

Goldstein Awards Application 2019

Summary of application

Insight Bristol's CSE Model

Scanning:

On the 24th March 2016, the Bristol Safeguarding Children Board (BSCB) published a Serious Case Review for Operation Brooke, a large-scale police investigation into the sexual exploitation of children in Bristol. The review published a series of findings which highlighted a number of national, institutional and operational failings. Embedded in their areas for improvement lay a need to move to a proactive, as opposed to reactive response to Child Sexual Exploitation (CSE). It was therefore decided that the Insight Bristol team, an interagency analytics hub, would utilise predictive analytics to identify potential victims and perpetrators of CSE.

Analysis:

Insight Bristol broke down the problem of producing an accurate, reliable and pragmatic CSE predictive model by identifying how the findings of the Serious Case Review into Operation Brooke could be overlaid onto the Problem Analysis Triangle (PAT). Statistically modelling the profile of victims through analysis of existing victims with Barnardo's Against Sexual Exploitation (BASE) workers allowed the team to better understand underlying risk across the population as a whole. The application of the project analysis triangle allowed the team to construct a model which effectively and accurately identified emerging CSE risk.

Response:

It was important to design a response to simply and efficiently give all of the key agencies the key outputs from the model, whilst simultaneously aiming to mitigate the areas for improvement outlined by the BSCB. The deployment of the model was therefore designed to meet a number of targets. These ranged through from alerting key agencies about risky individuals who aren't currently being worked with, to aiding commissioning and service planning.

Assessment:

The iterative, agile approach adopted by the Insight Team was key to assessing the success of the model. Feedback was constantly gathered, considered and then incorporated, ensuring the model to be as functional and operationally relevant as possible. Critical academic analysis was also sought out, as well as advice from experts in the field of CSE.

Number of words: 321

Description of project

Scanning:

On the 24th March 2016, the Bristol Safeguarding Children Board (BSCB) published a Serious Case Review for Operation Brooke, a large-scale police investigation into the sexual exploitation of children in Bristol. The Case Review centred around two trials that concluded in November 2014, where thirteen offenders were sentenced for a combined total of 116 years imprisonment. The review published a series of findings which highlighted a number of national, institutional and operational failings, with the area-specific findings as follows:

- The multi-agency system “is not set up to provide an effective response for adolescents (including those at risk of CSE)”, with a, “fragmented and reactive response”.
- There was a pattern of focusing on stopping victims having further involvement with perpetrators as opposed to preventing the abuse in the first place. There was also not enough done to disrupt the behavior of the perpetrators.
- The current systems did “not reliably identify patterns in individual and group behavior. This “reduces the chances of a timely response in the detection of victims and perpetrators of child sexual exploitation and leads to a more reactive rather than proactive approach”.

The common theme of these local findings clearly pointed to a lack of a proactive response to child sexual exploitation (CSE).

CSE is in many respects a ‘hidden issue’; it is extremely difficult to ascertain the true scope and scale of the problem. Many victims don’t see themselves as such, and therefore don’t cooperate with the police. Often victims are led to believe that they are in a loving and consensual relationship with the perpetrator, who then takes advantage of this emotional attachment to exploit and manipulate their victims. In other cases, perpetrators coerce their victims through violence and threats. It is sometimes the case that victims are unable to effectively report their perpetrators as they only know nicknames or aliases, and can only provide physical descriptions of their attackers. Furthermore, there is frequently a delay in the reporting of CSE offences; in known offences, on average it will take a month to report the incident from the incident taking place. It is also challenging to identify victims of online exploitation. The result of all of these factors is a significant underreporting of CSE within the City of Bristol.

What is certain, however, is the devastating impact that CSE has on individuals and communities. Exploited children face a significant reduction in their emotional wellbeing, and a marked deterioration in their physical and mental health. Positive outcomes are hugely affected by CSE, and, in what is described as ‘deeply troubling’ by the Office of the Children’s Commissioner, of the 2409 victims reported in a 2012 report, 155 were identified as also being perpetrators of sexual exploitation. It has been found that victims are often encouraged by their perpetrators to ‘recruit’ new victims through their social network. Beyond sexual exploitation, victims are also often coerced into committing criminal acts involving drugs and theft by their exploiters. CSE is undoubtedly a scourge, one that breeds crime and impacts victims for the duration of their lifetime.

In order to deliver the proactive and pre-emptive response that the BSCB identified as necessary, it was decided that Insight Bristol, a multi-agency analytics hub between Avon and Somerset Constabulary and Bristol City Council, would design, develop and deploy a predictive analytics model which would proactively identify potential victims, perpetrators and facilitators of CSE. The following six objectives were established for the model:

1. Use predictive analytics in order to:
 - i. Identify and prioritise CSE victims and children at an emerging risk of CSE.
 - ii. Identify and prioritise CSE offenders and facilitators.
2. Ensure that all agencies are considered by fusing data from Avon & Somerset Constabulary, the School Census, Housing, DWP, Social Care & Early Help, Safer Bristol, and other local authority datasets, in order to

uncover hidden areas of threat harm and risk.

3. Use the Problem Analysis Triangle method in order to statistically summarise CSE risk in a victim or offender's social network, inferring hidden victim/offender relationship links using incidents, intelligence, crimes and locations.
4. Provide a practical feed into intelligence scanning and tactical tasking processes in order to maximise safeguarding, investigation and disruption opportunities.
5. Work hand-in-hand with the Intelligence & Tasking Directorate's deployment of IBM iBase, delivering greater efficiency and depth in the creation of analytical products.
6. Facilitate the scalability of Avon & Somerset Constabulary's operational response to CSE, by automating the identification and prioritisation of CSE threat harm risk across the whole dataset.

Analysis:

Insight Bristol broke down the problem of producing an accurate, reliable and pragmatic CSE predictive model by identifying how the findings of the Serious Case Review into Operation Brooke could be overlaid onto the Problem Analysis Triangle (PAT). This was done in collaboration with Operation Topaz, Avon & Somerset Constabulary's operational response to CSE. The findings were as follows:

“The multi-agency system is not set up to provide an effective response for adolescents... (it provides a fragmented and reactive response.” This can be assimilated to the Victim side of the PAT. As a partnership team sitting between Bristol City Council and Avon & Somerset Constabulary, Insight Bristol had already pioneered the Think Family Database, which fuses person-level data from over thirty sources. The Think Family Database was born out of the national Troubled Families programme, which rewarded councils with funding for proactive responses to vulnerability. The approach taken by Bristol City Council was to work with its partners, including Avon and Somerset Constabulary, to amalgamate as much public sector information as possible on social issues relating to an individual. This addressed fragmentation of data and response in a two-fold manner; predictive modelling could draw upon this whole dataset of variables, including Constabulary, the School Census, Housing, DWP, Social Care & Early Help and Safer Bristol data; and the Think Family Database already has an expanding user group across the partner agencies providing the data. This ensures that agencies are tasking from a single version of the truth.

Another problem related to defining a sizeable, stable Victim Cohort upon which to train a predictive model. As outlined within the 'scanning' section, CSE operations tend to be intelligence-led, often without a reported crime. Complaints, when they arise, are often historic, which may present data quality and consistency issues across multiple and legacy systems, superseded ways of working and former operating models. We therefore decided to use the cohort of young people who had been allocated Barnardo's Against Sexual Exploitation (BASE) workers. This group represents children who have been referred, triaged for risk by CSE professionals and accepted onto the BASE programme – a group which is therefore not reliant on a reported or prosecuted crime for its definition. In order to capture how the target cohort was represented in data prior to multi-agency problem solving, the cohort was regressed to their BASE start of service date. This captured how they were represented at the time of peak threat harm risk, prior to mitigation. We could then identify which indicators were typically held of CSE victims and their families at this peak risk level, with common indicators shared by victims identified: these indicators include children going missing; educational attendance or exclusion; NEET and risk of NEET; criminal offences and ASB; Social Care and Early Help episodes; teenage pregnancy; homelessness; Child Protection Plans; Looked After Children; Children In Need plans; Domestic Abuse at individual, perpetrator and incident levels; mental health; and drug/alcohol incidents or support. The above indicators are also aggregated for the individual's family within the model. The resulting police and partner dataset contains over ninety indicators describing the risk relating to an individual. All subjects are scored and prioritised according to their statistical similarity to subjects in the target victim cohort. The final prioritisation takes whichever score is highest – the child's predictive score or their derived index score. This method combines the best of machine learning with the operationally-acquired expertise of experienced Criminal

Intelligence Analysts.

“There was not enough done to disrupt the behaviour of the perpetrators... preventing the abuse in the first place.” This is linked to the Offender side of the PAT. One issue with risk modelling CSE is that there is no offender cohort of sufficient size and stability upon which to train a predictive model to identify emerging offenders. There is also no single profile for an offender type that enables preventative action – reliable CSE indicators are historic events that have already happened, and the variety of offender types involved in committing and facilitating CSE problematizes the identification of leading indicators. We addressed this by incorporating organisational learning from Operation Topaz that a victim’s CSE risk is adversely affected by the proximity of offenders in their social network. We built a statistical model of each subject’s social network that provided the number of associates carrying CSE risk, a measure of that risk, and a measure of the proximity of that risk to the subject in their social network. Early identification of potential offenders can therefore be carried out via their social network, prior to the commission of crime and the submission of intelligence.

It is important that subjects identified by this method are not labelled as CSE offenders or potential offenders without further investigation and intelligence development. Subjects carrying a large amount of CSE risk in their social network may be witnesses or informants to liaise with; unwitting facilitators through their social circle who may require minimal management; or vulnerable people at risk of criminal exploitation themselves. The model cannot differentiate between these groups, and so the output of the model that has been designed to identify offenders has been branded “CSE Risk – Subjects Over 18.”

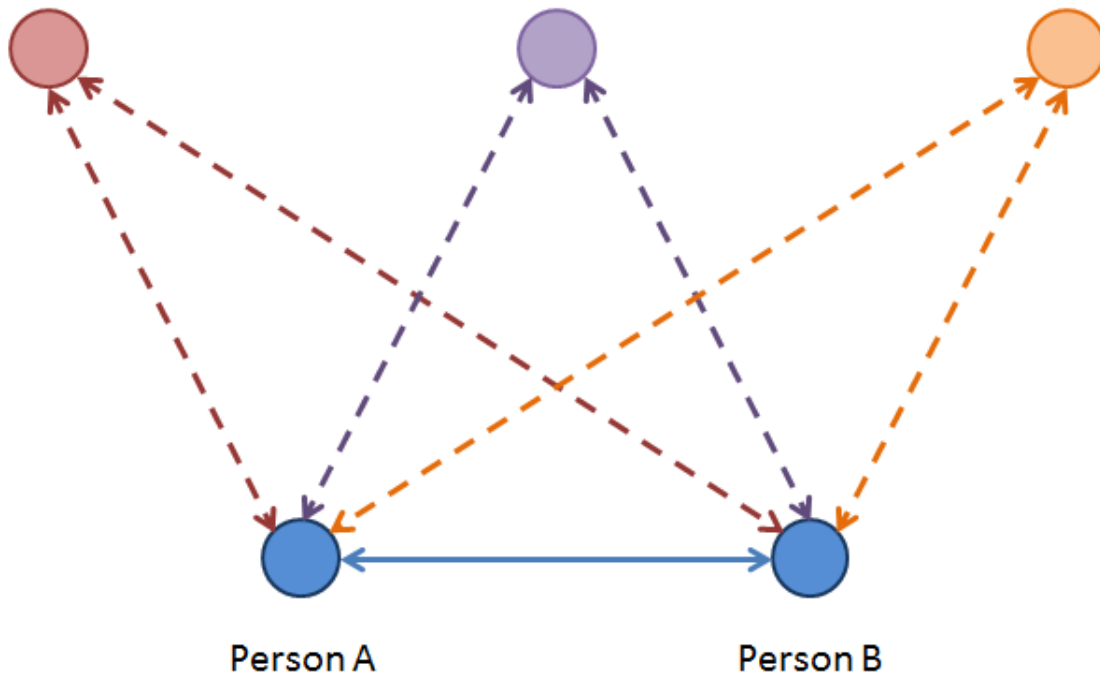
“The current systems do not reliably identify patterns in individual and group behaviour.” This was relevant to both the Place side of the PAT and the Incident/Crime at the centre of the triangle resulting from the proximity of victims and offenders at the same location. In its infancy, the CSE predictive model focused on aggregating linear event or state data in order to highlight individual vulnerable subjects who were statistically similar to the target cohort. This approach only told half the story, as it neglected the lateral risk in the victim’s social network, their presence at high risk locations or hidden victim/offender relationships implied through mutual links to the same events. This therefore led to a more sophisticated model to be developed which could take these factors into account.

The creation of a statistical model of a subject’s social network detailed above was problematized by the lack of relational linking in Avon & Somerset Constabulary data. Crime and Intelligence Recorders link victims and offenders to crime and intelligence, but not necessarily to each other – even when a clear social, criminal or familial relationship can be inferred from the context of the report. We compensated for this data quality issue by statistically inferring relationships from events, which we defined as any spatial-temporal occurrence that brought two people together at the same time, at the same place. The current model therefore relates two subjects linked by three or more events, whether that is the same crime; the same intelligence report; the same missing person report; or that they are the subject of PNC research relating to the same command and control incident. The threshold of mutual links to three or more events mitigates against false positives that may arise using this method. Our method of statistically deriving relational links is shown in the diagram below:

Intelligence

Occurrence

Call Card



The application of the project analysis triangle allowed the team to construct a model which effectively and accurately identified emerging CSE risk.

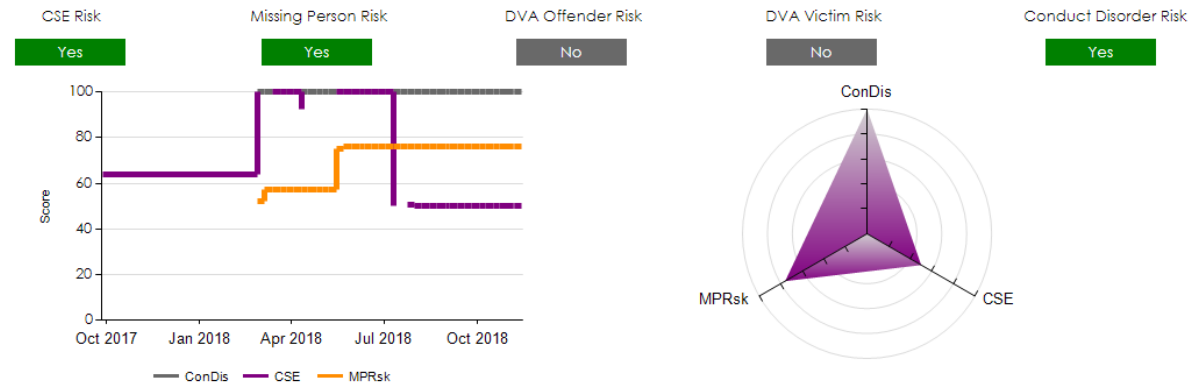
Response:

A recurring phrase within CSE literature reads, “protecting children from sexual exploitation is everyone’s responsibility”. It was therefore important to design a response to simply and efficiently give all of the key agencies the relevant information needed, whilst simultaneously aiming to mitigate the areas for improvement outlined by the BSCB. The deployment of the model was therefore designed to meet the following five targets:

- Proactively identify at risk individuals who aren’t currently known to the police/council, and to start working with these individuals/families.
- Notify existing caseworkers that CSE might be a risk factor in their cases that hasn’t been considered.
- Give information to Operation Topaz to aid in their investigations.
- Make CSE risk information known to other safeguarding stakeholders: school safeguarding leads, GPs, etc.
- Assist in service planning, through a strategic understanding of demand.

There are two outputs given by the CSE model. A simplified version delivers the final risk score calculated alongside a text summary of each subject’s risk indicators into the Think Family Database (TFD). Therefore, any staff member who has access will be able to see if the individuals that they are working with, or any individuals that they have safeguarding concerns for, have been deemed at risk of CSE by the model. This text summary allows individuals who are less technically literate to understand the risk factors relating to individuals and their area. Also incorporated into this ‘risk profile’ are a number of other predictive risk models produced by the team, including Missing Person risk (the risk that someone who has already gone missing once will go onto be a recurring missing person) and Conduct Disorder risk (where an individual hits a certain number of flags relating to conduct disorder). The resulting risk profile gives lead professionals a broad understanding of the different aspects of risk being held by their cases,

allowing them to deliver a more nuanced and targeted approach.



Indicator	Indicator Date	Value	Notes
MPRsk	12 Nov 2018	76.000	High
CSE	24 Sep 2018	50.200	This person was deemed at risk because they are monitored by Op Topaz; they associate with subjects at risk of CSE; call scripting identified CSE and other risk factors when they were reported missing to the Police This person has the following contextual risk factors - they are on a CP Plan This person has the following contextual risk factors relating to at least one member of their family - they are on a CP Plan and have been reported as homeless; they are unemployed; they are linked as involved and victim to DA incidents and offences; they have mental health problems and are receiving support for drugs or alcohol

Historically, access to the TFD has been only logistically possible for Bristol City Council staff and selected Police Staff due to network and information security. However, the value of sharing information between agencies is increasingly clear when discussing safeguarding children, as also identified by the BSCB in their areas for improvement. Therefore, getting access to external safeguarding stakeholders such as school safeguarding leads and designated safeguarding GPs has become a priority for the Insight Team. Advanced technology is being utilised in order to do this, which when finished will equip schools and safeguarding leads with the the ability to look up and understand the whole spectrum of issues, including CSE risk, affecting a child. Schools will also be able to easily generate a list of all pupils who are marked as attending their school who've been flagged as at risk of CSE, whilst also being able to see what other agencies are involved in the care and protection of those children. The result of this is a more unilaterally aligned approach through interagency working to tackle CSE, as a number of agencies will be seeing the same information.

A more comprehensive and technical output is delivered to the Intelligence Analysts attached to Operation Topaz. They use IBM iBase and i2 methods to develop the intelligence picture around subjects with either the highest risk or those with emerging risk. As the Insight Team adopts an Agile methodology when delivering projects. Topaz were given access to the model in the summer of 2018, and there is continuous feedback between them and the Insight Team to further develop the model and understand how it can be improved to help them with their investigations around CSE.

The advantage of fusing together the CSE output with the indicators within the TFD is that strategic service planning around CSE can be undertaken. The team has ranked each school in terms of a number of vulnerability factors, including the amount of CSE risk held, which has informed a number of strategic responses within the Council and the Police. The Early Help branch of the council have used this banding to allocate social care resource to schools, and the Insight Team is currently exploring applying these bandings to assist with the allocation of PCSO resource.

Assessment:

In order to assess the effectiveness of access to the CSE model, the team first had to be confident that the model actually works in identifying those at risk of CSE. This was done through running the model to see if it identified all of those included in the targeted cohort; those who have a BASE worker attached to them and therefore have had a history of sexual exploitation. This training set was used for the baseline because there are comparatively few current examples of children who are confirmed victims of successfully prosecuted cases of CSE. Most complaints take place when the children are older, reported on an historic basis – which makes it less likely that sufficient current data will be held on individuals that can be used for predictive modelling. The BASE cohort represented children who had been subject to a triage and referral process, resulting in a worker being allocated according to agreed need. One issue identified early in the project was that a mere cohort list of vulnerable children with allocated BASE workers was insufficient to use as a baseline. BASE workers manage and mitigate the risk around the children they work with; this manifests as noise in statistical modelling because the allocation of the BASE worker reduces the risk. This issue was resolved by capturing the BASE Start of Service for each subject, enabling us to regress the data back to the time that each individual was adopted into the cohort. The final baseline for training the machine learning was therefore the BASE cohort at their Start of Service date – in other words, at their point of peak unmitigated, unmanaged risk.

In order to be successful the model would need to identify both the BASE cohort at the start of their journey, and also individuals who weren't currently known to the Police/Council. When run, 69% of the individuals identified were of the target cohort, meaning that 31% of those identified were individuals that had CSE risk that were previously not identified as having CSE risk associated with them. The model also identified 94% of individuals within the target BASE cohort. The team could therefore say with confidence that technical performance of the model was sound in identifying individuals who have a potential risk of CSE. When deploying the model to lead professionals, the missing 6% of known individuals could be accounted for by incorporating into the model all individuals with markers on them by Operation Topaz, as to ensure that no known risky individuals were missed out of the model. This type of assessment, known as 'precision and recall' is a quick and easy way to test the technical effectiveness of the model, however, what is much more important is assessing the impact that the CSE model has had on the working practices of the people that it has been given to, and how it has helped to remedy the areas for improvement outlined by the BSCB.

The CSE model was developed in accordance with the Agile methodology. This involves regular, immediately useable deliverables; a continuous end user collaboration and feedback; and frequent, regular project milestones referred to as 'sprints.' The Agile methodology delivers regular checkpoints for quality assurance from both developers and data scientists within the Insight Bristol team, as well as from end-users. Issues are therefore identified early and addressed in a prompt manner, with feedback immediately possible after the next iteration. The iterative approach to the way in which the Insight Team works has been critical to the success of the model. When initially deployed to lead professionals, the output of the model was sent as a datasheet. However, feedback showed that having access to multiple systems, and having to continuously log in and out of different applications/spreadsheets meant that engagement with the model was significantly reduced amongst staff. It was also found that this method was difficult to update: weekly spreadsheets were being sent out, and it was very easy for people to exchange old spreadsheets without the most up to date information. The exchange of outdated information posed a serious and obvious risk, eliciting the need for a more streamlined response from the team. These issues were responded to by integrating the output of the model into the TFD, which meant the need to log into multiple systems was eliminated, and that everyone was always seeing the most up-to-date information.

Another example of benefits derived from this iterative approach related to the aggregation and fusion of multi-agency Think Family data. It was recognised that by only using Think Family data from Bristol, only individuals in Bristol could be identified by the model. The model therefore required the rationalisation of data between Bristol and Somerset Think Family programmes. As the two programmes collected and aggregated data from different sources in different ways, the data needed to be aligned where possible, resulting in the re-training of the model

onto a smaller set of indicators that were common between both programmes. Sacrificing some of the depth of the model allowed for an identification of individuals across a higher area within the force, deriving benefits across both counties. Future consideration is being given using each programme's complete array of indicators, resolving the difference using two predictive models.

There has been significant effort from the team to receive formal external evaluation for the effectiveness of the model. A PhD student was commissioned spent 3 years evaluating the use of predictive analytics in policing, with a special focus on the work being done by the Insight Bristol team. This form of evaluation has a number of positive outcomes: a deeper understanding of the identification and response to CSE is held within the academic sphere, and the work done by the team can be recognised nationally as exemplary, allowing other police forces and teams to adopt the approach taken by the Insight team. Insight Bristol have held a number of knowledge transfer days with other 'Troubled Families' teams, allowing for the sharing of this best practice and the protection of children nationally, not just within the City of Bristol.

One of the key findings of the BSCB was the need to implement a proactive, rather than reactive, response to the challenges of CSE. As a result of using the model alongside other responses to vulnerability (such as the TFD), a total of 243 cases were allocated to key workers in the period 2017-18. The risk scores, including the CSE, played a key role in facilitating the allocation of targeted support to these families.

In conclusion, the CSE model created by the Insight Team has done much to address the findings of the BSCB in the wake of Operation Brooke. Though there is still work to do in terms of streamlining the model and embedding its outputs into working practices, significant steps have been made within all of the areas for improvement outlined by the BSCB. The multi-agency approach to the formation of the model has allowed for the 'fragmented' interagency response to be overcome, with all stakeholder agencies involved in CSE having the ability to see the same, complete picture as other agencies. Through network modelling, perpetrators and risky individuals can now be highlighted at a rate not possible for human analysts, allowing for the disruption of patterns of behavior and greater intelligence work. Finally, it is important to consider that despite the considerable impact already seen as a result of the model, it is still in its infancy; as new technologies become available and are embraced even more perpetrators will be identified, and more children will be protected from the scourge of Child Sexual Exploitation.

Insert supporting diagram/charts etc. here

Number of words: 3899

Endorsement by Senior Representative

Please insert letter from endorsing representative:

As the manager of the Insight Team I am proud of the work that has been done to create the CSE model. It has allowed for the identification of a number of vulnerable individuals, potential perpetrators and is a fine example of the power of predictive analytics. I am looking forward to seeing how the model develops going into the future, where it will continue to help to identify hundreds of vulnerable children and help to disrupt and prosecute the very worst offenders in our society.

Tom Fowler