered on its own merits.

Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review?

The nature of the intended method(s) is such that the end-use is to provide information rather than predictions that feed into a process on an on-going basis.

Would it be considered to be used in the public interest and to be ethical?

In the face of the harm generated through CSE and in the interest of the more efficient allocation of resources, this project would be in the public interest.

Explainable:

Is information available about the algorithm / decision-making rules and the impact of each feature?

Information about the algorithms used would be available, however no decisions would be made *per se* and the type of analyses are such that measures such as feature importance are not applicable.