**ETHICS COMMITTEE – Formal Group Meeting**

**Wednesday 4th December 2024**

**10:00-13:00**

Meeting held virtually via Microsoft Teams

**Present**:

|  |  |  |
| --- | --- | --- |
| Marion Oswald |  | Chair of Ethics Committee |
| Jonathan Jardine |  | Chief Executive (OPCC) |
| Derek Dempsey |  | Ethics Committee |
| Tom Sorrell |  | Ethics Committee |
| Jennifer HousegoClaire Paterson-Young |  | Ethics CommitteeEthics Committee |
| Jack Tracey |  | Criminal Justice Policy Lead (OPCC) |
| Edward Hunter |  | Policy Intern (OPCC) |
| Davin ParrottOctavian Bordeanu |  | Data Analytics Lab (WMP)Senior Data Scientist (WMP) |
| Tom Joyce Matt WelstedMarc WilliamsIan Parnell |  | Chief Superintendent – Head of Corporate Development (WMP)Assistant Chief Constable (WMP)IT and Digital (WMP)Chief Superintendent – Head of Force Contact (WMP)  |
|  |  |  |

**Apologies:**

|  |  |  |
| --- | --- | --- |
| Kerry Reidy |  | Ethics Committee  |

|  |  |  |  |
| --- | --- | --- | --- |
| **1** | **10:00** | **Welcome**The Chair opens the meeting and welcomes members. It is suggested that members should have received a new draft of the terms of reference for the committee as a result of the last meeting. A member makes the committee aware that the force has received a letter from Amnesty International about predictive analysis and AI in which they call for a ban.  | **Marion Oswald** |
| **2** | **10:05** | **Ethics Committee TORs and Expansion**The presenter suggests the value of the Committee and expansion of advanced technologies and data has encouraged an intention to expand the remit of projects brought before the Committee. A focus on operational ethics will be maintained, however.  Both WMP and the OPCC express an intention to formalise the governance route of the Committee’s advice and embed its advisory role within governance structures. To do so it was suggested operational leads attend the Committee, there was a need to respond to the BRAID report and that a clear process of how Committee advice directly links to wider governance structures should be established to enhance the understanding of community interests and real-world impacts whilst increasing accountability in a structured way. As such a redraft of the TORs is necessary and a name change from “committee” to “panel” suggested but core principles will remain the same**Questions and Responses**A member expresses concerns that renaming to “Ethics Panel” and removing data from the title is misleading and reduces identifiability. * The presenter accepts there is a need not to lose the technology/data focus and that the Chair will retain authority on what is brought before the committee. Meanwhile, it is suggested “committee” suggests the group makes decisions whilst “panel” emphasises the advisory role better.

A member suggests that due to the Committee’s focus on digital technology, membership skills should better reflect this focus.* It is suggested the new draft of the TORs attempts to define the projects brought before the committee as the use of data technology and innovative policing models.

A member raises that the increased workload risks making the group’s work and advice superficial without time to investigate thoroughly.* The presenter suggested the definition of the Committee’s remit in the new TORs must ensure it remains manageable and significant but admits there are considerations for expanding membership.
 | **Matt Welsted** |
| **3** | **10:30** | **Andi-Esra**The presenter explains that Andi-Esra is a language understanding model which has successfully mitigated the issue of 101 calls overwhelming the capacity to respond, boosting response rates through:* Understanding the reason for the call and resolving transactional demand types
* Redirecting callers to appropriate personnel
* Recognising signs of vulnerability and prioritising those calls

The presenter emphasises the ambition for Andi-Esra to improve satisfaction rates alongside response rates as the current IVR system has a high level of abandonment.The presenter continues by suggesting that WMP adopts responsible AI practices adding that they don’t see AI as a solution to everything but are grounded in technology and continue to be data driven.The presenter concluded by explaining that the proof of concept has been running for 12 months and a formal report on the results and data from it has been submitted to the Police Chief Scientific Advisor’s Office.The new formal project, titled Andi-Esra 2.0 have been built and are scheduled to go live in December 2024 and be used for 100% of 101 calls. The new features made available in Andi-Esra 2.0 are:* A capacity for customer feedback
* Dashboard reporting
* Being used to replace the web chat on WMP’s website
* 12 new call flows through which to direct the topic/purpose of the call
* The ability to take caller’s details whilst waiting in the queue, saving time once put through
* The potential to be used as a WMP switchboard

**Questions and Responses**A committee member asks whether the formal report will be made available to the public or restricted to just the committee.* The presenter expresses the belief that it will be made public but will check and report back.

A committee member enquires whether a comparison of andonment rates between Andi-Esra and IVR has been conducted.* The presenter replies no such analysis has been done, nor will it be possible as there is no data being collected on IVR.

The member continues by asking whether callers are made aware they are speaking to a chatbot, suggesting they should be.* The presenter reveals that it is not explicitly stated but is fairly obvious.
* The presenter continues that WMP will adhere to standard, national practice in this regard.

A committee member expresses concerns that vulnerable people will struggle to communicate with the AI and having to subsequently repeat themselves to a human operator may be retraumatising.* The presenter notes the concerns of the member and agrees there is a need to understand the challenges faced by different individuals.

A committee member inquires as to the disaster recovery systems that are in place.* The presenter suggests that the system functions through AWS cloud systems who manage disaster recovery whilst removing most points of failure.
* The presenter adds that a manual failsafe is also available that can divert all calls to human operators.

The member asks how often this failsafe procedure is practiced.* The presenter suggests it isn’t.

A committee member asks to what extent the system can be monitored to ensure it is functioning correctly and if any aberrant behaviour has been seen.* The presenter explains that there have been situations where the technology has misunderstood what has been said.
* The technology will default to a human operator when it cannot understand.
* No formal assessment of performance has been conducted.

A committee member expresses concerns about how it may make callers uncomfortable talking to a robot and frustrated should they become stuck in an AI loop.* The presenter seeks to emphasise that Andi-Esra also aims to improve caller satisfaction.
* As a part of this, any detected frustration will automatically reassign the caller to a human operator.
* The caller also has the option to request to be reassigned.
 | **Ian Parnell****Marc****Williams** |
| **4** | **11:10** | **Video Analytics**The presenter explains that the project consists of using object recognition technology in order to increase the efficiency and effectiveness of analysing CCTV recordings. The technology processes the video, highlighting certain characteristics and objects which can subsequently be searched for and presented instantly.The presenter uses the example that one could search for all the red cars that appear in a section of footage and be shown all those appearing in the video immediately.The presenter emphasises that this will have no implications of the evidential process in court as a human will always review the results to determine the evidential value, not the technology.In demonstrating the effectiveness of the technology, the presenter exhibits a case study in which 24 hours of footage took a human seven days to analyse whilst the technology took 26 minutes to process and 35 minutes to be reviewed by a human. The technology also identified two additional sightings that the human missedThe presenter explains to the committee that, at present, footage is being double processed with both humans and the technology reviewing it in order to help identify areas in which the technology is weak. The presenter concludes by suggesting that the technology is currently limited to high-harm crimes but that there is an interest in its use for less serious inquiries. The technology also has the potential to be expanded later for real-time object recognition in order to identify weapons, disorder or diver behaviour. **Questions and Responses**A committee member asks whether it is likely to be rolled out on a national basis.* The presenter suggests other forces are also trialling it but no evaluations have taken place.
* The presenter has contacted the National Police Chief’s Council in order to coordinate with any upcoming evaluations.

A committee member expresses concerns about the technology’s real-time potential contributing to a surveillance state.* The presenter assures the member that they technology is not currently used in this way and that prior to any expansion, a community engagement process would take place.

The member continues in reference to the technology’s use for less serious crime, asking if it has the potential to overwhelm the police and at what point would they choose not to use it.* The presenter argues that criticisms over the lack of action in response to shop-theft demonstrates the need for the technology in less serious crime and that there is an ethical responsibility to use any advantage to solve all crimes.
* The presenter continues that cost would likely be the first obstacle to increasing the technology’s use rather than overwhelming the police.

A committee member reemphasises the concerns around the implications this technology has on a surveillance state, arguing with DVLA and passport records, his technology could be used to identify anyone. As such, they suggest it should be reserved for only the most serious of crimes.* The presenter notes the concerns and suggests they will be taken under consideration.

A committee member asks if there will be an evaluation in the next 12 months, investigating the prevalence of false negatives and the level of transparency from the provider about how the technology works.* The presenter suggests this is possible as the supplier appears keen to work together with WMP and open to collaboration and receptive to recommendations.

A committee member seeks to reaffirm that the committee should remain focussed on single use cases so as not to set a precedent for future decisions, referencing that this proposal is for serious crimes so advice should not be transferred to other crime types. Less serious crimes should be addressed separately  | **Matt Welsted** |
| **5** | **11:45** | **Break** |  |
| **6** | **11:55** | **Stalking Algorithm** Previous recommendations:* The committee expressed concerns about the accuracy of the tool and the possibility of high numbers of false negatives/positives.
* The Committee was unclear about the function of the parallel victim ‘harm’ model and how it would be implemented alongside the offender model.
* The Committee requested more information before being able to advise.

The Chair thanks the presenters for the comprehensive paper, commenting that the grounding in literature and research was particularly impressive.The presenter explains how the existing challenge in policing is that most female murder victims are killed by their partner. Behind each murder is often a long and detailed history of domestic abuse incidents leading up to it. The presenter therefore explains that this model aims to estimate a probability that certain individuals go on to commit or be a victim of high harm crimes given the fact they have been a victim or perpetrator of stalking or harassment. The MARAC intervention process at the moment asks a number of questions associated with risk of domestic harm but there is huge demand and is largely unscientific process for assessing risk and potential for future harm.The presenter explains that since the model was last presented to the committee the specificity of the technology has been recorded as .95 and the precision as .54.The presenter suggests the literature review has, additionally, helped identify keyword indicator variables which could be used alongside a large language model in an attempt to gain more information from free text. However, the presenter reveals this effort was later abandoned as the increase in accuracy was not worth the exacerbation of labour costs.**Questions and Responses**A committee member asks whether broader categories of data would make the LLM method more useful?* The presenter suggests that as information was taken from incident logs, a large breadth of information was already utilised.

A committee member asks if more data is made available would the model’s precision improve?* The presenter suggests it is unlikely due to the fact that within the tens of thousands of cases the algorithm looks at there will always be those outliers not following the pattern whilst others continually endure.

The member continues by asking how useful the current output of the model is given the improvements made to the positive outcomes?* The presenter explains that in its current form, the technology has the potential to be used as one of many decision-making factors for referrals and prioritisation into MARAC.

A committee member suggests that if the algorithm depends on continuous addition of data of potential victim, how will this continual updating be managed, and if there a potential for overwhelming services.* The presenter explains that this tool is to be used to prioritise and better target those most at risk rather than to identify and act on any level of potential risk.
* The model does rely on the continual flow of data but the presenter explains how absence of data is itself an indicator of behaviour because if no data is coming in, the individual is unlikely to escalate to high harm.

A committee member inquires as to whether the information gathered from the LLM is useful independently from the predictive model, potentially as context?* The presenter agrees it could be useful but suggests the cost is likely to outweigh the benefits

A committee member asks what the next steps are in terms of the use of this analysis?* The presenter explains the ambition is to produce a model with an associated dashboard which informs MARAC and other partnerships.

A committee member suggests that, in using a new system of prioritisation, resources will have to be reallocated. Subsequently, they ask if there has there been an evaluation of the process by which this will be conducted?* The presenter agrees that some real-world testing needs to be done and engaged with but suggests the police will work alongside NPCC advice.
 | **Davin Parrott****Tom Joyce** |
| **7** | **12:25** | **BRAID Report**The presenter explains the report was designed to look at the Ethics committee and assess its impact on technology whilst identifying challenges it faces.The report concluded that the committee helped identify potential challenges to project operationalism and allowed the lab to anticipate concerns. The report also had a number of concerns and recommendations, however:* The report understood there was a strong human rights focus amongst members but was concerned by the need to broaden the particular rights focussed on by the membership.
* The report suggested that the contents of the committee was not accessible for the community to engage with, particularly through the minutes.
* There was a suggestion that projects needed to return to the committee more often after being implemented to grant a better understanding of how advice has shaped practice.
* More community representation within the committee was suggested.
* The report expressed fears that the committee may infringe on the operational independence of the police rather than staying focussed on ethical concerns.
* It was suggested that with the focus of the committee expanding, the name should be changed to reflect this.
* The report also suggested the creation of performance metrics and increased transparency over how advice is implemented.
* Similarly, the publication of an annual report outlining and summarising the work done by the committee along with a brief commentary on the issues faced was recommended.

**Questions and Responses**A committee member questions the extent to which the report goes beyond the responsible use of AI and instead assesses the committee more completely?* The presenter accepts the title is misleading and the report acted as more of a preliminary study of the committee and how it reviews responsible AI.
* The presenter continues by emphasising the focus was on the committee more than AI itself.

A committee member suggests that the annual report could go further and be used to suggest areas for the PCC to set strategic direction or hold to account, perhaps being distributed to the Accountability and Governance Board.* The presenter agrees that the report will be vital.

The member continues that there should also be included, a possibility for escalation routes for the chair to raise issues with the commissioner/chief/internal audit around the implementation of advice.A committee member expresses the view that a lot of the recommendations can feed into the new drafting of the TORs which could then be used as a document through which members can hold one another to account. The member continues by suggesting that the recommendations must be balanced against practicality, however. | **Claire Paterson-Young** |
| **8** | **12:50** | **Any Other Business**A committee member expresses concerns about WMP still using twitter despite the unethical actions from Elon Musk and controversy surrounding the right-wing extremism and rising misinformation on the website. The member asks if WMP intends to stop using the site in the near future.* Another member explains WMP will still use it as it remains useful for distributing warnings and information due to the public access it has that hasn’t been replicated elsewhere. The member does, however, explain that the comments have been turned off for all posts in an attempt to alleviate some of the issues.
 | **Marion Oswald** |
| **9** | **12:55** | **Committee Comments****Andi-Esra**The Committee suggest that there is a concern around transparency as to whether callers are aware they are talking to a robot, with further suggestions that the lack of human contact could negatively impact satisfaction amongst callers and the public.Multiple members note that further evaluation of the technology is needed, that data concerning only the number or proportion of calls being processed and the speed in which they are, not being enough. Particularly, members expressed concerns about the lack of data or comparison between abandonment rates with and without Andi-Esra.One member raised concerns about the disaster recovery processes of the technology, suggesting it could be a likely target for malicious actors and have disastrous affects should it go down. A final concern raised by a member suggested there could be an issue with the technology in regards to linguistics, it potentially being difficult for the technology to understand inputs from those with thick accents. Alternatively, it was suggested that language barriers could create additional difficulties communicating with the technology. **Recommendations:*** The Committee recommend that there is an acknowledgement and transparent indication that the caller is not talking to a real person at the call’s inception.
* The Committee raise that further evaluation is needed and suggest that abandonment rates of the technology should be investigated and compared to the current approach.

**Outcome B – Proceed with minor amendments****Video Analytics:**The most salient concern amongst the Committee regarded the application of the object recognition technology beyond investigating serious crimes. The implications it could have on the right to protest and ideas of a surveillance state further decreasing public confidence were both explicitly raised. The committee suggest the ethical implications of the technology become less acceptable when applied to less serious crimes.One member raised the possibility for the technology to produce false negatives as a possible issue whilst another agreed and added false positives to the claim also.One member suggested the transferring and disclosure of data to and from the provider of the service could cause issues with data protection and further contribute to claims of a surveillance state, especially if it could be applied to facial recognition. **Recommendations:*** The Committee recommends that the technology should be limited to combatting serious crime.
* As a result, members also recommend that the definition of serious crime must be made clearer and more obvious so as to reduce the possibility for abuse of the technology.

**Outcome B – Proceed with minor amendments****Stalking Algorithm:**One member suggested there was concern as to how the algorithm may influence and inform the decision-making process.Multiple members expressed concerns about how the algorithm is dependent on up to date data and enough victims in order to remain continually operational. Inquiries into the management system of this data and victims were made as a result. Overall the Committee suggested it was too early to judge the algorithm and suggested instead, that the evaluation process and interface needs to be completed before this occurs. Consequently, however, they did express concerns over how uncertainly in this process will be communicated.**Recommendations:*** The Committee expressed a belief that the victimisation of child sexual exploitation should be included in the algorithm.
* The Committee request that the proposal be brought back to the Committee after a clear and structured evaluation process once the algorithm is sufficiently tested and enough feedback has been gathered.

**Outcome E – More information required from the Lab to be able to advise** | **Marion Oswald** |
| **10** | **13:00** | **Next Meeting: March 2025** |  |