![west midlands police and crime commissioner]

ETHICS COMMITTEE- Formal Group Meeting
Wednesday 13th December 2023
10:00-13:00
Meeting held virtually via Zoom

Present:

Marion Oswald (MO)	Chair of Ethics Committee
Thomas McNeil (TM)	Assistant Police & Crime Commissioner
Jamie Grace (JG)	Ethics Committee
Sian Thomas (ST)	Ethics Committee
Tom Sorrell	Ethics Committee
Claire Paterson-Young (CPY)	Ethics Committee
Peter Fussey (PF)	Ethics Committee
Jennifer House-Go (JH)	Ethics Committee
Malcolm Fowler (MF)	Ethics Committee
Simon Rogerson (SR)	Ethics Committee
Jack Tracey (JT)	Criminal Justice Policy Lead (OPCC)
Eliza Ogden Barnsley (EOB)	Secretariat (OPCC)
Davin Parrott (DP)	Data Analytics Lab (WMP)
Steve Clark (SC)	Data Scientist (WMP)
Sam Todd (ST)	Value & Business Architect (WMP)
Marc Williams (MW)	Head of Architecture for IT & Digital (WMP)
Miri Zilka (MZ)	University of Cambridge

Apologies:

Anindya Banerjee	Ethics Committee
Kerry Reidy	Ethics Committee

1	10:00	**Welcome** The Chair opens the meeting and welcomes members.	**Marion Oswald**
2	10:05	**Update on Actions** Jack Tracey updated the committee on the previous meeting, noting that the stalking risk tool had returned to the committee for this meeting as requested and highlighted key issues arising from the interim meeting with WMP re: ongoing FR projects.	**Jack Tracey**
3	10:15	**OPCC update re: WMP HMICFRS report findings** APCC McNeil updated the committee on HMICFRS decision to place WMP into 'Engage' status. He highlighted that: - The OPCC had repeatedly requested special grants from the Home Office to support actions which would have mitigated some of the issues that were later highlighted by the report	**Tom McNeil**

		(including sexual offences and VAWG) but these were not received - Whilst the PCC takes the recommendations very seriously, he believes that the assessment is incorrect given recent progress in the areas identified as causes of concern (such as reducing backlogs etc.) - APCC McNeil also notes that limited resources and lower funding compared to similar force areas has meant that WMP is increasingly stretched APCC McNeil also updated the committee on the likely transference of police governance powers to the West Midlands Combined Authority following the elections in May. - He explained that if this goes ahead it would mean that the Mayor will hold all the powers that the PCC currently holds. - Notes the PCC feels due process has not taken place and therefore potentially taking legal action - APCC reassures the committee that he has had conversations with key stakeholders to highlight the success of this committee to ensure it has the best chance of continuing after May - He also highlights that it could present new opportunities for partnership work	
4	10:30	**101 Voice Assistant "Amy" (Alexa Technology) to service 101 requests** The Chair praises the report submitted to the committee which highlights the ethical considerations of the project and how these can be mitigated. The presenter explains some key background context: - That 'Amy' is a 101-voice assistant that answers phone calls (like Alexa) - WMP used to have high abandonment rates and waiting times. Whilst these have improved recently, this is mostly due to more responders being employed - It has been shown to the force executive who agreed to proceed with a 2-month proof of concept pilot with live 101 calls where 10% of calls will be sent to 'Amy' The presenter identifies the key aims of the project: - Improve resilience around peak and fluctuating call volumes - Release call handling capacity back to deal with more nuanced calls - Improve CDI compliance and customer satisfaction	**Marc Williams**

- In the longer term, help to potentially establish a multi-language service
- Want to ensure it can be used by all forces not just WMP
- Understand and categorise WMP's non-emergency request demand

The presenter explains that there are three use cases that they hope 'Amy' will help to achieve:
1. Request Fulfilment (e.g. needing to speak to a certain department or officer, lost and found property, parking issues and people in custody).
2. Request a case update (Amy completes a crime reference and date of birth check). This then allows the caller to leave a voicemail which she then transcribes and emails, but Amy does not provide this update due to data security issues
3. Prioritise vulnerable callers (currently no system for 101 to prioritise callers). This would be achieved by using key words that have been suggested by force contact.

In all contexts, if the caller is not understood the first-time they will be transferred to a real call handler and it has been ensured that flows are specific and simple questions.

The presenter explains the technology behind 'Amy'.
- It is built on LEX B2 technology from AWS. This aligns with WMP and national standards
- Calls are recorded. Any data is always in WMP accounts and under WMP ownership
- Native cloud technology is used through Amazon Connect
- Can speak 27 languages (not yet switched on)

Questions and responses:
A committee member commends that this is a proof of concept and asks how it can be guaranteed, given the fact that cloud technology is used, that the data does not go outside the UK.
- The presenter explains that AWS defines geographical 'regions' where data cannot pass
- States that Amazon, from a legal perspective, is very clear that the police data cannot leave this region
- The presenter also explains that the police assured landing zone has a number of security measures and is therefore safe to hold 'official' police data

A committee member asked about whether there had been research into the demographics of 101 callers with regards to their digital literacy and other factors.

- The presenter explains that no work has been done to assess the demographic of 101 callers
- The presenter acknowledges that 'Amy' will not necessarily work for everyone

A number of committee members raised their concerns about the gendered name of the tool and presumed voice of 'Amy' and the ethical implications of this.
- The presenter explained that feedback on other AI systems has suggested that humanisation of such technologies is needed.

A committee member asked how 'Amy' is being trained and how potential ingrained bias could be mitigated against.
- The presenter states that in the process of creating 'Amy', careful consideration has been given to ensure that sensitive data is not shared.
- The presenter therefore reassured the committee that the 'opt out' clause ensures that no sensitive data is fed back into the training of 'Amy'

A committee member raised a concern about impersonation and denial of service technical errors that could have a large impact.
- The presenter acknowledges that impersonation is possible but notes that the request for the caller's crime reference number and date of birth provides a check whilst also maintaining usability.
- The presenter also reassures the committee that there is a further human check as the police officer has to personally phone and speak to the caller.
- The presenter also acknowledges that if 'Amy' is rolled out on a larger sample then the risk of errors would be considered in more detail.

A number of committee members raise concerns about how vulnerability is measured by 'Amy' and whether the use of key words could wrongly prioritise non-vulnerable calls or negatively impact certain demographics.
- The presenter explains that force contact has suggested what words are usually used by the most vulnerable callers
- The presenter acknowledges that an assessment of vulnerability is very subjective and reassures the committee that patterns of calls are not being used to assess vulnerability

	- The presenter explains that they have ensured that 'Amy' gives the caller examples of how to explain their reason for calling to try and ensure that vulnerability is not missed
	- The presenter acknowledges that vulnerability could be missed but states that they believe it will help a lot of callers too, especially given that there is currently no vulnerability prioritisation of 101 callers
	- The presenter reassures the committee that assessment will be done to see if defined vulnerable calls do use the words identified by force contact
	
	A member asks whether there will be any assessment into 'Amy's' feedback process and whether there will be other assessments into things such as the number of abandoned calls or an increase of 999 calls from the same number.
	- The presenter says that whilst there may be potential for a public feedback process in the future, there was insufficient funding for it to be used at this stage
	- The presenter acknowledges that it may not be appropriate, given the nature of calls, for there to be public feedback but reassures the committee that other analysis will take place
	- The presenter also states that they will seek further approval to assess whether dip-sample call-backs to assess public opinion might be possible
	
	Violent Crime Hotspot Policing RCT
	The presenter summarised the recommendations from the previous committee meeting which were as follows:
	1. Committee members expressed concern about the potential lack of statistical significance of the results presented and questioned the implications of the use of a 'cross-over' RCT (each area acting as its own control). It was noted that the method used had been determined by the Home Office based on methods recommended by Sherman et al.
	   - The presenter explained that because the data were collected from a cross-over design RCT, the analyses cannot be done as if they were derived from a parallel track method
	   - With regards to the potential lack of statistical significance, the presenter stated that the entire posterior of the estimate of interest should be considered.
	2. Committee members also raised concern about the exclusion of certain datasets, in particular those relating to the night-time economy (excluded during the hotspot identification
		Davin Parrott

stage, not from the analyses), and the short distance for the displacement check.
- The presenter explains that the exclusion of the night-time economy dataset is not an issue as night-time economy areas are patrolled regardless
- The presenter explains that the 50-meter displacement check was used as anything greater than 50 meters meant that the areas would merge, especially given the nature of such areas being very busy such as shopping centres etc
- The presenter explains that the hotspots have been reassessed, also now considering robbery, which was previously not included, and there has been no evidence that the hotspots have moved
3. The Committee recommended that prior to further use of the report results, that additional models/results are produced using a parallel track RCT method, and including relevant violence data previously excluded from the model, and increasing the length of the displacement check, for comparison to the existing report.
- The presenter explains that the data received under the cross-over method design cannot be analysed as if it was a parallel track RCT method
- The presenter explains that they are unable to re run any RCT due to limited resources

The presenter explains WMP are asking for findings and that these would be provided and communicated more widely within WMP with the caveat that that further research is being undertaken including using other forces data.

Questions and responses:
A committee member airs that, whilst acknowledging that they do not know how statistical significance is measured, they believe that the suggestion that statistical significance is just a matter of philosophical opinion does not properly address the committees concerns.
- The presenter explains that due to the overly dichotomous nature of the term 'statistical significance' the view taken by the lab is that an assessment should instead be based on resource/time effectiveness.

In relation to the previous question, a committee member acknowledges that whilst they agree with the presenter's comments regarding statistical significance, the committees concern is that

some of the statistics show only 50% accuracy which is concerning. They therefore suggest that further analysis is needed to understand the impact of this.

- The presenter explains that they had not responded to this specific nuance of the point due to the previous meetings recording being corrupted (which impacted the minutes/ advice provided).
- The presenter also explained that differing levels of effect size are present at different levels of confidence which overall lead to positive return on the activity.

The Chair highlights that it is important that the caveats, confidence levels and limitations around the data are very clear and communicated ahead of further analysis that is due to occur. It is also important that this is communicated in an accessible way in order for effective policy to come of this.

5	11:45	Break	
6	11:50	**Update: harmful stalking and harassment offenders – estimation of future risk tool** The presenter informs the committee that due to changes in categorisation of different offences by the Home Office this means the whole tool needs to be re built. The presenter tells the committee that there are unfortunately not the resources to do this currently. It is further confirmed that the new categorisation means that if the stalking or harassment is committed in conjunction with another offence it is likely categorised only as the higher order offence. Further analysis therefore needs to be done to assess the implications of this. The lab agrees to come back to the committee with any further information that they learn about these developments. **Questions and responses:** APCC McNeil asks whether the recent HMIC review of WMPs management of sex offenders has any relevance to this area. He also asks whether WMP, subject to all ethical queries, have all the resources they need to help implement this project. - The presenter confirms that this would need to be looked into in more detail - The presenter confirms that subject to vetting, they do have the relevant resources	**Davin Parrott**
7	12:00	**External presentation: predictive tools research** The presenter explains that the research analyses risk assessment instruments, these are tools that assess whether someone is likely to	**Miri Zilka**

reoffend based on their current circumstances and characteristics, and how this was impacted by differential arrest rates.

- The presenter explains how these tools are often used for bail and sentencing and in the UK, by probation officers, for pre-sentence reports
- The presenter explains that the issue with these tools is that they can only work to actually predict re-arrest rather than re-offending. This is a concern as there is not a definitive link between arrest and offending generally and that the link between arrest and offending varies depending on characteristics
- The presenter explains that this can cause issues as two people with almost identical risks of re-offending could be treated very differently as their probability for re-arrest, which is falsely used as probability of re-offending, is predicted to be higher or lower.

The presenter summarises the methodology and the key findings:

- There is up to a 10% higher probability of a black person being identified as high-risk despite having the same estimated re-offending risk as the white person they are compared against.
- This can also be seen when other characteristics are compared.

Questions and responses:

A member asks what implications could there be for police use of data.

- The presenter explains that it can be used quite differently by courts and police forces
- The presenter argues that the way these tools are used in courts are most vulnerable to these differential assessments.
- The presenter explains that the data these tools are based on is inherently imperfect and therefore should not be seen as a definitive predictor of re-offending, at most it should be seen as an estimate of re-arrest.

A member asks whether the presenter feels these tools would be beneficial for more narrow crime types such as stalking or harassment.

- The presenter notes that, generally, predictions of violence offending can be more reliable.
- However, they note that there should not be an over reliance on these tools as violent crimes can often be committed spontaneously

		- The presenter also notes that it is important to consider how a person's tendency to report a crime can vary greatly amongst different communities - The presenter believes that thorough analysis needs to be completed to assess the actual accuracy of these tools - The presenter also warns that officers need to be given clear advice on what actions they could can take if they are given this information, otherwise the tools are not helpful A committee member asks what practical uses could come about from this research. - The presenter suggests that the usefulness of these tools should be assessed on a case-by-case basis - The presenter highlights that these tools are often used far more by probation officers than they are judges, likely due to the time sensitive nature of probation officers work compared to judges - The presenter also warns that these algorithms can often appear to be objective when they are often very subjective, therefore the context they are used in is key	
8	12:30	**Committee discussion and recommendations** **General Recommendations:** - In light of the recent change to the definitions of stalking offences, the Committee recommends that the Lab implements a process to review the implications of any offence definition or other data changes and requests that they notify the committee when there are changes that impact existing projects. **Harmful stalking and harassment offenders – estimation of future risk tool (Outcome E- requests more information from the Lab to be able to advise):** - The Committee acknowledged that the Lab are waiting to hear from the Home Office about offence definition changes that will likely impact the project. They also acknowledged that the Lab need the time and resource to rebuild the model to take account of the changes. - The Committee request that the project returns to the Committee following the update from the Home Office. **Violent Crime Hotspot Policing RCT (Outcome C- proceed with major amendments):**	

- Bearing in mind the potential limitations/caveats in the research method as noted in previous meetings, the Committee recommends that any onward communication of results be subject to a clear statement of such limitations/caveats, including in relation to the method adopted and the exclusion of NTE data, which the Committee recommends should be approved by a senior member of WMP with the relevant expertise who it is recommended should confirm in writing that they have understood the aims, limitations, and benefits of the project.

101 Voice Assistant "Amy" (Alexa Technology) to service 101 requests (Outcome C- proceed with major amendments):
- The Committee commended the thought that had gone into the project and the level of understanding the presenter showed as well as their willingness to properly consider ethical considerations and the questions proposed.
- The Committee recommends that further analysis is done into the demographic data of callers ahead of any further roll-out of the pilot. This should consider factors such as digital literacy levels, those who are at risk of harm and more general factors. This will also help to assess what languages need to be available for future roll-out.
- The Committee recommends that there is a broader consideration of the key words that would categorise a call as vulnerable (with a broader cohort outside force contact including officers/staff with expertise in stalking and domestic violence offences) to understand and mitigate against improper or incorrect categorisation negatively impacting certain demographics. The Committee, however, do commend the planned efforts to do more analysis around the key words and the desire to introduce a prioritising system that is not currently available.
- The Committee emphasised this is a first step and recommend that the emerging degrees of risk associated with roll-out are recognised, reviewed, and assessed. Particularly relevant here is the potential issue of a denial of service attack.
- The Committee recommends a public feedback process as well as analysis into the number of abandoned calls and potential increase in 999 calls from the same number.
- The Committee would like to see legal advice and DPIA ahead of any future roll-out.

		- The Committee also recommend a further assessment into the gendered name and presumed voice of 'Amy' and the impact this would have. - The Committee suggests the project reach out to Amazon to hear their comment on these issues and request further analysis on potential issues such as regional accent recognition and bias testing.	
9	12:50	**Close**	