Ethical review to support Responsible Artificial Intelligence (AI) in policing:

A preliminary study of West Midlands Police's specialist data ethics review committee

Research Report

Marion Oswald,* Claire Paterson-Young, Pauline McBride, Michael Maher, Muffy Calder, Gitanjali Gill, Elizabeth Tiarks and William Noble

11 September 2024

The project is part of the BRAID (Bridging Responsible AI Divides) Programme and funded by the Arts and Humanities Research Council (AHRC), part of UK Research & Innovation (UKRI)

***Corresponding author and project lead** marion.oswald@northumbria.ac.uk

Arts and Humanities Research Council

BRAID
Bridging Responsible AI Divides

Acknowledgements

The research team would like to express their gratitude to Simon Foster, the West Midlands Police and Crime Commissioner, Chief Constable Craig Guildford, Tom McNeil, the former Assistant Police and Crime Commissioner for the West Midlands, Jonathan Jardine, the Chief Executive of the West Midlands Police and Crime Commissioner, and the wider Police and Crime Commissioner and West Midlands Police teams for their support of this research. The research team would like to thank all the interviewees from the Data Ethics Committee, Data Analytics Lab, the police and the community, who generously gave their time to be a part of the research. Finally, the team would like to acknowledge valuable support of BRAID and the AHRC for making this important research possible.

Table of Contents

Executive Summary

The West Midlands Police and Crime Commissioner (WMOPCC) and the West Midlands Police (WMP) have for the past five years maintained a Data Ethics Committee to advise on the design, development and deployment of advanced data analytics and AI capabilities. This Committee comprises people drawn from backgrounds in academia, industry, public/third sector and policing. Since 2019, it has met at least on a quarterly basis, advising and making recommendations on each occasion on several projects and proposed tools, from in-principle analysis to tools ready for operational use. Its papers and minutes are published via WMOPCC. This interdisciplinary research used mixed-methods (including 26 interviews) to review the **impact and influence** of the Committee, and to recommend to national bodies, other forces and to WMP/WMOPCC factors that affect how best to go about using independent advisors in this context. Lessons from the Committee's experience, together with a single structured framework could inform a **coherent and consistent national approach**. The Conclusions and Recommendations (for **national strategy, police, Committee members, community representatives, academia and research funding bodies**) fall into the following themes:
- A Data Ethics Committee with diverse independent voices can **contribute positively** to the validity and responsibility of policing AI, thus **supporting operational policing**. It can develop **understanding** within the police of key ethical, scientific, legal and operational issues for planning and implementation. This will be successful only if the Committee has a clear function, is **fully incorporated** into the system of oversight and scrutiny, visibly championed by the Chief Constable & PCC, and suitably supported by a secretariat, robust process and communications;
- This will be successful only if membership includes genuine representation from the **community** that the police serves, there is transparent engagement, and time taken to allow members to understand the **technical and legal** aspects of the work.
- This will be successful only if the **operational context** is explained by operational police officers, and time taken to understand **how AI outputs will be used**, so as to enable potential benefits, risks/harms and proportionality to be assessed in the same conversation. Attention must be paid to police responsibilities for **public safety** (and how AI may support these responsibilities) as well as to risks related to **privacy, fair trial and freedom of expression**.
- Police forces, PCCs and national bodies embarking on such an approach will need to be prepared for **ambiguity**. There are often **no 'black and white' answers** to ethical, legal or technical questions raised by policing AI, such as reconciling privacy and security priorities relevant to the assessment of the proportionality of using suspect data.

Glossary

Term	Definition
Accuracy	A measure of the performance of a machine learning model as a classifier, defined as the total number of correct classifications (both true positives and true negatives) divided by the total number of classifications output by the system
AI	Artificial Intelligence: our working definition includes advanced data analytics and machine learning
BFEG	Biometrics and Forensics Ethics Group (BFEG)
The Committee	West Midlands PCC and West Midlands Police Data Ethics Committee
DAL	West Midlands Police Data Analytics Lab
NDAS	National Data Analytics Solution
Precision	A measure of the extent to which positive predictions output by a machine learning model (as classifier) are correct, defined as the number of true positives divided by the number of true positives and false positives.
Sensitivity	A measure of how many inputs belonging to the positive class (actual positives) are correctly identified as such by a machine learning model (as classifier), defined as the number of true positives output by the system divided by the number of true positives and false negatives (together, the actual positives). If sensitivity is high, then the system found most of the actual positives
Specificity	A measure of how many inputs belonging to the negative class (actual negatives) are correctly identified as such by a machine learning model (as classifier), defined as the number of true negatives output by the system divided by the number of false positives and true negatives (together, the actual negatives). If specificity is high, then the system found most of the actual negatives
SRO	Senior Responsible Officer
WMP	West Midlands Police
WMOPCC	West Midlands Office of the Police and Crime Commissioner
WMVRP	West Midlands Violence Reduction Partnership

1. Introduction

The deployment of AI by the police, while promising more effective use of data for the prevention and detection of crime, brings with it considerable threats of disproportionality and interference with fundamental rights. The West Midlands Office of the Police and Crime Commissioner (WMOPCC) and West Midlands Police (WMP) data ethics committee aims to **bridge the gap between ethical reflection, scientific rigour, and a focus on human rights, thus contributing to responsible AI in policing**. Democratic legitimacy and public trust around West Midlands Police's use of AI is partly dependent on the ethics governance in place and the public assurances that are made. This six-month project brings together a diverse team of researchers in Law, Computer Science, Social Innovation, and Policing, with extensive experience of theory and practice of real-world ethical approaches to data analytics and AI in sensitive contexts. The partnership with WMOPCC presents a unique opportunity to analyse the operationalisation of AI tools in policing, and the impacts of advice from its data ethics committee. A specific focus is on the impact of the deployment of policing AI on human rights of vulnerable groups.

Research context

Established in 2018, the West Midlands Police Data Analytics Lab (DAL), the force's internal data science function, has developed a range of projects involving the use of advanced analytics. This report summarises these projects by the use of the term 'AI' which in this context, relates to models that make use of advanced data analytics and statistical techniques and some use of machine learning. AI has the potential to enable the effective allocation of resources, or the adoption of new crime prevention models. Examples of projects that have been developed include a system for forecasting the risk of re-offending, and data analysis tools for investigation of domestic abuse. Ethical questions frequently arise concerning the implications of the use of AI (House of Lords Justice and Home Affairs Committee, 2022). Without ethical scrutiny, there is a risk of new technologies contributing to injustice or exacerbating issues, such as institutional discrimination (Borgesius, 2020). The use of such technologies in law enforcement offers the prospect of enhanced insights from crime data but also creates risks for individual rights and of disproportionate impacts on vulnerable members of society.

West Midlands Data Ethics Committee

To help to address such issues, the Data Ethics Committee (the Committee)[1] was established in 2019 by the WMOPCC, championed by the former Strategic Advisor and later Assistant Police and Crime Commissioner for the West Midlands, Tom McNeil, and both Police and Crime Commissioners (PCC) elected since 2018. The Committee seeks to provide practical, independent advice and recommendations to the DAL, regarding the ethical risks that may arise from their projects. This specialist committee was the first of its kind in UK policing and presents an ongoing experiment in scrutinising and advising policing projects that involve data and technology. The Committee's role is to act as a 'critical friend'[2], aiming to put rights at the heart of the technical programme, whilst increasing the transparency of data science and highlighting legal, ethical, and societal challenges.

The Committee's role is purely advisory; the Chief Constable and/or the PCC remain responsible for decisions as to whether the project proceeds. The Committee's secretariat is provided by the PCC which means the Committee is therefore effectively part of the PCC's strategic oversight function, although it has no independent existence, budget or powers.

The Committee is comprised of WMP and PCC representatives, non-WMP policing representatives, lay members and professional academics from a wide range of disciplines such as Law, Computer Science, Safeguarding, Ethics and Philosophy. The Committee is attended by members of the DAL, and a Senior Responsible Officer (SRO) from operational policing, when required, to link the data science projects to the real-world policing demands and strategy. The DAL and SRO provide briefing papers and presentations to the Committee, which are then discussed at the scheduled meetings. The Committee has in the past considered projects led by other organisations or groups. Such examples include the National Data Analytics Solution (NDAS) project (a national project funded by the Home Office), and organisations such as the West Midlands Violence Reduction Partnership (WMVRP).

---

[1] https://www.westmidlands-pcc.gov.uk/governance/ethics-committee/.

[2] A 'critical friend' is used to denote an advisory relationship between two parties that allows the advisor to look both 'inwards' (through comment and advice) and 'outwards' (through responsibility and legal accountability) (Murji, 2011).

The Committee meets every three months and is chaired by a Committee member, Professor Marion Oswald from Northumbria University, who was appointed by the WMOPCC and has remained in this position since the Committee's inception. The agenda and briefing papers are sent out prior to the meeting and accompanied by the minutes from the last meeting and a data tracker. The data tracker is updated to keep track of the projects that have been presented to the Committee, including the questions and recommendations made regarding the projects' design and potential impact. Whilst these meetings are private, most documents are available on the WMOPCC website[3] and can be viewed by the public, reflecting the Committee's commitment to transparency.[4]

Terms of Reference have been developed to lay out the roles and responsibilities of the Committee, appointments of members, principles, and key questions that the committee will put to the DAL. The principles, set out below, are aligned with those of the National Statistician's Data Ethics Advisory Committee[5]:

- The use of data has clear benefits for users and serves the public good.
- The proposed use of data should not discriminate against any individuals or groups of people (e.g.., based on race, nationality, religion, gender, sex, disability, age, social background or otherwise).
- The proposed use of data is necessary and proportionate.
- The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.
- The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity, quality and human rights.
- Data used and methods employed are consistent with legal requirements such as data protection laws, equality laws, the Human Rights Act, public law and rights of judicial review and the common law duty of confidence.

[3] https://www.westmidlands-pcc.gov.uk/governance/ethics-committee/ethics-committee-reports-and-minutes/.

[4] Where topics or the discussion is considered sensitive, these are excluded from the public documents.

[5] https://uksa.statisticsauthority.gov.uk/the-authority-board/committees/national-statisticians-advisory-committees-and-panels/national-statisticians-data-ethics-advisory-committee/.

- The data used and research and analysis methods employed are sufficiently accurate, appropriate and rigorous in order to draw reasonable conclusions.
- The views of the public are considered in light of the data used and the perceived benefits of the research.
- The access, use and sharing of data is transparent, and is communicated clearly and accessibly to the public.
- The use of data will never be used to harm, cause emotional or physical danger, stigmatise and/or victimise a particular individual or group of individuals.

The West Midlands Police Data Analytics Lab

The DAL is led by data scientists and its role is to develop programmes of work that use data more intelligently to contribute to the strategic and operational goals of the WMP, for example, the prevention of crime or the effective allocation of resources. The projects developed cover a wide range of remits, such as the identification of crime hotspots, identifying individuals involved in criminal networks, and models that predict or estimate future offending, in order to inform police actions and interventions.

The DAL, together with the relevant SRO and other operational officers, presents projects to the Committee and receives comments, questions, advice, and recommendations through a free-form discussion. After this discussion, the Committee will make one of six recommendations for the project's development and potential operationalisation as specified in the Terms of Reference:[6]

- It advises approving the project.
- It advises approving the project with minor amendments.
- It advises approving the project with major amendments.
- It advises rejecting the project.
- It requests more information from the DAL to be able to advise.
- So far as it relates to analysing the outcomes of an actioned policy or strategy arising from the DAL's research, it advises on a continued or changed course of direction (e.g. if

[6] https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2019/07/Ethics-Committee-Terms-of-Reference-as-at-1-April-2019.pdf?x32340.

serious problems have been identified necessitating stopping the intervention, or encouraging further implementation of the intervention based on evidence of positive outcomes).

After each meeting, the DAL will consider the relevance of the Committee's advice and recommendations for the project's design. The DAL will return to a future meeting with an update on the project or answer questions that were unable to be answered in previous meetings. Some projects developed by the DAL can be 'in-principle' and will not necessarily be operationalised if limitations in their development cannot be overcome. Therefore, the Committee may often offer advice and recommendations for 'in-principle' models to aid their development.

Within policing, operational decisions are ultimately decided by the Chief Constable, and, therefore, the responsibility for decision-making about use of the DAL's projects is independent from the Committee and the WMOPCC.

Research methodology

Aims and research questions

This project investigates the influence of the Committee upon technology design, identification of human rights concerns and the incorporation of the interests of vulnerable groups. The research considers the potential of other frameworks to improve the process, and the challenges that could shape future research. The project does not address ethics in the abstract but is grounded in the real challenges of real applications of AI tools in policing, focusing not only on outcomes but also on processes which may generate trust or fairness by exercising and displaying good governance.

The key research questions informing the project are:

- RQ1 *Influence on technology*: What influence has the Committee had on the design and operationalisation of WMP AI policing projects?
- RQ2 *Human rights issues*: What human rights related issues were identified by the Committee and how were these issues dealt with in the design and operationalisation of AI tools?
- RQ3 *Vulnerable groups and data*: How, if at all, are the interests, views and concerns of vulnerable groups incorporated within the ethical review process?

- RQ4 *Challenges of ethical review*: What issues and challenges have Committee members and police representatives encountered in the committee review process?
- RQ5 *Potential of other models to improve the committee process*: In what ways could the use of the factor's framework (Janjeva, Calder and Oswald 2023) and matrix evaluation model (Oswald, Chambers and Paul 2023) improve the development of Responsible AI in policing?
- RQ6 *Research challenges*: What challenges emerge from the research which would need to be addressed in larger research projects investigating embedded ethics processes?

Methods

This section offers an overview of the mixed-methods approach that draws on an intersectional lens to assist in the critique and interrogation of the AI design, data collection, tools, and techniques (see however, Challenges to the intersectional research approach). Figure 1 at the end of this section summarises the research approach.

Interviews

Semi-structured interviews balance the flexibility of unstructured interviews with the consistency offered by structured interviews. It allows for a set of predetermined questions to guide the interview but also promotes natural conversation and the exploration of topics that arise spontaneously during the interview. This flexibility was key to the research as it enabled interviewers to probe deeper based on respondents' answers, facilitating a richer understanding of the subject (Lincoln and Guba, 1985). This balance allowed for a comprehensive exploration of the research topic while maintaining some consistency across interviewers. It also enabled interviewers to obtain detailed and nuanced insights into participants' experiences and perceptions of ethics, AI and policing that created space for the discovery of unexpected information (Lincoln and Guba, 1985; Noaks and Wincup, 2004). The conversational nature of semi-structured interviews also contributed to participant comfort, especially with sensitive topics (Noaks and Wincup, 2004).

Participants were identified and recruited with support of the WMOPCC and the research team's professional networks. Participants were approached by a member of the research team to identify whether they would be interested in taking part in the project. Consent was sought from potential participants from the research team's professional networks to share their contact details with other members of the research team. This was to minimise any potential conflicts of

interest where a member of the research team had a pre-existing relationship with a research participant.

This recruitment process involved an initial contact via email to introduce the research, provide a copy of the Participant Information Sheet (Appendix E) and Consent Form (Appendix F), and reinforce the participant's right to decline participation without justification or repercussions. If the participant indicated they would like to be interviewed, then further correspondence was sent to schedule an interview time and location for the interview.

Semi-structured interviews were conducted with 26 participants from three key stakeholder groups including the Committee (n=10), DAL and Police (n=6), Community representatives (n=10). Community representatives supported diverse groups in the following areas: Rights and Equality (Race, LGBTQ+, Disability); LGBTQ+; Religion and Faith; Mental Health; and Youth Justice. Interview schedules were created for stakeholder groups, aligned with the research questions, to promote in-depth discussion with stakeholders (Appendix A). Interviews were audio recorded and transcribed by an approved transcription service and/or a member of the research team. Interview participants volunteered to participate with anonymity and confidentiality assured; therefore, the results presented in this report are anonymous and any names that appear in quotations are coded. The research team labelled quotes from participants based on stakeholder group, namely: Committee representatives as 'Ethics Committee representative (number)', DAL and Police representative 'Data Analytics Lab and Police representative (number)', and Community representatives as 'Community representative (number)'.

Thematic Documentary Review of Committee Minutes and Documents

A thematic documentary review of existing published papers and minutes of the Committee and ongoing proceeding(s) during the life of the project (building on Oswald's 'observing participant' research (Oswald, 2022)) was conducted by the researchers. The thematic documentary review seeks to understand the Committee's recommendations on DAL proposals aligned with principles set out in the Committee Terms of Reference (West Midlands Police and Crime Commissioner, 2019, pg. 8-9 (section 48)).

Ethics

Ethical approval was obtained from the University of Northampton Faculty Research Ethics Committee and Northumbria University Research Ethics Committee. Ethical considerations for this research were associated with confidentiality and anonymity; voluntary participation; informed consent; data protection and storage; and the safeguarding of participants. This includes:

- Informed consent – Participants received detailed information on the project, with a Participant Information Sheet outlining the purpose, benefits, risks, and funding behind the study before they agreed or declined to join. Participants completed a consent form prior to participating in interviews.
- Anonymity and Confidentiality – Participants' anonymity was protected, and any identifiable information removed from the transcripts prior to analysis. Participant information was stored in accordance with Data Protection and GDPR.
- Data protection and storage – Information is stored in line with the Data Protection Act and GDPR. Data security was maintained by having personal information accessible only to the research team. Data is stored on the University of Northampton and Northumbria University servers. Subsequently, all data will be deleted and destroyed, based on JISC guidance on managing research records.
- Safeguarding of participants - Physical, social, psychological, and other types of harm were not expected to occur as part of this research project. However, a distress protocol was detailed for the research to ensure relevant support and contact information. In addition, details of the complaint procedure were available.

Data analysis

Documents

Analysis of publicly available minutes and technical documents from the Committee were grounded in the principles of thematic analysis. Themes were continuously reviewed and refined as documents were reviewed, with the aim of identifying key projects for review that covered a range of topics and significant themes that arose in the minutes.

Interviews

Interviews were transcribed and imported to Microsoft Excel (removing any identifiable information). Transcription was undertaken by an approved transcription service or a member of the research team. Interviews with Committee members, police, DAL, and Community representatives were analysed using Braun and Clarke's Reflexive Thematic Analysis (Braun and Clarke, 2022). Reflexive Thematic Analysis can be applied across a wide range of research questions and topics, but the underlying assumptions that guide a researcher – including epistemological and ontological assumptions – can also influence the analysis. Under the Reflexive Thematic Analysis approach, researcher reflexivity[7] is prioritised. This reflexivity extends beyond how a researcher is positioned in relation to their work and includes reflexivity about the assumptions and beliefs underpinning the researcher's approach to the topic, the participants, and the data (Braun and Clarke, 2019). Reflexive thematic analysis is fundamentally grounded in interpretivist traditions, whereby data are not presumed to reflect an objective reality that can be directly observed. Rather, data are comprised of themes that reflect a '*pattern* of shared meaning, organised around a core concept or idea, a central organising concept' (Braun et al., 2019, p. 844, emphasis in original).

Data analysis proceeded over multiple steps, or phases, including familiarisation and coding, theme development, review, and definition of themes, and finally, producing the analysis. Each phase was iterative and grounded in the data. Patterns were identified within and across interviews to enable a deeper understanding of the shared meanings and latent constructs that helped to explain the research question(s) under investigation, for example, how human rights were interpreted and built into ethical review. In the initial phases of data analysis, data was categorised into codes. Each code captured an important and specific segment of the data, reflecting the research question. At later stages of analysis, codes were refined and abstracted into themes that captured latent constructs within the data (Braun and Clarke, 2019; Braun et al., 2019).

[7] 'Reflexivity' denotes a process where the researcher examines their own thinking, emotions, perceptions, reactions, and motives in a circular manner to reduce bias in data analysis (Ide and Beddoe, 2023).

Technical observations

To complement the research, members of the research team carried out technical observations of AI tools to contextualise the work of the DAL and better understand how the Committee's advice might impact on the design of AI projects. The observations were fed into the reflective sessions to aid theme development and the contextualisation of the collected interview data.

The observations though useful were limited in scope. Six projects created by the DAL had been selected for analysis.[8] Dashboards had been created for three of these but of these, only two were 'live': the IOM model (in beta) and the Offender Checking tool. Researchers were unable to view the dashboard for the Offender Checking tool for reasons of confidentiality. However, it was possible to view the dashboard for another project, the Crime Seasonality planner.

A summary of the research design is set out overleaf.

[8] These were RFSDi/IOM, Youth Crime MSV, Mental Health related incident analysis, Stalking and Harassment, Community Tensions, Organisation offending checking tool for PSD.

Summary of research design

Stage	Definition	Title	Description
Review	Understanding and Scoping	Analysing Committee proceedings and relevant literature	Desk-based scoping
			Collation of Committee recommendations, technical and human rights issues
Observe	Exploring and comparing	Observation of technical design and comparison with Committee recommendations	Observation of technical design
			Compare with Committees' structure and recommendation
			Collation of Committee recommendations, technical and human rights issues
Understand	Analysing and creating	Stakeholder Perspective	Semi-structured interviews with stakeholders including members of the Committee, Data Analytics Lab, and West Midlands Police and Crime Commissioner
		Investigating stakeholder views and thematic analysis	Semi-structured interviews with Community representatives or under-represented groups
			Semi-structured interviews with stakeholders from Scrutiny Panels
			Reflective data review of semi-structured interviews
Communicate	Validating and disseminating	Final project report, and further academic communications	Report of findings for all research Action plan for future research

Figure 1. Summary of the adopted research approach

2. Abridged literature review

The following sections offer an abridged literature review designed to integrate and synthesise peer-reviewed literature that is relevant to the context of the research. It is split into four overarching topics: understanding AI and its role in policing; the ethics of AI in policing; intersectionality and vulnerability; and the implications of AI-supported policing for human rights.

Understanding AI and its role in policing

Although AI is now part of mainstream vocabulary, there is no consensus about the definition and scope of the term (Wang, 2019; Marko Grobelnik, Karine Perset, Stuart Russell, 2024). Stone et al. (2016) suggest that in the absence of a widely accepted definition '[p]ractitioners, researchers, and developers of AI are … guided by a rough sense of direction and an imperative to "get on with it"' (p.12). The result is a spread of approaches and sub-fields. According to Wang (2019) 'for the larger community of computer science and information technology' it is sufficient to define AI by reference to these approaches and sub-fields (p.7). For policy makers and regulators, on the other hand, a working definition is essential.

Guidance issued by the judiciary for the Courts and Tribunals of England and Wales defines AI as '[c]omputer systems able to perform tasks normally requiring human intelligence' (Court and Tribunals Judiciary, 2023). The definition of AI systems in the EU AI Act definition is based on the Organisation for Economic Co-operation and Development (OECD)'s definition and defines an AI system as a:

> 'machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.' (Regulation (EU) 2024/1689)

This research takes a similar approach, using 'AI' as a descriptor for a wide range of technologies which perform tasks normally associated with human intelligence. In a similar way to Oswald (2022), this research treats 'advanced data analytics, predictive policing, natural language processing, network analysis and facial recognition' as forms of 'intelligent' analytics falling within the scope of 'AI'.

AI in policing

The use of advanced data analytics and AI in policing is not new (Babuta, 2017). Various factors may account for the adoption of such methods. Santos (2017) traces increased interest in crime analytics to a shift in focus to problem-solving and the prevention of crime. Lack of resources may be a factor; Perry (2013) claims that predictive policing allows the police to do more with less. Increases in computational power and volumes of data make it more feasible to carry out sophisticated analytics at scale (Babuta and Oswald, 2021). Information overload may make it challenging to analyse volumes of data except by such methods (Babuta and Oswald, 2020).

Afzal and Panagiotopoulos (2024) provide a recent review of the use of data analytics in policing. Most of the approaches they identify have been adopted or explored in the UK. Thus, in a UK context advanced data analytics are used in:

- **exploratory or predictive crime mapping** – identifying or forecasting the location/timing of crime (Babuta and Oswald, 2020, 2021)
- **exploratory profiling** – carrying out harm scoring to assess the risk presented by offenders (Babuta and Oswald, 2020, 2021) including by means of network analysis (Taylor, 2023)
- **social media analysis** (Edwards and Urquhart, 2016; Babuta, 2017)
- **predictive profiling** – forecasting which persons are likely to re-offend, engage in high-harm crime, or become a victim of crime (Babuta and Oswald, 2020, 2021)
- **surveillance automation** including facial recognition (Veale et al., 2019; Radiya-Dixit and Neff, 2023) and automated number plate recognition (Babuta, 2017).

Certain initiatives by UK police authorities do not however fall neatly into the categories proposed by Afzal and Panagiotopoulos (2024). These include the use of techniques for content analysis and enrichment in the investigation of crime (e.g. analysing seized media for police investigations, synthesis of information gathered from multiple sensors)[9] and a pilot of an AI-enabled system to handle calls for service (Vallance, 2024).

[9] The Data Analytics Lab used Latent Dirichlet Allocation (LDA), a natural language processing technique, to identify topics or themes in crime reports. This analysis was carried out as part of a broader process of identifying and extracting data to be used as features or inputs to a predictive model.

Further high-profile examples of the use of AI in policing include:

- Predpol – a commercial tool which provides forecasts of geographic areas likely to experience higher level of crime (Ferguson, 2019).
- HART (Harm Assessment Risk Tool) – jointly developed by Durham Constabulary and the University of Cambridge, the tool employed random forest modelling techniques to forecast the risk of re-offending (Oswald et al., 2018).
- The Gangs Violence Matrix – an intelligence tool developed by the Metropolitan Police Service ('MPS') to identify and risk-assess gang members and identify those at risk of victimisation (Metropolitan Police, no date). After the Information Commissioner found that MPS contravened several principles of the Data Protection Act 1998 in its use of the tool and served an enforcement notice on the service, the tool was redesigned and the MPS announced its decommissioning in February 2024.[10]
- Generative AI – commercial provider Axon has launched a large language model designed to generate police reports based on audio from a body worn camera.[11]

Ethics review of AI in policing

The field of ethics and governance within policing is one in constant flux as it responds to institutional change, modernisation, and, occasionally, scandal (Snelling, Macvean, and Lewis, 2023). Accountability is critical within policing and requires a transparent process to ensure that the public understands where, and by whom, decisions are made and policy is developed (Wilson et al., 2023). Ethical panels play a key role in this process, ensuring that decisions align with societal values, legal obligations, and moral standards, whilst acting as an impartial 'critical friend' (Dixon, 2018). However, the relationship between ethical values and policy is complex. Introducing ethical considerations to policy development in policing may slow down processes when quick solutions are needed or attract scrutiny if impartial stakeholders find policies ethically indefensible (Snelling, Macvean, and Lewis, 2023). Moreover, ethical considerations are themselves intricate. The two major ethical theories, deontology and consequentialism, embody inherent tensions in the development of policing technology. Deontology's reliance on values-

[10] https://policeprofessional.com/news/mps-to-decommission-gangs-violence-matrix/.

[11] https://apnews.com/article/ai-writes-police-reports-axon-body-cameras-chatgpt-a24d1502b53faae4be0dac069243f418.

based choices can be challenging when values contradict each other, for example, privacy and security when training datasets for facial recognition. Alternatively, consequentialism may justify a minority suffering to ensure that a greater good is achieved or lead to decisions based on a miscalculation of a future outcome (Goldberg, 2022). As policing continues to incorporate AI technology in its strategy, ethical considerations are becoming both more pressing, and valuable due to the uncertainty of the outcomes from its implementation. There is consensus across the literature regarding the principles on which AI should be ethically governed:

- Transparency
- Reliability and Safety
- Justice and Non-Discrimination
- Accountability.

This is balanced against shared concerns:

- Vagueness
- A lack of technical transferability
- Unenforceable principles (Corrêa, 2023).

This has led to more explicit calls for an assurance that oversight is considered, along with factors relating to societal and environmental wellbeing (Ethical Guidelines for Trustworthy AI, 2019).

Ethical vagueness

Oversight is additionally required as AI ethics are generally limited to broad and unspecific principles (Morley, et al., 2023). This can lead to vagueness and ethics 'operat[ing] at a maximum distance from the practices it actually seeks to govern', leading to a gap between academic literature and high-minded ideas, and real-life use and the technological developments on the ground (Hagendorff, 2020; Morley et al., 2020; Munn, 2023). Whilst the gulf between the ethical and technical makes effective AI ethics difficult, so does the lack of binding principles (Corrêa, 2023). Poorly implemented ethical principles can act merely for PR purposes as opposed to governing the development of AI (Hagendorff, 2020). Therefore, whilst these 'soft governance' principles are a good first step in AI ethics (Morley et al., 2021), this 'unjustified laissez-fair attitude' (Morley, 2023, p.418) means that AI practitioners and policy makers are allowed to ineffectively self-govern, rendering compliance with ethics as a mere check-box exercise (Hagendorff, 2020).

The negative consequences of this can be further exacerbated by the fact that the AI industry and AI ethics constitutes a type of **'male-dominated justice ethics'** which favours certain principles over others (Hagendorff, 2020, p.103). Similarly, given that the AI and tech industry is an 'a-ethical' space that contains biases, harm can be exacerbated or created, and ethical concerns can be essentially ignored (Munn, 2023, p.871). There is therefore a need to ensure that ethics committees involved in AI and policing are given ample opportunity to develop specialist knowledge that straddles the **real-world consequences of AI and the positive or negative impacts it may have on effective policing**. Consideration should also be given to the experiences of ethics committees in other institutional disciplines and the challenges related to the standardisation of ethical approaches (compounded by the limited number of AI and policing ethics committees). Opportunities should be sought for opportunities for cross-collaborative learning and the development of specialised knowledge, and logistic considerations (Marcus et al., 2015).

AI ethics toolkits

AI ethics toolkits can assist practitioners, users and wider stakeholders to assure the ethical issues surrounding AI including the review of 'existing power dynamics' (Radiya-Dixit and Neff, 2023, p.1334). Whilst the format of these toolkits varies, their shared objective is to take a context-specific approach to hold AI developers accountable, ensuring users are well-informed and to make 'systems fairer, more robust and transparent' (Durmus, 2021).

There is a plethora of AI ethics toolkits, including **audits, information sheets and model/system cards**. Some toolkits are confined to specific areas or are aimed at specific groups. WEF's Empowering AI Leadership, for example, is aimed at board directors and AI Now Institute's 'Algorithmic Accountability Policy Toolkit' at legal and policy advocates (Durmus, 2021). While these toolkits have specific functions, some toolkits can be applied in a variety of contexts and by a variety of users. For example, Radiya-Dixit and Neff's (2023, p.1341-1342) 'socio-technical audit' was used in relation to police use of facial recognition technology. However, their 'score cards' can be customised for other contexts. In a similar vein, Brown, Davidovic and Hasan (2021) suggest a framework for ethical audit in which algorithms are scored on the basis of certain metrics, and the scores used to assess the impact of the algorithms on stakeholders.

Information sheets or model cards allow developers to record information about 'both the technical aspects of any given tool and the organizational process of its implementation' (Franks

and Cameron, 2023. For example, Gebru et al. (2021), propose the use of datasheets to document characteristics of datasets and Mitchell et al. (2019) the use of model cards to document characteristics of AI models.

AI ethics toolkits can improve transparency and prompt ethical reflection. However, there are weaknesses in these processes. In particular, they are typically targeted at stakeholders who are internal to technology companies (Wong et al., 2023) and geared to **internal record-keeping** rather than processes of external verification. They are best treated as **only one component** of a broader ethics assurance framework (Powell and Oswald, 2024).

The need for specific data ethical oversight in policing

The need for data ethics oversight in policing arises 'from the fact that some means of law enforcement and crime prevention involve information technology with great potential for intrusion and other kinds of harm (Sorrell, 2024, p.16).' The West Midlands oversight model has been said to embody much best practice for such oversight; the House of Lords Justice and Home Affairs Committee recommended both a national body to act as a resource for best practice and regional committees to scrutinise the use of new technologies by police forces through their lifecycle (House of Lords Justice and Home Affairs Committee, 2022). The **2023 AI Covenant for Policing** advises Chief Constables and Police and Crime Commissioners that they may consider using specialist committees for advanced data analytics and AI projects (National Police Chiefs Council, 2023). Such committees should 'wherever possible, work in public, be independently chaired, include experts in data ethics and medical ethics, and have community representation', with their remit including the review of 'accuracy, reliability, security, safety, performance, evidence-based decision capability and feedback ability of all AI.'

Only a small number of police forces – West Midlands, Essex, Thames Valley/Hampshire and Police Scotland to date – have established public or semi-public specialist committees or data ethics functions as far as the researchers are aware. In addition, the Biometrics and Forensics Ethics Group (BFEG), an advisory non-departmental public body sponsored by the Home Office, has recently expanded its remit to include large datasets and machine learning, and to provide an advisory resource for police forces in respect of complex data analytics and AI projects.

Sorell identifies that the **stakeholders** relevant to the Committee and BFEG differ – for West Midlands, the police itself, regional citizens, data scientists, and civil society groups are stakeholders; for BFEG, the Home Office, Ministers and regulators are currently the main

stakeholders. It appears yet to have been determined whether BFEG could – now or in the future - provide a national or regional independent oversight structure of the kind recommended by the House of Lords report. Furthermore, there is **no single agreed framework** used by these committees. Oswald (2022) comments that the West Midlands process in practice operationalises key factors that must be considered in the human rights necessity test, and advises that technical and statistical aspects of policing AI should not be isolated from the legal, contextual, operational and ethical considerations, as each will contribute to how technology is evaluated. Sorell (2024) argues that legal concepts of necessity and proportionality may have to be supplemented by thinking about the way in which severe harm may generate obligations to prevent it, where prevention may be assisted by AI models.

Intersectionality, marginalisation, and vulnerability

AI may reproduce or amplify human biases, entrench social inequalities and disproportionately impact on marginalised and vulnerable persons and communities (Ulnicane, 2024). Facial recognition technologies have been shown to work poorly for dark-skinned women (Buolamwini and Gebru, 2018). In the US, systems for algorithmic risk assessment have been found to disproportionately impact Hispanic communities (Hamilton, 2019). Hotspot policing can increase 'the entry of predominantly low-income, often minority ethnic men into the criminal justice system' (Longstaff et al., 2015). Intersectionality is emerging as an important framework to address concerns about such discriminatory effects (Ulnicane, 2024).

Intersectionality, as a critical framework, was developed by Kimberlé Crenshaw to expose 'how single-axis thinking undermines legal thinking, disciplinary knowledge production, and struggles for social justice' (Cho et al., 2013; Crenshaw, 1989; Crenshaw, 1991). Crenshaw argued that the tendency to treat race and gender as mutually exclusive categories contributes to the marginalisation of Black women in the United States. Intersectionality, as a recognition of the **multidimensionality** of Black women's experiences, is a prism for seeing the way in which various forms of inequality often **operate together and exacerbate** each other. Patricia Hill Collins describes intersectionality as a framework for understanding the experiences of marginalized individuals within a 'matrix of domination characterized by intersecting oppressions' (Collins, 2000; May, 2015; Crenshaw, 1991).

Intersectionality commands an important place in academic discourse as a diverse and generative critical framework examining 'multiple between-group differences, charting shifting

configurations of inequality along various dimensions' (Atewologun, 2018). There is consensus that intersectionality analyses positions and individual and relational experiences as 'reciprocally constructing phenomena' within race, class, gender, sexuality, ethnicity, and related structural systems of oppression that perpetuate social inequalities (Collins, 2015; McIntosh, 2012). Intersectionality 'makes clearer the arithmetic of the various forces – the offsetting, ameliorating, intensifying, accumulating, or deepening' impacts of power, dominance, oppression, and inequality at individual and social levels. As Collins states, utilising intersectionality as an 'analytical strategy' helps us to better understand gender, race, class, ethnicity, and similar categories in relational terms rather than in isolation and thereby address more effectively interpersonal and structural systems of oppression.

Intersectionality, human rights and responsible AI

Increasingly, the intersection between AI, vulnerability and human rights considerations with respect to policy position and legislative proposals is gaining traction that addresses the AI fairness infrastructure (Rodrigues, 2020). Expanding both the conceptualisation and operationalisation of intersectionality will create an evidence-based 'fresh dossier' that will reflect not only 'existing human prejudice and discriminatory behaviour, but also based on new grounds such as profiling identities based on a combination of behavioural and demographic characteristics' (European Union Agency for Fundamental Rights, 2022). Consequently, responsible AI should be designed, assessed and reframed through the intersectional vulnerability framework that take into consideration both the adversities people experience and factors likely to make them more vulnerable. Responsible AI, in this reframing, is **not narrowly focused on the AI technology pipeline** but involves 'centering marginalized people and valorizing critical knowledge production that makes room for their voices' (Ovalle et al., 2023).

Definitions of marginalisation and vulnerability

Marginalisation and vulnerability are complex, inter-connected and multifaceted concepts that do not have a universally accepted definition. **Marginalisation** is defined as the 'position of individuals, groups or populations outside of 'mainstream society … distanced from power and resources that enable self-determination in economic, political, and social settings' (Schiffer and Schatz, 2008, p.6) whereas the **spectrum of vulnerability** includes the 'inability (of a system or a unit) to withstand the effects of a hostile environment, such as ecosystem and human

society'.[12] Vulnerable population includes 'those individuals or groups who have a greater probability than the population of being harmed and experiencing an impaired quality of life because of social, environmental, health, or economic conditions or policies'.[13]

Vulnerability through different approaches situates individuals in different economic, social, cultural, and institutional relationships. For example, vulnerability involves being 'susceptible to harm or under threat of harm'; it is 'a universal condition that humanity has in virtue of our physical or social nature' and 'involves one or more specific attributes, contexts, or group type' (Wrigley and Dawson, 2016, chapter 7, section 7.2).

There is limited difference between people identified as marginalised/vulnerable and for clarity, the research approach for this project has focused on the **terms 'vulnerable' and 'vulnerable groups'.** The reason being marginalisation is dependent on the historical and socio-economic societal context. The vulnerability framework focuses on establishing the parameters of government's responsibility for societal structures, institutions and its relationships that are responsive to structural inequality and reduction of risk of harm especially within the targeted vulnerable community (Fineman, 2019). As Crenshaw states, 'addressing the needs and problems of those who are most disadvantaged [vulnerable]' involves fostering communities and technologies where multiple voices feel valued to address inequalities and injustices in AI. By focusing on intersectional responsible AI that is specifically centred on vulnerable communities, emphasis is placed on prioritising harm reduction thereby promoting justice for vulnerable communities (Ovalle et al., 2023). From data to design to implementation, from technologist to user, Responsible AI should foster an inclusive system that promotes an equity-centred, long-term sustainable approach.

Challenges to the intersectional research approach

One of the goals of the research was to incorporate intersectionality as part of the data analysis, to ensure that participants were not viewed as being part of a homogenous social category and, instead, that the researchers took into consideration the complex interplay of social and personal identities which may impact of how they engage with the Committee and the wider relationship

[12] https://www.ipbes.net/node/42125.

[13] https://www.ipbes.net/node/42127.

with the policing ecosystem (Behizadeh, 2024). To achieve this, a semi-structured interview methodology was adopted, and community representatives were selected from a variety of different advocacy organisations. Conversations of intersectionality, however, were not initiated by the interview subjects and therefore do not feature strongly in the research. This may be due to community representatives' roles in supporting **specific groups** which meant they were, therefore, more focused on ensuring the challenges and issues associated with AI in policing for those groups were identified in the research (see: Research limitations). Interviewees did acknowledge, however, that multiple, and intersecting, identities could have an impact on individual experiences with various policing bodies. Therefore, although the research does not explicitly address intersectionality and the related challenges, there is an implicit recognition amongst the involved stakeholders, that vulnerable groups face compounding challenges and vulnerabilities. These issues and challenges are also acknowledged, either implicitly or directly, by the **human rights framework**, which is discussed in the following section.

Implications of AI-supported policing for human rights

The UK has international obligations under various human rights treaties and instruments, including the **European Convention on Human Rights** (the 'ECHR') (Colvin and Cooper, 2009). When the Human Rights Act 1998 ('HRA') came into force, the Convention rights became part of UK domestic law (Colvin and Cooper, 2009; Fenwick and Baker, 2017). Certain Convention rights are understood to be 'absolute', where state intervention will never or only rarely be lawful. Other rights are 'qualified' (Colvin and Cooper, 2009); state interference with these rights may be lawful depending on whether interference:

- is in accordance with the law
- has a legitimate aim
- is necessary in a democratic society to achieve the legitimate aim (the 'proportionality' limb of the test) (Murray, 2020).

Where the rights conferred by the HRA are at stake, **proportionality** is assessed according to a structured four-part test (Ramshaw, 2019; Janjeva, Calder and Oswald, 2023). The test is said to involve 'an exacting analysis of the factual case advanced in defence' of the act or measure giving rise to the interference to assess:

> *(i) whether its objective is sufficiently important to justify the limitation of a fundamental right; (ii) whether it is rationally connected to the objective; (iii) whether a less intrusive measure could have been used; and (iv) whether, having regard to these matters and to the severity of the consequences, a fair balance has been struck between the rights of the individual and the interests of the community*.[14]

Carter (2021) points out that 'when considering whether the benefit is worth the cost of limiting the right' (the 'balancing' part of the proportionality test), judges:

> *must necessarily be informed by some empirical understanding of the real-world impacts of the law in question* (Carter, 2021).

Human rights in practice

Section 6 of the HRA imposes a requirement on all public authorities (including police authorities and courts) to act in a manner that is **compatible** with the Convention rights, and therefore public authorities must have policies, procedures and practices to ensure that they comply with this duty.

Murray (2020, p.158) notes that 'using human rights law to inform states' decision-making processes is not straightforward'. For some commentators this difficulty flows from the nature of the proportionality test itself (Urbina, 2017; Chang and Dai, 2021). For others the difficulty is of a more practical character and concerns both the nature of the information which must be considered in assessing proportionality (Carter, 2021) and the task of weighing up this information (Janjeva, Calder and Oswald, 2023). A public authority must, minimally, weigh up information about 'the extent of the benefit to be achieved and the degree to which the right will be impacted' (Carter, 2021, p.60), in effect carrying out a 'pre-deployment impact assessment' (Murray, 2020). It may be difficult to 'account in advance for all possible contextual factors that may be relevant in each individual case', exacerbated when the rights intrusion results from the use of automated analytics (Janjeva, Calder and Oswald, 2023).

[14] *Bank Mellat v HM Treasury (No 2)* [2013] UKSC 39 [20] (Lord Sumption JSC)

Human rights at stake in the use of AI-supported policing

Various ECHR rights may be at stake in the context of the development and operationalisation of AI tools in policing. Beyond the ECHR, the right not to be subjected to **purely automated decision making and profiling** (Data Protection Act 2018, s.49) is also applicable, as well as the **Public Sector Equality Duty** under the Equality Act 2010. Perhaps most relevant to the criminal justice context is the right to a fair trial under **Article 6** ECHR (Moreira, 2022). Both the presumption of innocence (Sachoulidou, 2023) and equality of arms (Stoykova, 2023), which fall within the ambit of Article 6, can be impacted. Police use of AI may make it harder for the defendant to effectively challenge evidence, especially where there is a lack of transparency (Palmiotto, 2021). The right to liberty under **Article 5** ECHR is relevant too, where there are concerns about the reliability of tools and the risk of unlawful arrest or detention (Golunova, 2023).

Rights under **Articles 10** (freedom of expression) and **11** (freedom of assembly) can be impacted by police use of AI tools in surveillance at large gatherings and legitimate protests, as well as analysis of social media use, all of which could potentially have a 'chilling effect' (Leslie et al., 2021). The right to respect for family and private life under **Article 8** ECHR is also at stake through the use of AI tools in policing (Grace, 2018). In the Court of Appeal case of Bridges[15], the Court of Appeal found the interference with Mr Bridges' Article 8 rights by police use of Automated Facial Recognition was not in accordance with the law (Purshouse and Campbell, 2022). Other legal challenges alleging breaches of Article 8 by police use and retention of data (where there has been no charge and/or no conviction) are also relevant.[16] There have been differing outcomes, but unanimity about the engagement of Article 8 in these contexts. The use of AI tools to process and analyse large amounts of data held by the police makes the question of the lawfulness of interference with Article 8 rights both more complex to analyse and more likely to arise.

The prohibition of discrimination under **Article 14** ECHR can also be impacted by police use of AI tools, where concerns arise about such tools amplifying existing discrimination (Borgesius, 2020). It has even been suggested that decisions about the types of offences police target with AI tools could be a possible source of discrimination, where these offences are more likely to be

[15] *R (on the application of Bridges) v Chief Constable of South Wales Police* [2020] EWCA Civ 1058

[16] *Catt v United Kingdom* [2019] ECHR 76

carried out by poorer, more vulnerable members of society, rather than targeting white collar crime (Castets-Renard, 2022).

Conflict and relationship between rights

It is also important to note that different rights can be in conflict in the criminal justice context. For example, in rape trials there can arise an issue of balancing the defendant's right to a fair trial under Article 6 ECHR against the complainant's right to privacy under Article 8 ECHR, where the defence wish to adduce sexual history evidence of the complainant (Sous, 2020). Grace argues that AI tools could 'exacerbate and inflame human rights tensions that already exist in criminal justice settings' (Grace, 2021). Therefore, as well as considering the impact of AI-supported policing on specific human rights in the criminal justice context, there must also be consideration of how police use of AI tools might impact the way the balance is struck between competing rights.

In addition to rights which are at stake in the use of AI-supported policing, there are also rights which might be **supported** through the use of AI tools. The state has a positive obligation to conduct an adequate investigation into alleged breaches of **Articles 2 and 3**, for example, the police must properly investigate an allegation of a serious sexual assault (Bates, 2020).[17] The use of AI tools could support this obligation. However, if not used in a way which is compliant with legal obligations, there is the potential for an adverse impact on the investigation – it could potentially lead to evidence being inadmissible or opening up the potential for a challenge on the basis of a breach of other rights (Scottish Human Rights Commission, 2021).

3. Research findings

The following sections present the research findings, organised across three thematic areas: **The effectiveness of the Committee including findings from technical observations; Community engagement; and Ecosystem communication**. The presented data is drawn from the interviews, document review, technical observations and where appropriate, triangulated with literature. Each section includes a summary that highlights the main findings, and practitioner

[17] In *Commissioner of Police of the Metropolis v DSD and another* [2018] UKSC 11, the UK Supreme Court confirmed that two women who had been subject to serious sexual assault by John Worboys were entitled to compensation following police failures in the investigation.

advice, which structures those findings in a manner which may be beneficial to other ethics committees.

3.1 Effectiveness of the Committee

The findings in this section are particularly relevant for our research questions 1, 2 and 5 though they have cross-cutting relevance for all our research questions. They explore these questions with reference to the effectiveness of the Committee, examined across three themes: whether the Committee fulfils its remit particularly in relation to **human rights** (RQ 2); whether its advice and recommendations are **followed**, including findings from technical observations (RQ1); and whether it contributes to an **increased understanding** of the ethical issues at stake (RQ 2 and 5). The themes align with measures of success that are adopted in other institutional environments, where effectiveness is explored according to criteria which include performance of core functions, delivering change, and increasing education and awareness among stakeholders (Coleman and Bouëwhile sseau, 2008; Crico et al., 2020).

The Committee's remit

The question of whether the Committee fulfils its remit is understood through two perspectives, whether the Committee provides a **thorough ethical analysis** of projects that come before the Committee for consideration and whether it puts '**rights at the heart**' of ethical review (West Midlands Police & Crime Commissioner, no date). The Terms of Reference of the Committee set out its remit – what it is supposed to do, and, in broad terms, how it should do it. Those terms set a benchmark against which the effectiveness of the Committee may be assessed. The role of the Committee is to provide 'independent advice … regarding the ethical issues arising from the Lab's work', that is the development and operationalisation of a broad range of tools (grouped, for our purposes, under a broad definition of 'AI') for use in a law enforcement context (West Midlands Police and Crime Commissioner, 2019). As one Committee member explains:

> … *it's deliberately* ***not meant to be a generalised Ethics Committee****. It's meant to be around* ***data science and tech****…* - Ethics Committee representative 6

The Committee is tasked with 'conduct[ing] a **thorough ethical analysis**' and 'provid[ing] **practical and independent advice**' (West Midlands Police and Crime Commissioner, 2019). The terms of reference set out a list of ten principles which the Committee should apply when delivering its advice. These include whether the proposed use of data produces benefits or

causes harm, legal requirements, necessity and proportionality, the quality of the data and the data analysis, the risks of new technologies, non-discrimination and taking the view of the public into account.

The Terms of Reference also set out a non-exhaustive list of questions. The implication is that, where appropriate, the Committee might put these questions to the DAL during ethical review of a project. Like the principles, these questions explore issues relating, for example, to **data quality, human rights compatibility, bias, necessity and proportionality and the views of the public**. If a focus on outcomes is implicit in the principles it is explicit in the questions.[18]

The remit of the Committee commits it to a form of review which extends beyond a narrowly conceived notion of 'ethics'. The Terms of Reference mandate an **analysis grounded in legal considerations (especially as regards human rights), good data science, an outcome-oriented analysis of benefits and harms, and an eye to public acceptability and engagement**. This mandate is reflected in what the Terms of Reference say about the composition of the Committee, namely that it should comprise persons with various backgrounds including data science specialists, a human rights expert, and members 'who are residents of the West Midlands' (West Midlands Police and Crime Commissioner, 2019). Such different dimensions of the Committee's mandate and intended focus are reflected in one member's summary of key considerations for the Committee:

> *And so, … those are like the tiers I would say … make sure it's lawful, make sure the data science is rigorous, whatever that might mean given the project and then thirdly, make sure that you are doing good.* Ethics Committee representative 3

Providing ethical scrutiny

There is widespread consensus among interviewees that the Committee provides meaningful scrutiny of projects:

[18] The assessment of lawfulness under the human rights regime entails a focus on outcomes. As Janjeva et al. note 'Understanding the potential impact of practical applications of automated analytics is essential in weighing up utility and harm, and therefore to assessing proportionality' (Janjeva, Calder and Oswald, 2023). See also (Murray, 2020).

> *… [the committee] enables us [the Data Analytics Lab] to have our work scrutinised, questioned by a group of experts and people with some local background and interest in the West Midlands on behalf of the public.* Data Analytics Lab and Police representative 2

This scrutiny moves beyond simply challenging the explicit design and implementation of projects, but includes help in identifying and addressing problems or areas of concern that may **not have been anticipated** during the design stage:

> *So, I would say that the impact that we make is quite general … We tend to bring up things that the police find relevant on many different occasions*. Ethics Committee representative 2

> [*the Committee*] *… would point out questions that we haven't thought of … which is always useful.* Data Analytics Lab and Police representative 3.

As regards the spread of issues addressed by the Committee, interviewees report that it raises questions about impact on human rights (legal considerations), data quality, bias (questions about good data science) and disproportionality (outcomes):

> *Human rights concerns are prevalent, especially around privacy, expectation of privacy, to fundamental rights around 'innocent until proven guilty'. Those concerns do come up from time to time. Storage of data and 'just in case' storage of data … [are also raised]*, Ethics Committee representative 10

> *common themes … in the past have been data quality and anonymisation. And latterly … it's more to do with … privacy, potential impact on different demographic groups* … Data Analytics Lab and Police representative 3

The Committee, therefore, devotes significant consideration to questions of 'scientific validity, legal proportionality, and operational context' (Oswald, 2022). The Committee also endeavours to consider the public view, though there are concerns about the extent to which it can do that effectively without greater diversity in its composition:

> *So, do I think the Committee tries really hard to think about how the public thinks about these things? Yes, but by definition,* ***they can't do that for everyone, and are there voices missing? Yes, would be my answer***. Ethics Committee representative 10

One interviewee registers concern about whether sufficient attention is devoted to outcomes:

> *I want to know what the **worst-case scenarios** are of either the data science being flawed, but also the outcomes in how it is used which, quite frankly, is still not given enough attention in the Committee, in my opinion*. Ethics Committee representative 6

Technical vs operational considerations

This observation is noteworthy since ethical review of a project, whether from a human rights perspective or with an eye to whether the project results in public benefits or harms, involves consideration of outcomes. **There is a suggestion that the lack of visibility of likely outcomes sometimes results not from reluctance on the part of WMP to provide this information, nor from a failure on the part of the DAL to recognise the significance of outcomes, but because representatives of the DAL do not have control over how their tools are operationalised**. One DAL representative explains:

> *... in some instances, the Committee might be slightly treading into, as opposed to the project per se, they might be looking more at the, what the police would do as BAU [business as usual], which obviously... we essentially have no control over because it is police BAU. So ... some of that stuff might be ... sort of outside of the purview of the lab ... but obviously, it is a concern of the Committee as to what the final use would be of most of these things*. Data Analytics Lab and Police representative 3

Another interviewee suggests that questions about the projects such as

> '*why do you want it? And what are you going to do with it*?' (outcomes-related questions) are '*actually better asked of the business, as in the operational officers in my context, the cops, than... of the data scientists*.' (Data Analytics Lab and Police representative 6)

The interviewee expresses a wish:

> *... to move the ethics panel's conversation away from the technical - what is the ... data scientific process that you've used in that. There'll be a space for that of course. But more... **they should be holding the business lead to account** ... it's not right that [name redacted] argues why the business needs a particular product. [name redacted] is a data scientist ... not a police officer*. Data Analytics Lab and Police representative 6

While the DAL may have no control over operational activities, **the suggestion that the DAL is not well-placed to answer questions about the purpose and intended use of projects they develop is surprising even if police officers are better placed to explain the business need**. Such information would seem to be highly relevant for the design of the projects. Information about purpose and intended use is also essential for legal compliance, both for the completion of Data Protection Impact Assessments (DPIAs) [19] and for assessment of lawfulness under the HRA.

Interviewees express concerns about the failure of the projects to return to the Committee after operationalisation. **Such failure would limit the ability of the Committee to assess whether there are 'positive and/or negative outcomes arising from an implementation of operational proposals arising from a project's findings' and 'what steps should be taken to remedy identified problems or improve outcomes'** (West Midlands Police and Crime Commissioner, 2019).

The interviewees' descriptions of the spread of issues raised lends some support for the assessment by one Committee member that '*generally speaking, … it's doing its core functions*.' (Ethics Committee representative 6) Nevertheless, there is room for improvement in certain respects notably in relation to:

- the involvement of **lay members** to enhance the ability of the Committee to meet its mandate to consider the public view.
- a seeming lack of clarity about what information needs to be shared, when, by and with whom, about the **intended purpose and use of projects and the likely impacts**, whether for the purpose of ethical review or to guide the development of the project.
- the absence of clear protocols about 'live' projects coming back to the Committee for consideration **post-implementation**.

[19] DPIAs must take into account 'risks to the rights and freedoms of natural persons' (ICO, no date b). These risks are not limited to risks to data protection or privacy but include risks to other rights and freedoms. The ICO's 'AI and data processing toolkit' alerts to the risk of '[f]unction creep over how personal data is processed is caused by not defining what [sic] purpose you will use your AI system' (ICO, no date a).

Putting rights at the heart of ethical review

It is not part of the remit of the Committee to provide legal advice. However, the Terms of Reference expressly state that the Committee 'will, of course, consider relevant legal issues … as part of its analysis' (West Midlands Police and Crime Commissioner, 2019). Despite the generality of the requirement to consider legal issues, the Terms of Reference make it clear that the Committee is expected to routinely evaluate project proposals for conformity within the context of human rights.[20] Several questions in the 'non-exhaustive list of questions' in the Terms of Reference are oriented to a human rights analysis:

- Will anyone's privacy or human rights be infringed?
- Will anyone be discriminated against?
- Is the proposed project necessary and proportionate (e.g. is the analysis reasonable and/or is the use of personalised data, where applicable, proportionate?)[21] (West Midlands Police and Crime Commissioner, 2019).

The focus on human rights and non-discrimination is also reflected in what is said in the Terms of Reference about the make-up of the Committee, in particular the inclusion of an experienced human rights expert and an expert in privacy, bias and discrimination (West Midlands Police and Crime Commissioner, 2019). The Terms of Reference, therefore, have human rights at their core.[22] Interviewees suggest that this focus is reflected in the approach adopted by members:

> *generally I think most members of the committee are quite concerned that there isn't bias or inadvertently that bias isn't introduced into the data in such a way that it would affect potentially the rights of individuals. And then more broadly, does the end justify the*

[20] Yeung makes a strong case for a human rights-centric approach to securing ethical AI. Such an approach, she suggests, is an 'important element in the overall sociopolitical landscape needed to build a future in which AI systems are compatible with liberal democratic political communities in which respect for human rights and the rule of law lie at its bedrock.' (Yeung, Howes and Pogrebna, 2020) See also (Oswald, 2022)

[21] The tests for necessity and proportionality are core to an assessment for conformity with fundamental human rights. It is also essential to carry out an assessment as to necessity and proportionality where the processing of personal data is at stake.

[22] We note that the 'Toolkit for Responsible AI Innovation in Law Enforcement' developed by INTERPOL and UNICRI states that 'To help determine if an AI system affects human rights, law enforcement agencies should involve ethics and human rights experts in carrying out a human rights impact assessment.' (UNICRI and INTERPOL, 2024).

means? Is it proportionate? The use of data that can affect individuals - is it proportionate and necessary? Ethics Committee representative 4

what the Ethics Committee ensures is that there's a human rights focus on what the lab does. Which I think is a positive thing. Data Analytics Lab and Police representative 4

Breadth of rights considered

Some interviewees expressed concern that the Committee is overly focused on certain human rights, particularly the rights conferred by Articles 8 and 14 of the European Convention on Human Rights (respectively, the right to respect for private and family life and the right to protection from discrimination):

I would say they're very much focused on ... privacy rights and discrimination rights. Less attention is given to utility of the capability and the degree to which it prevents harm, which ultimately is your article 3 right. Data Analytics Lab and Police representative 4

*People don't normally talk about right to a **fair trial**, but it's clearly kind of implicit in everything. Is this fair treatment of them in the process or fair assumptions to make about people in a context that could harm their life? I don't really think any others really come up really. It's privacy versus right to life.* Ethics Committee representative 6

A focus on the rights to privacy and protection from discrimination is unsurprising. It is noted that when asked about potential human rights implications of the use of AI in policing, a majority of the Community representatives interviewed highlighted implications for the right of privacy. An ethics committee must inevitably concern itself with legal issues relating to the privacy implications of the use of such data. The potential for discriminatory effects resulting from the use of policing data in general, and stop and search and arrest data in particular, is well-documented. Police representatives drew attention to these concerns:

*So, the data quality might be perfect, but that doesn't mean that there's not bias within the data. And again, **stop and search** is a prime example of that. You can have very, very accurate stop and search data, but the people who were being stopped, are disproportionately from one particular group, so it's not the data itself, which is dodgy, it's the activity which the data reflects.* Data Analytics Lab and Police representative 6

Concerns about possible discriminatory effects of the use of AI in policing were shared by one Community representative:

> *So, is AI going to be targeting certain sections of the community, demographics, equality strands, protected characteristics, what are you using it for?* Community representative 10

Nevertheless, the criticism that the Committee is overly focused on privacy and discrimination merits further attention.

One Community representative expressed concern that a consequence of the use of AI in policing is that 'we are no longer innocent until proven guilty, cause the concept assumes that everyone is under suspicion' (Community representative 6), with implications for the Article 6 right to a fair trial. In fact, given the spread of projects examined by the Committee, one might expect the minutes to record questions relating to the right to liberty (Article 5), the right to a fair trial (Article 6), freedom of expression (Article 10) or freedom of assembly (Article 11). Of these, **a search of the minutes reveals only one express reference to Article 6 and the right to a fair trial**. These rights may not be relevant for some of the projects considered by the Committee. For example, they are unlikely to be relevant (or, at least, directly and immediately relevant) for the DAL's explanatory or exploratory projects. However, they are likely to be relevant to varying degrees for predictive projects, including those that seek to predict the likely location of crimes or those persons who are likely to commit certain crimes.

One might also expect the minutes to record comment or discussions about the implications of Articles 2 (right to life) or 3 (prohibition on torture). Both rights impose **positive obligations** on the state to take operational measures to protect victims or potential victims of crime, investigate crime and take measures to prevent crime (Colvin and Cooper, 2009). **A search of the minutes discloses only one explicit reference to Article 2.**

Oswald notes that '[t]he majority of the projects reviewed by the committee engage ECHR rights, not only Articles 5 (liberty and security), 6 (fair trial), 8 (respect for private life) and 10 (freedom of expression), but also Articles 2 (life) and 3 (freedom from torture and inhuman and degrading treatment) in relation to positive protective or preventative obligations.' (Oswald, 2022) It may be, however, that there is a reluctance on the part of the Committee to seek to offer a full-blown rights analysis when it is not part of their remit to offer legal advice. As one interviewee puts it:

> *… it really is beyond the limits of the Committee to do their [West Midlands Police's] legal work for them*. Ethics Committee representative 6

Operationalisation of projects and the impact on rights

The question of whether the risks to the right to liberty, right to a free trial, freedom of expression and freedom of assembly are theoretical or real will depend not only on the technology but on how and for what purposes the technology is deployed. For instance, the risk of a **'chilling effect'** (the discouraging of a legitimate exercise) on freedom of expression and freedom of assembly may be context-specific and so difficult to measure and assess. The Committee has expressed frustration about lack of information about the 'how' and 'why' of operationalisation on more than one occasion.[23] For example, in the minutes of the 21st July 2021, it was recorded that *'It's really hard for as [sic] a committee to look at the ethical impact of using the model without understanding the clear path effectively to the operationalisation of the model and how it's actually going to be used within the Force'*. Interviewees also describe how certain human rights issues are raised in Committee but are **not explicitly articulated** in terms of human rights:

> *… everything's inherently ultimately about discrimination, I guess, but people don't really talk about the human rights provisions when we're talking about discrimination, it is a bit more specific then. That's where I guess conversations about the use of police powers become most pivotal, really*. Ethics Committee representative 6

> *So going back to Most Serious Violence* [the name of one of the projects] *when we're having this conversation about the fifty four percent, and I was like, that's not enough, I guess you could frame that as a human rights issue. I mean, I don't think the conversation went, oh, you're infringing on their human rights, but it does infringe on their right to privacy and family life, doesn't it? With us knocking on the door*. Data Analytics Lab and Police representative 5

According to one interviewee, even where the right to privacy is at stake:

[23] For example, the minutes for 21 July 2021 record the following comment: 'It's really hard for as [sic] a Committee to look at the ethical impact of using the model without understanding the clear path effectively to the operationalisation of the model and how it's actually going to be used within the Force.'

> *they [the committee members] would never say I'm raising this under the Article 8 right to a private family life. But the nature of the conversations the communication it always kind of felt as though it was in that that area of a person's right to a private life ultimately.* Data Analytics Lab and Police representative 1[24]

Similarly in the minutes it was found that the Committee raised questions about whether predictions as to the likely location of knife crime 'would be part of a rationale for s.60 Criminal Justice and Public Order Act 1994 authorisations'. This consideration is highly relevant to the right to liberty but the Committee did not expressly invoke the right.

A reluctance to expressly invoke particular rights might be attributable to a commitment to the 'pragmatic approach' described by Ethics Committee representative 8 and mandated by the Terms of Reference (West Midlands Police and Crime Commissioner, 2019). The minutes, moreover, are not a verbatim record of the Committee's deliberations. It may be that the full spread of rights that might be engaged did not form part of the Committee's deliberations, but the scarcity of express references to the rights at stake makes it difficult to assess which rights were considered. For some interviewees the absence of an explicit human rights discourse is not problematic:

> *... in these discussions I'm more interested in people's intuitive sense of right and wrong because human rights do really go to that, they kind of are the most obvious principles really.* Ethics Committee representative 6

However, the lack of explicit reference to the human rights issues addressed by the Committee impacts on the utility of the minutes as a source of learning and aid to transparency. If the Committee's deliberations are to inform an understanding of rights implications and serve (where appropriate) to raise and record relevant legal issues, then explicit reference to the rights at stake is appropriate. This does not necessitate a requirement for a full human rights analysis within the Committee discussions, but that its recommendations are

[24] The Committee has met on 20 occasions to date [July 2024]. Minutes have been published for 19 of these (all but the most recent). A search for the term 'privacy' disclosed only ten instances of the term in the minutes and in three of these cases the issue appears to have been raised by the presenter, not the committee. 'Discrimination' or 'discriminatory' are mentioned nine times. Articles 2 and 6 are each expressly raised in the committee's comments only once, and Article 8 is mentioned three times. It is worth noting that three of these express references (one to Article 6, two to Article 8) were made in written comments concerning the Home Office Retrospective Assisted Facial Recognition project.

supplemented with a rationale which notes human rights, and other key technical, legal and social issues, including issues relating to disproportionate impact on vulnerable groups and those at risk of intersectional discrimination.

The implementation of the Committee's advice and recommendations

To measure the Committee's effectiveness, it is important to understand how, and if, its recommendations and advice are implemented and operationalised. One interviewee describes the process of exchange of questions and answers between the Committee and the DAL:

> *… they [the Committee] ask questions so, then we ask questions and then go back to them, … it allows us to ask the question, because obviously at the time, it didn't occur to us or whatever.* Data Analytics Lab and Police representative 3

The interviewee provides an example of the effects of this exchange:

> [speaking of the review of the IOM model by the Committee] … *when we originally went, the question arose as, okay, you've got these features. Do any of these features actually allow… are they a back door to identify an ethnicity? So, we then went away and basically did a multinomial on it, using the features that were in the model and no, they're not. So, provably using this data you could not identify anybody's ethnicity using the features that are in there.* Data Analytics Lab and Police representative 3

Such exchanges are beneficial. They aid transparency and may contribute to good practice. For instance, if the DAL intends to exclude ethnicity from the features of a model, it is important for them to be able to demonstrate that none of the features that are used operate as a proxy for ethnicity. Bias analysis and the removal of features which are a proxy for protected characteristics are important mechanisms for the development of responsible AI (Wang, 2021; ICO, no date a). Nevertheless, the interviewee makes an important point when they maintain that the exchange of questions and answers '*doesn't fundamentally change what's being provided or done.*' **If the Committee is to contribute to responsible AI in policing, it is essential that it not only raises the right questions and makes appropriate recommendations, but that these inform the development of the projects**. In fact, a slew of comments tends to suggest not only that '*[t]he committee is heeded'* (Ethics Committee representative 10), but that the Committee's advice does inform the development of projects. One interviewee suggests that while the DAL revised

projects in response to the Committee's recommendations, the revisions did not necessarily address the problems that had been identified:

> [describing the exchange between the Committee and the DAL] ... *and they're saying, well, you know it's not right for all these good reasons but then this unit's not really understanding what they're hearing and so they're making revisions that don't necessarily address the problems and so that whole thing ... despite the extremely high quality of the Ethics Committee, to my mind wasn't working*. Ethics Committee representative 5

However, another had a much more positive view of the effects of the Committee's recommendations:

> *almost everything we've considered, we make better. I mean, ... there will be things tweaked in it or questions raised that make people sometimes remodel. Almost invariably, you know, time after time that kind of thing happens.* Ethics Committee representative 2

One interviewee, indeed, expresses frustration that the Committee has exerted too much influence over the development and operationalisation of projects:

> *individuals involved in this process have got confused around what their role is. so ... [there are] data products that ... [the] data analytics lab have produced that are not live but are built and the reason they're not live is because they are waiting for permission from the data analytics lab [sic]. That's just wrong because that is not the process that we have. ... West Midlands Police is an independent body and the decision-making sits with the chief constable, ... The ethics panel is there as a tactical advice mechanism, a challenging tactical advice mechanism, don't get me wrong, but they don't... they do not make decisions.* Data Analytics Lab and Police representative 6

Comments by another interviewee tend to confirm that some projects did not proceed following the Committee's advice. The interviewee presents this outcome in a positive light:

> *And there have been... that is probably a real testament to, I would say, how seriously the committee is taken ... going back to an earlier question about, 'do police, does policing listen to what the committee says'? In the master document you will have seen there are numerous projects that did not continue; they ceased. And that's probably the biggest testament to that.* Ethics Committee representative 10

Shaping decisions

Some of the most compelling statements about the impact of the Committee are from police representatives who describe how the committee's recommendations **shaped their own decisions** about projects:

> *... nothing springs to mind immediately where there was a recommendation, and I thought, no, I'm not going to do anything about that. There were lots and lots of examples of where there were recommendations that I did do something about.* Data Analytics Lab and Police representative 1

Interviewees mention examples of this process including a project which was questioned by the Committee and did not proceed. The Mo]st Serious Violence project, an initiative of the National Data Analytics Solution (a centralised advanced analytics capability for UK policing) did not proceed, in part, because of the Committee's concerns about the **low accuracy scores** of that model. The project was later to be found to have an error in the coding:

> *... the committee I think rightly asked for one project to be cancelled altogether on account of it being highly inaccurate ... it's called most serious youth violence, I think*. Ethics Committee representative 6

> *It turned out there's an error in the* code [in the Most Serious Violence project]*, but we were going to the Ethics Committee to say, well, with most Midlands police data, we can give you a fifty four percent prediction that, you know, that [name] is going to go and commit an offence in the next two years with a gun or a knife. And the Ethics Committee were constantly coming back and saying fifty four percent ain't good enough. So, but ultimately, there was an error in the* code*. And the actual figure was much, much less than that anyway ... And we did shelve that project partly as a result of that* ... Data Analytics Lab and Police representative 5

Interviewees offer other examples. Stop and search data without positive results was removed from West Midlands Police Integrated Offender Management model because of the Committee's concerns. Speaking of the IOM model, one police representative notes:

> *... the natural default position of many people on the committee is start with privacy, start with discrimination and that can be very healthy and there's been positive outcomes as a result of that where specific data sets have been taken out of models to ensure against*

> *any potential bias and actually, through that process, the model has been refined to a point where the accuracy and effectiveness was- was not degraded, but the assurance around potential discrimination was achieved*. Data Analytics Lab and Police representative 4

The briefing papers confirm that stop and search data without positive finds has been removed from the model. Another Committee member refers to the Committee's advice about the use of **suspect data** in the IOM model:

> *So, there's a project called the Integrated Offender Management tool… the committee saw that proposal quite early on and made the observation that some of the data collection and input was at quite a low threshold, effectively people being suspected of an offence, if I remember correctly, not being charged with offences. And so, the committee's advice was the tool needs to be built using a body of data, a dataset that only contains people being charged*. Ethics Committee representative 3

However, a Data Analytics Lab and Police representative notes:

> *… obviously, for example, on the IOM model, it was originally just looking at people who'd been… who had reached the offender status, so they've been charged with stuff. As it transpires, there's issues with that, but … that decision was originally taken because at the time and that was a long time ago now, but at the time, it seemed like that was probably the slightly more ethical thing to do*. Data Analytics Lab and Police representative 3

The minutes and briefing papers suggest that suspect data relating to people who were not also offenders was not included either in the original model or after the rebuild of the model between July 2019 and January 2020. Internal and external evaluations of the model were carried out over the course of October and November 2022. The relevant reports were presented to the Committee in November 2022. The internal evaluation report recommended the use of suspect data on the basis (i) that offender managers already use suspect data, and (ii) use of such data both 'helps to identify opportunities for interventions aimed at preventing further criminality' and makes decisions 'more robust'. The Committee expressed unease about the inclusion of suspect

data and sought clarification. At the time of writing discussions are ongoing between the DAL and the Committee about the proposed inclusion of suspect data.[25]

Tick-box compliance?

Although many interviewees have strong views on the extent to which the Committee's advice makes a difference, one Committee member expresses a degree of uncertainty:

> *So, the question you have to ask yourself, then, is, … 'is this what I term 'tick-box compliance', or is it really making a difference'? I hope it's making a real difference, and listening to some of the things, … I get the impression that it is making a difference, but I, as a committee member, can't actually see that, I don't think. But that might be just me.* Ethics Committee representative 8

Ideally stakeholders and other interested persons should be able to access sufficient information to draw their own conclusions about the effectiveness of the Committee. Neither the minutes nor any other (publicly available) document systematically records, for each project, whether the Committee's recommendations were followed. Members of the public, journalists and researchers will not always have the luxury of speaking directly to persons involved in the work of the Committee, nor the resources to carry out a detailed, longitudinal analysis of the minutes.

Technical observations

The research team are grateful to the DAL who allowed technical observations of the beta version of the RFSDi/IOM (Recency, frequency, severity harm score/Integrated Offender Management) tool and the fully operational Crime Seasonality Planner. Both tools were developed in-house by

[25] Inclusion of suspect data was discussed at the committee meeting in June 2024. The Data Analytics Lab propose to use suspect data (data relating to someone officially flagged as a suspect but not yet prosecuted or charged) in the RFSDi element of the RFSDi/IOM model. The perceived advantage is that inclusion of this data will, in some cases, result in a more accurate harm score and contribute to better policing. Some members of the committee continue to have concerns relating to implications for the presumption of innocence, the possibility that the system may vastly increase the number of people flagged as high risk, and the implications for police workload and resources. The Committee requested assurance in respect of circumstances in which an individual might be flagged solely on the basis of suspect data: file:///C:/Users/VGYK3/Downloads/ETHICS-COMMITTEE-Minutes-and-Advice-12.06.2024.pdf

DAL. These observations allowed the researchers to draw additional conclusions regarding the impact of the Committee's advice and recommendations on technical developments.

RFSDi/IOM model

The RFSDi/IOM tool is intended for use by Local Offender Management Units. It has two elements. The RFSDi element calculates a harm score for offenders. The harm score is calculated with reference to factors relating to the **recency, frequency and severity** of crimes as well as information about misuse of drugs and intelligence about the offender. The IOM element is a **predictive** model that forecasts the probability of an offender engaging in high harm crime over a coming two year period. The outputs for both elements are displayed on an interactive dashboard.

An Internal Evaluation Report prepared by the DAL suggested that the IOM element is not much used by offender managers, although the model may contribute to decision-making in borderline situations (West Midlands Police, 2022). This Report also showed that the **precision** of the IOM model performance had deteriorated since initial testing of the model while accuracy, sensitivity and specificity of the model were largely unchanged. It was noted that a value of 0.6 was set as the cut-off for classifying a predicted class of high harm. The reported metrics were as follows:

Measure	Result
Accuracy	0.92
Sensitivity	0.75
Specificity	0.93
Precision	0.38

Figure 2 IOM Performance metrics calculated between September 2021 to April 2022 (West Midlands Police, 2022)

A review of the metrics from the model run in October 2021 also showed that precision had fallen 'precipitously', with DAL commenting that it was difficult to assess the accuracy during a period when only a small number of individuals had transitioned to high harm, and where there had been a large jump in those classed as high harm in January 2022 (the reason for which was unclear) (West Midlands Police, 2022). The initial results of a re-build of the model in 2022 showed significant improvements in precision rates (West Midlands Police, 2022). The DAL were particularly concerned that the model should have a high specificity rate, that is, that the model would **correctly classify actual negatives as negatives** (West Midlands Police, 2019). Should the low precision of the model continue, combined with the high specificity measure this might suggest that, taken together, the tool should only be used to identify people who are *unlikely* to

transition to high harm crime; rather than to identify those who might transition. The minutes for April 2019 note that the Committee commented positively on the DAL's focus on avoiding people being wrongly identified as at risk of transitioning to high harm offences. There is no record in the minutes for November 2022 or subsequent meetings that the Committee commented on the drop in the model's precision rate, which might highlight the **difficulty of assessing the meaning and consequences of changes in performance metrics**. It should be noted that the recent re-build of the model to include suspect data, as discussed at the Committee meeting in June 2024, was intended to address the low precision issue and the dissatisfaction of officers with a large number of 'missing nominals'.[26]

Furthermore, in the case of the IOM model, the DAL used Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan, 2003) for the purposes of feature engineering: converting raw textual data into features (in this case topics) that are suitable for input to a machine learning model (Arlitt, Khan and Blessing, 2019; Panda and Misra, 2021). Topics are probability distributions (or less formally, 'statistical mixtures') of the words that appear in the data. One disadvantage is that the method depends on selection and generation of an appropriate number of topics and that the Committee sought clarification of this aspect at a meeting in April 2019. The DAL provided a list at the following meeting. The list includes thirteen features labelled as 'topics' (e.g. 'topic1_max_value')[27] (West Midlands Police, 2019). The briefing papers for the IOM model mention LDA (West Midlands Police, 2020) but do not mention that it is used for feature engineering. In principle, the use of LDA in feature engineering introduces an **additional layer of uncertainty**: the predictive model is built on features that are themselves generated through probabilistic methods. However, the features are selected by reference to the model's performance. Only the top 50 features were used. It appears however that the features generated through LDA were not used after the IOM model was rebuilt (West Midlands Police, 2020). This technical aspect of the model again highlights the potential challenges for Committee members in understanding and tracking feature engineering aspects and changes, and the importance of ensuring that such systems contain functionality to allow tracking and logging of outputs, and the impacts of their use.

---

[26] file:///C:/Users/VGYK3/Downloads/3.1-Internal-RFSDi-IOM-Evaluation-v1.0-1-1.pdf

[27] The briefing paper describes this topic as the 'Maximum probability of pertinent IMS logs being topic 1'.

Crime Seasonality Planner

The Crime Seasonality Planner uses time series modelling to output forecasts of monthly variations in reported crime demand for individual crime types across the Force and at an NPU (Neighbourhood Policing Units) level. It is used for resource planning, annual leave embargoes, timing of crime prevention public campaigns, planning partnership activities with other agencies, and minimising the impact of internal changes during periods of expected high demand. The tool is operational.

The dashboard for the Crime Seasonality Planner displays the model's **error rate**, along with a visualisation that conveys the error rate according to a **probability yardstick**, thus communicating information about the reliability and certainty of model outputs to the police user and potentially reducing the risk of overreliance on the outputs of a model (Goddard, Roudsari and Wyatt, 2012; Passi and Vorvoreanu, 2022). A screenshot of the visualisation is set out below.

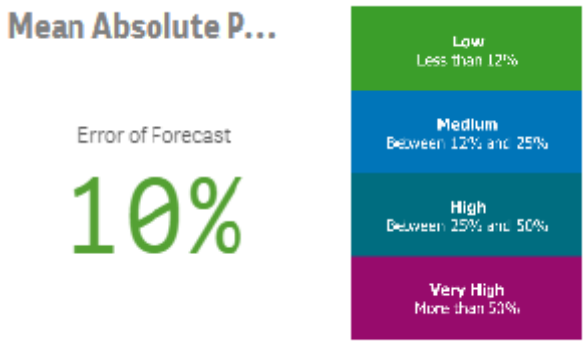

Figure 3. Screenshot from the Crime Seasonality Planner

The inclusion of information about the model's error rate and the probability yardstick is in line with recommendations made by Oswald that the output of AI models in policing should be treated akin to intelligence and graded accordingly. Reliability and certainty should, Oswald suggests, be measured by reference to a framework that should 'facilitate evaluation and critique by subject matter experts, with a model's errors and uncertainties highlighted' (Oswald, Chambers and Paul, 2023). The IOM dashboard demonstrates the feasibility of such an approach.

Improving understanding of ethical issues

Findings on this question reference three themes; first, whether in general the Committee contributes to knowledge development; second, whether the human rights discourse in the Committee is productive; and, finally, whether a structured framework for assessment of proportionality might contribute to the effectiveness of human rights discourse.

Knowledge Development

Various interviewees spoke of the Committee's role in delivering learning:

> *I did say … that I would be taking the learning from the West Mids group and applying it in in the development of our own Ethics Committee. So, it's been a kind of win, win, win for me.'* Ethics Committee representative 4

Interviewees also acknowledge the influence of the Committee in shaping their thinking and approach:

> *Obviously, we've been engaging with the Ethics Committee for a long time. So, when people come to us, we will kind of go, you know, we're not really prepared to stand in front of the Ethics Committee and put something to the Committee in those terms. You might want to think about how the Committee will respond to that.* Data Analytics Lab and Police representative 2

A police representative describes how the process of engagement with the Committee increased their knowledge and understanding:

> [speaking of the realisation that AI tools used in enforcement generate new data] *When I went to the ethics committee**, I actually had my eyes opened as a police officer** … There were some considerations that came our way for our review that originally, I didn't really understand, but as I got more into the project, I could absolutely see the relevance of why we were considering them*. Data Analytics Lab and Police representative 1

It is encouraging to see that the Committee has delivered insights and impacted on practice. No doubt it has contributed to an improved ethical understanding. However, the *extent* to which it has done so is not clear. In particular, questions arise about the extent to which the Committee's discourse is productive of a deeper understanding of ethical issues relating to human rights.

The Human Rights Discourse

One DAL representative suggests that over time, the exchange between the Committee and the DAL has had the effect of putting human rights considerations at the heart of the DAL's work:

> [explaining that the DAL now considers human rights issues] *Perhaps, you know, it's maybe … a bit of chicken and egg … Which came first, you know, is the lab only doing that because they know the Committee will look at it, or are they doing it because they want to do it and they think it's right to do and then there's an additional level of assurance … argue it either way, but … certainly if you … took a slice through it now rather than, you know, day one, you would see that kind of human rights consideration in in there.* Data Analytics Lab and Police representative 4

Nevertheless, some interviewees question the utility of human rights discourse as a means of resolving ethical questions. One interviewee describes the tests of necessity and proportionality as:

> *… defective in lots of different ways because necessity and proportionality are extremely difficult to define.* Ethics Committee representative 2

and criticises:

> *… a* ***ritual incantation of proportionality and necessity*** *which is derived from human rights instruments …*[28] Ethics Committee representative 2

Another interviewee comments that:

> *… even though I'm a big … fan of human rights law … I don't actually think they're that helpful for these discussions because they're so broad. So, you know the right to life. Well, you know,* ***the tool could be favoured and not favoured for that argument****. The right to a fair trial and all these things are so broad-brush. I just don't think they help narrow the discussion to the crux of issues enough. Anyone could level a broad-brush argument like*

[28] A search of the minutes finds 3 references to 'necessity and proportionality' and 2 references to 'necessity' (in the sense of the test for necessity in a human rights context).

that. And I don't think it'll get us to a place of agreement. Ethics Committee representative 6

These reservations about the utility of the necessity and proportionality test and the productiveness of human rights dialogue are significant. After all, law enforcement bodies have a duty to act in accordance with the law and in conformity with the human rights regime. As a result, WMP must carry out an analysis, for each project, as to whether its development and use is lawful. For every project where human rights are engaged, a proportionality assessment will be required. As one police representative explains:

> *… there is a multitude of ethical implications for every action and inaction that we do. So, in all of what we do, it should be a decision and that decision should be proportionate, necessary, accountable. Is it reasonable? Is it lawful? And is it the least intrusive approach that we can take? Which is our general overview for the… for the ECHR*. Data Analytics Lab and Police representative 6

If the Committee is to fulfil its role, it must be able to engage with and, where appropriate, critique that assessment. The Terms of Reference direct that the Committee should be solution-focused; it is to provide 'support in turning the outcome of theoretical moral debates into meaningful actions'. (West Midlands Police and Crime Commissioner, 2019) It is important therefore that the Committee's human rights discourse should not take the form of *'ritual incantation'* (Ethics Committee representative 2) but should have the effect of focusing issues and advancing debate. In the context of human rights assessments, many factors may contribute to **difficulties** either in reaching agreement or in focusing areas of dispute.[29] These may include:

- **lack of resource** - the Committee has much specialist expertise at its disposal including human rights expertise but members give their time on a voluntary basis
- **gaps in understanding** - analysis of the human rights implications of AI tools in a law enforcement context depends on an understanding of the uncertainties, inaccuracies or bias that may be introduced by these tools as well as knowledge of law and ethics. A holistic assessment is necessary

[29] Janjeva et al. (2023) note that making a structured proportionality assessment, that is, applying both the necessity and proportionality tests, is 'difficult in practice'.

- **gaps in information** – without information about why, how and for what purpose a particular tool is likely to be deployed, it will be hard to determine whether the tool is necessary, or what its impact is likely to be
- **uncertainty** on the part of those carrying out or critiquing a human rights assessment as to the nature of the human rights tests to be applied
- **difficulties in assessing levels of intrusion** on human rights – Janjeva et al. (2023), for example, point to four factors relevant to the use of AI-enabled analytics which raise additional considerations for the assessment of proportionality, namely, 'additional data processing, inscrutability of machine learning models, probabilistic outputs and the scale and retention of training data'
- **difficulties in the application of the proportionality test**, including lack of consensus about the nature and utility of the test, the fact sensitive nature of the test, lack of information needed to apply the test and the temptation to assume that if a policing purpose is established and/or the positive obligations under Articles 2 and 3 of the Convention are engaged, these trump other rights considerations.

No findings are made as to which of these factors, if any, impact on the productiveness of the Committee's human rights discourse but recall the comments by interviewees about lack of information about outcomes.

The merits of a structured framework for proportionality assessment

To 'supplement and futureproof' current approaches to proportionality assessment, Janjeva et al. (2023) offer a 'structured analytical framework for assessment of proportionality of privacy intrusion'. Toolkits exist for the application of necessity and proportionality tests in the context of European data protection law (European Data Protection Supervisor, 2017, 2019). Kennedy et al. (2020) offer a tool 'to evaluate proportionality and necessity in the use of restrictive practices in forensic mental health settings'. INTERPOL and UNICRI have developed a 'Toolkit for Responsible AI Innovation in Law Enforcement' (UNICRI and INTERPOL, 2024). The UNICRI/INTERPOL toolkit includes a risk assessment questionnaire which prompts users to address various risks including that of privacy intrusion. Unlike these instruments, the toolkit proposed by Janjeva at al. (2023) is specifically directed to assessment of the proportionality of automated analytics in a UK national security and law enforcement setting. Interviewees were invited to share their thoughts on the value, in general, of a structured framework in the evaluation of AI technologies developed and operationalised by West Midlands Police.

One Community representative was enthusiastic about a role for such a framework and considered that it would have a role in holding police authorities to account and in allowing those authorities to document their reasoning:

> [endorsing the use of a structured framework for proportionality analysis] '*... a hundred percent. [police authorities] wouldn't be able to just not [carry out an analysis] ... and then they wouldn't have any reason, or backup, or substance to why they're doing it. So, because it would be violating human rights, obviously there's a legal side to that, so you could have a complaint on that side but I suppose if you're following a framework, and you've got evidence, etc, as to why you've used it, I think ... that's how it would work best.*' Community representative 9[30]

Some Committee representatives considered that structured frameworks might be '*educational*' (Ethics Committee representative 5) and could form '***the foundation stones***' for ethical review (Ethics Committee representative 4). However, there were some reservations. One noted that some structured frameworks:

> *can be quite **tick boxy** and ... sometimes guide people away from some of the discussions I think need to be happening, again, like the policy outcome*. Ethics Committee representative 6

Another reflected on the implications for the volume of material that Committee members might have to review:

> *... it's a balance ... if you make the briefing pack too big the committee is never going to have time to read it before the meetings*. Ethics Committee representative 4

[30] The comment is insightful. Note that, faced with an argument that a public authority has infringed the Human Rights Act 1998 or a Convention right, the question for consideration by the court is 'whether there has actually been a violation of the applicant's ... rights and not whether the decision-maker properly considered the question of whether his rights would be violated or not.' *Belfast City Council v Miss Behavin' Ltd* [2007] UKHL 19; [2007] 1 WLR 1420 at paras 13-15. Nevertheless, although the court is the final arbiter on proportionality it may 'have regard to and may afford a measure of respect to the balance of rights and interests struck by a public authority such as the police in assessing whether the test at stage (iv) is satisfied'. *Director of Public Prosecutions v Ziegler and Others* [2021] UKSC 23; [2022] AC 408 at para 130. See also *Dalston Projects Ltd* (n 13).

Summary

The research finds that the Committee, in broad terms, fulfils its remit, has a clear focus on human rights, impacts on the development of projects, and delivers learning and insights. Nevertheless, the research highlighted potential obstacles to the effectiveness of the Committee, including a lack of diversity in membership, inconsistent communication about the purpose and likely outcomes of projects, and the absence of clear protocols for review of operational projects. Concerns were noted about the productiveness of the Committee's human rights discourse, that is, whether that discourse serves to focus or resolve debate. There are obstacles to the evaluation of the committee's effectiveness, in particular **the scarcity of explicit references to human rights in the minutes and the lack of (publicly available) records documenting, for each project, whether the Committee's recommendations were followed**. The comments from interviewees concerning the utility of a structured framework for evaluation, though useful, are in general terms; further research is necessary.

3.2 Community engagement

The following section examines the theme of community engagement. It aims to understand the community's perceptions of the Committee within the policing context, how the community may interact with it, and barriers to inclusion. This analysis is divided across three sub-themes, each examined separately, but all contributing to the overall process of community engagement. [31]

Accessibility and definition

The role of the community and community representatives within scrutiny panels and Independent Advisory Groups (IAGs) is characterised as one of a 'critical friend' – a supportive entity that can offer constructive criticism and insights in set cases (Dixon, 2018). An example of this is Community representatives' reviewing 'Stop and Search'[32] videos as part of their duties on scrutiny panels and providing feedback on police practices:

> *You know, with the stop and search, being connected to young people and stuff, I just wanted to be a part of that, and be involved, and give feedback of good practice, and not good practice with young people.* Community representative 9

Community representatives saw IAGs and Scrutiny Panels as a means of engaging proactively with the police and creating positive change:

> *we don't have to look at stop and search to see how there is a lot of mistrust with those sorts of powers and also, when it comes to policing of protests, things like that. So, we need an open and honest policy and dialogue to move forward*. Community representative 4

This is supported in literature, with both communities and police stakeholders recognising that greater citizen cooperation in policing procedures leads to greater trust and confidence (Hickman, 2009).

---

[31] Note that references to decisions of the Committee in this section relate to internal decision-making processes of the Committee, as the Committee's recommendations are advisory only.

[32] 'Stop and Search' refers to the search powers of an officer who has reasonable grounds of suspicion that an unlawful item is being carried (College of Policing, 2016)

Community representatives were generally unaware of the existence of the Committee however.

Those who were aware of the existence of the Committee were uncertain of its role:

> *I think I've heard mention of it before, but not any nothing of any major depth, just the fact that it does exist. But not really what it does or where it is or who's involved.* Community representative 2

The uncertainty of Community representatives with regards to the Committee's purpose and activities can create an additional barrier, as community members have limited time and resources to engage with policing mechanisms, thus choosing to dedicate time to panels and boards where the outcomes and potential impacts are direct. Community members reflected on the issues with involvement in committees:

> *Is this been done on a freebie? I'm using that language because there's there are many, many situations where the NHS or the City Council or the public director of public health, they're expecting people to come forward, volunteer to do some work and it's really a cost saving exercise because previously they used to employ people to do this.* Community representative 7

It is worth noting that the Committee may be competing with other policing panels for Community representatives. The first area of consideration for the Committee, therefore, is **the clear communication of what it is seeking to achieve, and the potential impact it could have on local communities**. Currently, the Committee conducts these activities through its online presence, for example, the publication of its minutes and associated documents. Although the Committee minutes and associated documents are available online, Community representatives were unaware of these:

> [Regarding Committee meetings minutes being online] *I'd be surprised if you'd spoken to anybody who's read those.* Community representative 1

> *I found some information online, but I have had to go delving for it. They have not been forthcoming, as to what they are proposing to do.* Community representative 6

Developing a welcoming and meaningful environment

When Community representatives engage with the Committee in the future, there will be a need to ensure the promotion of an inclusive and welcoming atmosphere to overcome barriers of distrust within the wider ecosystem. For example, Community representatives noted concerns over existing panels being tokenistic or unwelcoming:

> [Discussing previous experiences] *They want to have meetings where they turn up suited and booted, make nice comfortable kind of noises and no change happens because they don't want change to happen*. Community representative 5

> *'It's interesting. When I started about twenty years ago, custody visiting, you often came across custody sergeants who were very resentful of somebody's coming in and so forth. Not all of them, but there was a significant minority, and they could be quite awkward. And you found yourself either being sort of quietly bullied or having to be quite assertive. And obviously, not everyone's comfortable with that*. Community representative 3

The challenges in creating meaningful environments can be overcome, with Community representatives sharing examples where a positive environment has been created for the collaboration between policing stakeholders and the community. In each case, **the goal was to ensure that the Community representative was overtly valued, and that efforts were taken to diminish negative power dynamics**:

> *If you make polite criticisms or point out issues, they're not difficult in the least. They take note of them and try and do something about them, so they're very responsive. It's been a big turnaround. The same with the committee or the panel I sit on. The inspector is amazingly open*. Community representative 3

> *When they realise that they can speak to those within policing and it's almost like you can leave titles at the door when you go in, you talk to each other as individuals to be able to scrutinise, to push individuals, to really get some answers, especially if they need closure for incidents they've been involved in themselves that can do a lot to really build, amend bridges because they understand why something's happened. It wasn't just that the police couldn't be bothered investigating*. Community representative 4

Averting tensions with community groups

This pressure was also noted by Committee members who mentioned there could be tensions within meetings which need to be balanced between the community and the police:

> *The role of the committee or any Ethics Committee is about helping the force navigate these kinds of choppy waters*. Ethics Committee representative 4

The Committee can create and strengthen the relationships between community groups, who are collectively impacted by DAL projects, and the police (Brugge and Kole, 2003). Introducing a participatory-style Committee, that places value on community voices, can develop stronger and mutually advantageous relationships between the DAL, police and community partners (Cross, Pickering, and Hickey, 2014). Indeed, it was noted that a positive environment could be created by ensuring **Community representatives could meet Committee members prior to their official duties to help develop internal relationships** (Community representative 5).

The first area for the Committee to consider is how it positions itself within the policing and community ecosystem to ensure Community representatives can identify the opportunities of engaging with the Committee, and the associated benefits for Community representatives who have competing demands on time. Achieving this requires **clearer communication of the Committee's objectives and the impact of the Committee work on the community**, as well as the creation of an inclusive environment in the Committee. If this is accomplished, the Committee may be able to encourage meaningful community representation. However, this is a singular part of a circular process, with it being important for the role of the Community representative to be appropriately positioned within the Committee.

Capacity and influence

Capacity refers to the ability (or the perceived ability) of individuals to make a meaningful contribution to the decision-making processes that affect them. Understanding the capacity of individuals and/or groups requires a clear description of the definition of the group. It may be linked to the ability to offer 'mature opinions' based on an in-depth knowledge of a subject or the ability to contribute to decision making processes that align with the political-institutional context, for example, the adopted and accepted policing strategy (Lundy, 2007).

Utilising community knowledge

Recent literature has positioned capacity as something that should be self-defined by individuals and/or groups, ensuring that individuals and/or groups can engage with decision making bodies or processes of justice on their own terms (Mkwananzi et al., 2021). **This is important in areas where vulnerable, or other participating groups, may experience knowledge gaps in the technical aspects of the processes that may have direct impact on them, such as the use of AI**. In these cases, it is important to ensure that how individuals and/or groups express themselves is valued by committee members and is fed effectively into the decision-making process. For the Community representatives within the research, **the value of their contribution was the introduction of real 'lived experiences'** that may be beneficial for other members of the body that they participate in:

> *Whether it be funders or people asking you, nowadays, the common question is 'how is your work informed by lived experience?' So, as opposed to [someone] coming up with an idea about doing something, what everybody likes to see, nowadays, is how's that been informed by lived experience? For those survivors who've been through this, to say, you should be doing it like this, or have you thought of this, or have you thought of the consequences on victims or something?* Community representative 1

Community representatives (none of whom has any involvement in the work of the Committee) speculated that their lack of knowledge about technical aspects of AI would be relevant to what they could contribute to the work of the Committee. **Such lack of knowledge was characterised as a strength**, however; community representatives could approach projects from the perspective of the community, with limited knowledge of AI, but concerns about how it is to be used:

> *Certainly, some kind of lay scrutiny would be a useful tool…. a lot of the AI stuff is… close to being out of the reach of my intellect, to be honest with you, when you're getting to the- the actual bones of the AI. The way things are working and why, and where it sends the data, but obviously it's understandable by most people, however, this is another step down*. Community representative 2

> *Given that most of the Committee members aren't grounded in the communities of the West Midlands, you are going to miss some opportunities for nuance, but I think there's a* ***fairly good appreciation of the diversity that there is in the West Midlands****, which*

means that you do get specific discussions about specific communities and the impact on them. Ethics Committee representative 6

Both the DAL and Committee representatives supported this sentiment, noting the Committee was more effective when a collection of voices and perspectives were shared:

I'm also a believer that diversity isn't just good for its own sake, but there's a reason we have lots of instruments in a symphony. That different people make different noises, and that a trumpet can't do what a harp can do, and there's always a missing perception that might be the crucial one. I'm also conscious of that there are lay opinions out there that would be really beneficial, and that would perhaps even inform the, quote, unquote, 'expert' opinions in a slightly different way, and bring out slightly different inflections. So, do I think the committee tries really hard to think about how the public thinks about these things? Yes, but by definition, they can't do that for everyone, and are there voices missing? Yes, would be my answer. Ethics Committee representative 10

Everybody comes to the group for a different reason, so some people are there because they're considered to be a little bit of a subject matter expert on ethics data, ethics, and policing. [...] Other people have been brought into the group because they are local people and they are effectively lay members who don't have any specific subject matter knowledge but are able to [...] they all bring some experience and knowledge of something relevant to the the group by way of their life experience or other professional experience. Ethics Committee representative 4

Inclusion of community perspectives

The inclusion of viewpoints from the community within the Committee itself was recognised as valuable by the DAL and Police representatives. They reflected on the challenges of including community viewpoints in the development of their project, noting **the Committee may be a useful conduit in capturing these *missed* viewpoints** (Data Analytics Lab and Police representative 1). The capacity and value of the Community representatives on the Committee is, therefore, in raising lay persons' concerns:

If somebody knows a little bit about something, they've already got an opinion probably [and] is going to be the most worried. And walking down the street thinking well 'What's this AI thing, I don't know if somebody watching me all the time, and if I drop some litter

on the floor, I'm going to get a policeman knocking on. the door when I get home. I don't know what how it works. I'm- I'm scared of it'. That's probably the thing, they have a fear because they know nothing about it. So, [the Community representative] would be able to offer the view of someone that fears already knows a little bit about it. Community representative 2

Committee power dynamics

Research on ethics committees in medical science (Schuppli and Fraser, 2007) illustrates that **community members on ethics committees can feel 'outnumbered' in committees where most other members have expertise in specific knowledge areas, leading to disempowerment**. One way of addressing this issue would be for other members of the Committee to, essentially, act as intermediaries between the DAL and these Community representatives. This would involve understanding the underlying logics of any concerns, contextualising them, and utilising their own knowledge of the wider institutional or technological environment to feed them into their recommendations (Communi]ty representative 7).

This, however, is not a simple process to embed into an environment where the majority of those within the Committee have specialised skills that lay-members may not possess. Interviewees recognised there was a need for specialists to be involved in the Committee, and the benefits of individual perspectives to be supported or challenged:

We need that intersectional view where everyone has, I would say, a base level of understanding and then we have people who are a bit more specialist in those areas because then that way it's going to be a much more rich, more rich conversation. Community representatives 4

It's about working together in partnership to make sure that what you're intending to do as a force is actually considered in terms of the community that you are proposing to kind of impact on the work. Ethics Committee representative 4

This last quote also highlights the need to create a space that reinforces the position of people on the Committee so they can share ideas freely and do not feel 'exposed' when unable to contribute. **A concern of Community representatives was associated with the feeling of not knowing what was happening around them**, being alienated if discussion focused on technological aspects of the decision, or crowded out of conversation:

> *Everybody that's within the system, they're all comfortable with it because they know how it's working, whereas the [average] person knows absolutely nothing. They are going to be very uncomfortable with it because they know nothing about it.* Community representative 2

Influencing the Committee

It is important that once their opinions are heard, the Community representatives can understand and see the impact or influence they have had on decisions, and not only seeing the impact of their contributions, but also ensuring that the Committee is operating as expected. This is particularly important in previous cases where Community representatives have felt that their input has had little impact on the process:

> *Unfortunately, it happens in some way, they are a non-statutory body. So, how much influence or interest that the senior police take of our views is up to the chief constable and so far, they have largely ignored what we have to say.* Community representative 6

This can be achieved through collaboration across the Committee, with members initiating processes to create an environment that encourages high levels of engagement from all participants, ensuring that each member feels that they have contributed:

> *normally every member of the committee is really quite well engaged in the discussion, which leads me to think actually that they have managed to glean enough about what it is that they've been presented with to be able to offer views and opinion, which ultimately translates into the advice that the committee gives.* Ethics Committee representative 4

Once Community representatives join the Committee, the first step should be to clearly establish their role, whilst demonstrating how they contribute to the decision-making process, legitimising their position and developing their confidence in participation. Although the Community representatives held concerns that their lack of knowledge may hinder their participation in group discussion, findings from the research show that stakeholders do not regard this as a problem and, instead, view the value of Community representatives in their unique perspectives, rather than their technical expertise. It is important, however, that Community representatives **are able to see the influence and impact of their contributions**. Processes of confidence building may also be important to developing the Committee, ensuring that Community representatives receive the knowledge requires to engage meaningfully.

Dissemination and development

With Community representatives being able to access the Committee and become active and full participating partners in the decision-making processes, **there exist opportunities to utilise them as both advocates for the community, as well as develop community knowledge of AI**. As noted in 'Accessibility and Transparency', Community representatives noted distance between the Committee and the community:

> *I'd expect the Ethics Group at the moment is, I would call them quite [...] high-brow. You have time to bridge the gap between the Ethics Committee and the general public, who don't understand what's going on.* Community representative 6

With a limited understanding of the Committee's work and the wider implications of AI technology in the community, interviewees reflected on the need for the Committee to share knowledge and information with the communities it wants to engage with on what AI is and how it can be used (e.g. Community representative 8).

Methods of community engagement

It was noted by stakeholders, however, **that the development of spaces for knowledge sharing can be difficult, with questions on who should lead knowledge sharing activities and what skillsets may be required along with resource barriers**:

> *I know another force [..] They did do some workshops with young people around a model that they had built and they had very positive responses from those kids. But I'm not sure how they were selected, and the people presenting to them were the people who built the model. I'm not sure I'm personally skilled or experienced enough to know how to correctly conduct those kind of sessions and whether it should be the people who built the model who explain it, or propose it, or advertise it to the marginalised group. I mean, it's such an area of specialised expertise. I think that's something that we don't have the skill in the lab to do.* – Data Analytics Lab and Police representative 2

> *The all-singing and dancing version would be quite a wide scale public engagement, which just hasn't been plausible with resource and time at the moment.* Ethics Committee representative 5

Opportunities for engaging with the community were discussed by Community representatives who believe that members of the Committee could enter the community, delivering information and/or training sessions in-person to promote community engagement with the Committee. **Engagement with the community should be twin tailed, promoting valuable engagement of Community representatives in the Committee, and vice-versa**.

Another benefit of engagement with the community was the opportunity it would create for the wider policing ecosystem to talk about the issues they are facing, and the reasons for the implementation of AI into policing strategy. **The purpose of this is not to come to an agreement, but rather increase the community's knowledge of what is hoped to be achieved using AI**, as well as reinforcing the Committee's role:

> *I would say have courses or training or- or workshops in the community, by just explaining how AI works and all this, instead of actually giving a voice to say yes, I do, or I don't agree. It's not about agreeing, we are not out on the front line, they have no idea what you guys are dealing with*. Community representative 8

Contextualising the Committee

Engaging more fully with the community would have the added benefit of encouraging community members to engage proactively with the Committee and encourage other representatives to join. Additionally, this ties into the notion that the **Committee may benefit from a more explicit contextualisation of its role, that separates it from other community policing bodies**:

> *That could be a first step towards them wanting to get involved, more wanting to join the committee eventually or be part of an IAG or anything like that*. Community representative 4

One of the goals of this type of engagement is to turn Community representatives into community advocates, who feel empowered to take the work of the Committee to the community:

> *It's not a simply a case of the police saying this is what we're going to do... There needs to be dialogue. There needs to be an exchange, there needs to be a chance for those community and group leaders to take what they've heard or what what's proposed back to their groups for them to really have time to digest it, to think of any considerations to them, bringing them back to the table so that so that as these sorts of things develop, they*

develop with the community moving with them moving forward. Community representative 4

People vote with their feet, so you have to try and make sure that you are investing in the community in a way that is appreciated and that they can hear that their voices are heard within that committee so that you have a committed membership who are turning up and learning and keeping up to date and also developing an expertise about it in terms of every discussion you have is another bit of learning. Community representative 5

Empowering the wider community and developing opportunities for engagement may also create opportunities to close intersectionality gaps (Walsh et al., 2024). Creating space for Lab-Police-Committee-Community engagement promotes trust which can be eroded in environments with 'governance creep' (Gorman, 2011; Walsh et al., 2024). Community representatives and Ethics Committee representatives noted that although they maintain leadership positions within their community, they did not necessarily represent all those voices:

Everybody's lived experience is different. Community representative 4

We have had slightly more diverse representation on the committee before than we do currently and we have put resource into trying to recruit more diverse members which we've struggled to do. If we had more time, I think we would literally be able to approach particular individuals, but some of that would also be about being perceived to be doing the right thing rather than ensuring that the voice is actually being heard, if I'm being truthful. So, we're not there yet. Ethics Committee representative 6

The final stage in creating effective community representation is the dissemination of information and development of community knowledge. **This acts to both reinforce the accessibility and contextualisation of the Committee's goals, while building a knowledge base within the community that allows them to more effectively feed-in to the Committee**. Direct engagement can also allow more voices to be captured, and for future Community representatives to engage directly with the community and the police.

Summary

Drawing on the interviews, three steps have been identified in developing a mutually reinforcing cycle of community engagement (Figure 4). Beginning with Accessibility and Definition, **it is**

important for the Committee to clearly define its role within the policing ecosystem, distinguishing itself from entities such as Independent Advisory Groups and Scrutiny Panels, while emphasising the significant and positive impact it can have for the community. Currently, most interactions that interviewees had with the police were that of 'critical friend', whereas the interviews suggest that the Committee aims for Community representatives to act as both advocates and advisors, as anticipated in the Terms of Reference. Once this has been established, **it is important to create an open and welcoming environment for Community representatives that engages with their knowledge of social context and ability to contextualise the potential impacts, outside of the AI pipeline** (Ovalle, 2023). The Committee **may learn from the experiences Community representatives have had in other organisations, to allay concerns of potential tokenism**. This can be fostered in informal settings, or by evolving the format of the formal meetings.

Once included within the Committee, considerations of Capacity and Influence become crucial. Community representatives should be encouraged to engage with the discussions according to their capacity. **Community representatives voiced concerns that a lack of knowledge about AI might limit their contributions or have them sidelined during wider conversations**. However, this concern was not shared by other stakeholder groups who viewed the primary value of the Community representatives, **not in technical expertise, but instead sharing lived experiences** which allow them to make suggestions that may not be apparent to the other Committee members. This requires a role change for other Committee members, who should consider acting as 'translators', by contextualising community experiences within the institutional or technological context. It is essential, however, that regardless of whether these experiences impact on the Committee's advice and recommendations, **Community representatives are made aware of their influence and informed how their contribution was utilised**.

For the final step, the Committee should focus on Dissemination and Development. This involves promoting its work within the community and **using these promotional engagement opportunities to build the knowledge of those it seeks to represent**. This can be achieved through workshops, training seminars, or community meetings, with each aimed at directly engaging with the community and increasing their unde]rstanding of the use of AI technology in policing. By doing so, **the Committee can reduce power distance between itself and the community and allow alternative voices to be heard** which are not directly represented on the Committee. There are suggestions that the Community representatives themselves could help

lead these initiatives, turning them into advocates, making the Committee accessible to new members, whilst building future capacity.

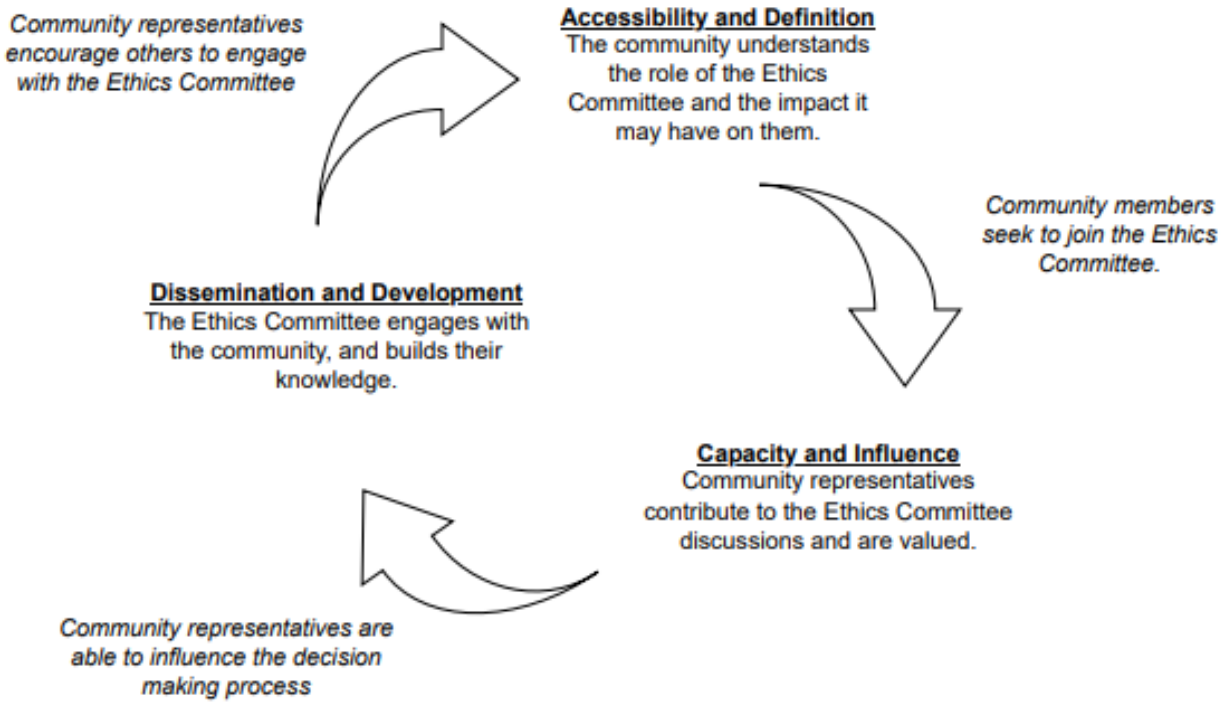

Figure 4. Illustration of the mutually reinforcing cycle of community engagement on the Data Ethics Committee.

3.3 Ecosystem communication

One emerging problem when considering engagement with AI technologies, is **the perception that the topic is 'opaque, complex, and requires a high degree of numeracy'**, which is beyond the general skill level of not only the public, but also researchers and scholars (Burrell, 2016; Jenkins, and Purves; Sieber et al., 2024, p.634). Within the policing ecosystem, understanding the technical aspects of an AI project is fundamental to understanding how it may be deployed, the potential impacts, and influence on broader police strategy. **However, these discussions must be balanced with the need to move beyond the technical details of a project to include legal, ethical, safeguarding, and social considerations**. It is, therefore, necessary to include both specialists (those with a deep understanding of a specific topic) and generalists (those with a broad knowledge of many topics) within the Committee. The following section examines communication between the Committee, DAL, and Policing stakeholders, across the impact of technical language, how projects are tracked, and internal tensions between the parties. These themes are united in the summary.

Technical language

Technical language can act as a barrier to participation by Community representatives (See: Community Engagement) and hinder effective communication among other Committee members, some of whom may rely on specialists within the Committee to interpret complex details. Committee and DAL and Police representatives noted that whilst technical language is necessary, **the diverse level of understanding and capacity within the Committee can slow down discussions or force the simplification of language**:

> *If you put one of the data scientists in the room with an operational police officer, they won't understand each other, and that's one of the biggest challenges around the communication. The same would be said between the lab and the Committee if you didn't manage that communication effectively.* Data Analytics Lab and Police representative 4

> *'Those on the Committee, their communication skills are exceedingly good and that I think is to do with shrewd appointments [...] I think with that communication, it is shifting registers, being aware if you [are] using technical language, and pause to explain it [...] so you're not using something incredibly technical.* Ethics Committee representative 1

The need for 'balancing' creates tensions in the Committee, with Committee members acknowledging the need for those with knowledge of the complexities of 'quantitative' approaches:

> *It wasn't balanced in the sense that we did need those new people, either public health background or statistical or mathematical background, because we needed more people to understand the complexities of the quantitative side of the evidence that would be put to us as part of these paper proposals. The lab did a very good job of explaining it to us as best they could, those of us without a statistical or mathematical knowledge base but we just needed that help as a committee. We needed a broader membership and we got it.* Ethics Committee representative 3

> *I think the only thing that has been a little bit problematic on occasion is that sometimes the papers that come through are very technical and complex. I think particularly some of the ones which focus on the data science models in relation to sort of specific project under discussion, they can be really, really complicated to understand. If I remember correctly, asked on occasion for those documents to be simplified to enable members of the committee to be able to understand them.* Ethics Committee representative 4

Creating opportunities to learn

A solution to challenges in interpreting and/or understanding technical language was considered by a Committee representative, who **believed there were limited opportunities offered for them to improve their knowledge**, with information being limited to Committee meetings, or training workshops that were held rarely:

> *We have had a couple of sessions where we as a Committee has been given training of some kind in relation to data protection, for example, or other relevant things [...] but that there doesn't tend to be anything outside of the meetings themselves or the papers that are issued apart from on a couple of occasions. I think over the years we've had like CPD sessions, where we've been given the opportunity to develop our knowledge in a particular area but, the contact is limited to the meeting itself and to the papers that are sent before and after.* Ethics Committee representative 4

The impact of the lack of intricate knowledge of data and AI has been limited, however, as the Committee has evolved, and as more knowledgeable and specialised members have joined.

Although the impact is limited by evolving membership, **there are inherent risks in relying on specific Committee members for technical expertise**. Such reliance can hinder the motivation and inclination of other members to learn the technical side of the AI discussion:

> *there was a time when we didn't have good communication about technical matters and attempts were made to try and explain some of the data science to Committee members without a data science background or indeed without a statistical background. Those efforts were much appreciated but didn't ultimately work very well. I'm talking about the process of the meetings in 2019-2020 – 'can you explain that to us in a way that we might understand?' - And attempts would be made and we wouldn't understand them [...] So now I think the culture of the meetings, the tendency of, if you like of the meetings has moved to with the recruit of more statisticians, data scientists, human geographers and economists, I think, I think we've now got to a point where I don't need to understand some of the technical stuff anymore*. – Ethics Committee representative 3

Within the broader ethics committee context, training opportunities have been used to improve skills related to ethical review, through knowledge exchange, open discussion, or case studies (both theoretical and real) (Cash et al., 2009). For example, the Health and Research Association (HRA) ethics committee has requirements for ethics committee members to participate in training and development, with the expectation that REC members complete 5 hours of training per year (HRA, 2023). Consideration should be given to the role of technical training, and whether there would be appetite and time allocation for developing knowledge in this area for the full committee, when technical experts are already members – **acknowledging the time and resources available to Committee members who are performing a voluntary role, and that the Committee performs an advisory role only**.

Currently, the balance between specialists and generalists has been perceived positively by DAL and Police representatives, who believe that the ethics review process is becoming stronger, although **there is potential reliance on certain individuals to clarify statements or explain areas to other members**:

> *the expertise that's on the committee is pretty substantive. So, where there is technical language used, I don't think I've ever had to say, 'what are you talking about'? because usually there's been someone that's been able to either clarify it, or to be fair to the lab, they clarify it, and they're cognisant of their language. But, for instance, when you've got*

folks like [name] and [name], who know this landscape like the back of their hand, they're able to really communicate it effectively. Ethics Committee representative 10

Tracking projects and outcomes

Interviewees reflected on communication issues affecting the Committee's ability to track how WMP acts on their recommendations, obtain feedback about the outcomes of 'live' (operational) projects and deliver advice on such projects. **Existing processes, including processes for tracking how the DAL responds to the Committee's advice and recommendations, were seen as deficient**:

That's probably one of the fundamental areas that the committee still has work to evolve on. It's very much the intention that projects come and inform us how ethics has informed them. That's literally written into the fabric of it. Some projects do return and explain how they've dealt with things... there's an expectation it comes back. And we track - you'll have seen the trackers and stuff, and obviously some of those are imperfect despite best efforts. Ethics Committee representative 6

I think the audit trail, and the feedback is lacking in this process. Ethics Committee representative 8

I think that's probably where I would like more information, or the committee doesn't get enough information. [...] Once we have sign-off on a project, I don't think we really ever go back and see, 'so we signed off on this, how was this used? What were the outcomes? Do the outcomes show any evidence that the wrong people are being targeted?' To put it very, very crudely. I don't know that we ever go back and look at them and maybe that is something that we should do. Ethics Committee representative 9

One Committee representative suggested that **retrospective analyses** of approved projects would benefit the decision-making processes and help guide the future work of the Committee:

We do probably lack a kind of retrospective analysis of the impact of all these projects, depending on how many were actually, finally, operationalised or how many were actually not operationalised, or some just might have fallen by the wayside because there were too many worries about it [...] I think a retrospective analysis would be helpful in- in trying

to guide the future work of the committee and direct interaction with the communities. Ethics Committee representative 9

A DAL and Police representative alluded to the fact that returns to the Committee are mostly conducted through direct requests or ad-hoc updates:

From the point of view of general data science project management, you would keep an ongoing check as to model degradation and things like that. ....There might be updates every now and then that are provided, they might ask about certain projects ...but there are no specific requirements, as far as I remember, in terms of going back once something's been through the process Data Analytics Lab and Police representative 3

Creating an audit trail

Processes for tracking how the DAL responds to the Committee's recommendations, and **how feedback about 'live' projects should be provided should be clearly set out in the Terms of Reference**. Consideration should also be given to how and when 'live' projects should be returned to the Committee for consideration, and any learning from the successful deployment of projects fed back**. Formalisation of this process would allow for the creation of an audit trail** that maintains the tracking of projects, with specific dates or other 'triggers' for projects to return to the Committee once operational. The introduction of this process may strengthen the inherent trust that the Committee operates under (Hedgecoe, 2012):

You do need to have this process whereby the model performance is fed back; we need to know how these models are performing and what the results are. [...] we'd like to get a twelve-month update of any of the operation, any of the model's that are operational, but I said it's happening on an ad hoc kind of basis. Ethics Committee representative 7

When some projects don't really progress at all, there's the tiny chance they're progressing and no one's telling us. But we do ask about the projects, so I'm hoping that's not the case but I wouldn't want to put my house on it. Ethics Committee representative 6

The formalised process should also address how opinions and outcomes are communicated to the DAL. Currently, DAL and Police representatives view the communication of the Committee's recommendations and advice as weak as explained below.

Communicating decisions and advice

The Committee's deliberations take place behind closed doors (but with published minutes and papers) to ensure that members can discuss ethical issues openly:

> *It's a discussion amongst a closed group. We'll bring our own thoughts, views, our own individual biases to the to the discussion, but we do our best, I think to kind of try and tease out all of the arguments and everything that we think should be taken into account and make sure that that's reflected in the advice that they we give.* Ethics Committee representative 4

From the DAL's perspective therefore, effective communication may be dependent on the quality of the minutes:

> *Sometimes, there can be frustration if something has been misinterpreted or misunderstood. We always have the opportunity to make some amendments to the minutes if we feel there's an error in them, so, we obviously go through the minutes. And where there are particular points that have been raised, then we always attempt the next time we come back to ensure that we've shown that we've considered them and made changes if we think it's appropriate to do so, and if we're able to.* Data Analytics Lab and Police representative 2

> *you might be able to read between the lines in terms of the concerns that have been noted in the minutes, so it can be dependent on how good the minutes are. Because obviously we're not party to the discussion that the Ethics Committee has about the project, so that tends to come out in the minutes and that's what we would rely on to get any sort of background context on the minutes. There have been occasions, rarely, where we've thought a concern was 'A' and it transpired it was actually 'B', because the minutes didn't bring out that nuance.* Data Analytics Lab and Police representative 3

Interviewees also suggested that the Committee's recommendations can be framed in such a way as to make it difficult for the DAL to understand the main thrust of the concerns:

> *[The] recommendations lack explanations [...] as in 'the ethical issues that we think you might want to consider are...' as opposed to '[a choice between] outcomes A to E'.* Data Analytics Lab and Police representative 3

The Terms of Reference set out the options for the recommendations that can be made by the Committee. These are in specified terms and are sometimes provided without accompanying explanation or summary of the Committee's concerns. **One possible solution is to operate a process that allows DAL members to return at the end of Committee meeting for an overview of the recommendations**. A benefit of promoting this qualitative and reflective discussion is that it may help create an environment that enables trust and open dialogue (Walsh et al., 2024) and break the cycle of iterative discussions between the Committee and the DAL.

This section highlights the need for processes and procedures for tracking whether and how the DAL implements the Committee's recommendations, for projects to return to the Committee, and for the learnings from successful projects to be integrated into the decision-making system. Implementation of such processes and procedures would aid transparency and enhance the Committee's ability to scrutinise projects. The section also highlights **the need for better communication of the Committee's recommendations and advice**. The formulaic nature of the recommendations and the lack of explanations or rationales for the advice and recommendations, present obstacles to effective communication and may contribute to misunderstandings. Consideration should be given to a process where the DAL have the chance to rejoin a meeting after the 'closed' discussions to engage directly with the Committee on recommendations. **The recommendations should also be supplemented with a brief summary of the rationales for the recommendations**.

Internal tensions

The final aspect of the communication theme is based on inherent tensions within the ethics process, **the friction that can occur between the two parties**, and competing visions of the Committee's role. The DAL Lab and Police representatives and the Committee representatives, recognised that the tensions between them were to be expected:

> *we understand where the committee is coming from, but we don't always agree with it, or we feel that it's misplaced, or missed the point, or not understood the operating environment within policing. We're clear that we're operating in this framework and people come with their own personal opinions about policing. But this is the environment that we're allowed to operate in*. Data Analytics Lab and Police representative 2

> *a system with no rubbing points, is no system at all. They're supposed to be rubbing. There is supposed to be friction and we are supposed to resolve those frictions and those*

misunderstandings which they very often are, or disagreements in a civilised fashion. Ethics Committee representative 1

There may, however, be instances where **the DAL felt that the relationship could be adversarial,** although this was not necessarily regarded as an issue**:**

It can feel a very challenging meeting for us, and it is something we kind of psych ourselves up for. It can feel quite adversarial, sometimes. And I don't have a problem with that. The ***purpose of the Committee is to scrutinise and question what we do*** *and not to be our friends*. Data Analytics Lab and Police representative 2

I think they've lost their way in terms of what their role is and therefore it's unnecessarily adversarial. Certainly, in the last one that I have been involved in personally, I had to go back and rewrite or not accept the outcomes from it because the language used was inaccurate. Data Analytics Lab and Police representative 6

The last quote speaks to an area of consideration, with **DAL and Police representatives feeling that the Committee has crossed into the operational independence of policing**. As illustrated by the following quotes, the contrary view is taken by Committee members who believe that operational aspects of the model can be regarded as relevant to ethical concerns and that what is useful may not be known until given voice:

there are times when the Committee stray into that bit around the operational independence of policing. I think there are times when they've been asked to look at the way in which a model has been developed and they've questioned whether policing should be doing this in the first place. Well, the chief has already decided that the chief needs to do this. That's the operational independence of the chief. The question is how best to go about it [...] There was a discussion at some point around the use of drones, and it turned into conversation 'well, I'm not sure police should be using drones'. Well, that kind of decision has already been made and that's not for the Ethics Committee to determine. It was around a specific model involving the use of drones. Data Analytics Lab and Police representative 4

The Committee is completely and utterly entitled to make any comment it wants and observation it wants, and that's important. Firstly, they might have something incredibly useful to say, but also it might artificially limit what they do say if they're thinking that

> *there's parameters about what they're allowed to talk about, and sometimes commenting on the law is a grey issue anyway, not because the law is grey, it might be ambiguous as to whether your comment really is a legal one or not*. Ethics Committee representative 6

The DAL and Police representatives acknowledged the merit in Committee comments but felt the practicality and achievability of recommendations requires a careful balance:

> *Sometimes, they're [advice and recommendations] a bit more of a wish list on the part of the committee. So quite often, they'll say 'this would be much better if you were able to use non-police data as well as police data'. Yeah, we know, but we haven't got non-police data. We've got police data. So sometimes, their recommendations [are] not achievable for us. So, it's a point of view maybe rather than a recommendation. I appreciate why they need to say it because it's not an incorrect comment*. Data Analytics Lab and Police representative 2

The role of the Committee

Much comment in literature sees ethical review as a dynamic process, aligning with members' reflections on the importance of open dialogue, and the broad contextualisation of the presented ethics issues (Schopper et al., 2015). Members of the DAL did note that the Committee has an important role within the policing ecosystem. **One interviewee believed the focus should be on holding the business to account, rather than examining the technical aspects or policing strategy**:

> *I kind of want to move the ethics panel's conversation away from the technical. There'll be a space for that of course. But they should be holding the business lead to account. 'Why is it you want to have that data set?', 'What are you going to do with it?', 'How are you going…', 'how are you going to justify balancing the rights of this group against the rights of this group?', 'Or the unintended consequences as it affects this group'. So, I try to not only get a better way of prioritising the work, but actually move the ethical discussions to the right place* …. Data Analytics Lab and Police representative 6

This position was shared to a degree by some Committee members who saw their **value in offering situational and contextual challenges**:

> *I'm always asking about 'what are the **worst-case scenarios**?' I want to know what the worst-case scenarios are of either the data science being flawed, but also the outcomes*

in how it is used which, quite frankly, is still not given enough attention in the committee, in my opinion, and certainly not in the public discourse about what these structures should be doing. Ethics Committee representative 6

You've got some people who don't think it's for us to comment, the police telling us it's not really relevant although they, you know, ebb and flow on how much they accept these arguments or not. Sometimes they do more than others. And then you've got another group who just don't seem to focus on that at all. It's almost like they're just stuck in their theoretical debate or the data science debate elements and not thinking about what its actual application is in the world and they might think I'm wrong to be focusing so much on the outcomes or the policy of a platform or whatever you want to call it. Ethics Committee representative 2

A compounding issue may be the high turnover of police workforce, which kept the Committee in a circular process of having to rebuild relationships that may help circumvent tensions:

The movement of police personnel is a real issue… they've picking up the baton and there are running with it, but this is in part to do with [the development] of a personal relationship. I don't think that's overstating it. Ethics Committee representative 1

People are in a job for two years and then they go and do something else, possibly in a different force…. It means that expertise is built up. Just at the point where people know what they're doing, they leave… the staff turnover is something fierce. And I don't think that helps anybody. Ethics Committee representative 2

The tensions within the Committee and DAL tend to be drawn from competing perspectives on the purpose and role of the Committee. Interestingly, while both parties expect there to be some tension, due to the nature of their roles, there were concerns, and added tensions with police representatives, when the Committee was believed to question policing strategy.

Summary

The Committee faces several challenges in its effective engagement with AI projects, **balancing the role as a critical friend, representative of the community, and providing ethical oversight**. One key issue is the use of technical language, which is necessary for an in-depth

understanding of AI projects, but a potential barrier due to the limited knowledge and capacity of some Committee members. To address this, the Committee has inducted specialists. However, this has led to a reliance on the knowledge of specialists by generalist members. Consideration should, therefore, be given to **creating a specific space for technical discussions during meetings to allow for in-depth and targeted consideration**.

Some interviewees **considered that it is not for the Committee to question police strategy**, taking the view that such questions present an inappropriate challenge to the operational independence of West Midlands Police. Others disagreed, **noting that the projects may impact on the scale or intrusiveness of policing operations and that it is appropriate therefore for the Committee to raise questions about how the outputs of projects might be acted upon**. A resolution to this could be the deeper inclusion and input from senior police officers who can ensure that the Committee is fully aware of the purpose of the proposed tool, and its intended use.

The second area of consideration is the further development of formal processes to track whether and how the Committee's recommendations have been followed, **for the return of operational projects to the Committee for review and for communication of the Committee's recommendations**. Currently, there are trackers available for projects. However, some stakeholders interviewed felt that they were ineffective. The two main critiques were that there was **no formal route for projects to be sent back to the Committee**, and the recommendations **were too formulaic, when more detail was required**. Consideration should be given to the format and content of recommendations to ensure that communication is clear and that the DAL can operationalise and feedback on recommendations. This communication should be at a level where, should projects be recommended for re-evaluation, the DAL can resubmit in the future and highlight the specific changes made with direct links to the recommendations. This may expedite the decision-making process. There should also be a way for projects which 'proceed' to **return to the Committee so the impact of their role in the community can be evaluated** and be fed into future decision making. Consideration should be given to a process map which can highlight what information and outcomes are expected at any given point in time.

4. Conclusions

Research Question 1- *Influence on technology*: What influence has the Committee had on the design and operationalisation of WMP AI policing projects?

The research found that the Committee exerts **influence**, though steps could be taken to improve transparency around the extent of that influence.

The research found that the Committee's exchanges with the Data Analytics Lab influences the design, operationalisation, transparency and good practice of AI projects in policing. However, there are challenges in terms of the time and expertise required for Committee members to understand the **technical detail** of the models (such as feature engineering) and thus the implications of the outputs.

The exchange between the Committee, WMP and the Data Analytics Lab influences the design and operationalisation of AI projects in more subtle ways. The Data Analytics Lab **anticipates** the concerns of the Committee in its approach to new projects. Senior police officers responsible for the delivery of projects speak of **insights** they have gained from these exchanges. The influence of the Committee extends beyond WMP when these officers join other policing bodies or take on national portfolios and strategic roles.

The extent to which the Committee's recommendations are followed is not immediately clear from a bare reading of the minutes. In the interests of transparency, informing public debate and improving trust, records should be maintained that not only detail the projects considered by the Committee and the Committee's advice and recommendations, but also the **impact** of those advice and recommendations on the development of the project. Where appropriate, these records should be made accessible to the public.

Research Question 2- *Human rights issues*: What human rights related issues were identified by the Committee and how were these issues dealt with in the design and operationalisation of AI tools?

The research found that the Committee has a strong human rights focus. Data Analytics Lab and Police representatives described how engagement with the Committee **enhanced or transformed** their understanding of issues with rights implications and informed their own approach, and that of the Data Analytics Lab, to the development of projects.

However, the research highlighted concerns about the productiveness of the Committee's human rights discourse, and that it may be **overly focused** on human rights relating to privacy and protection from discrimination. Interviewees, including Community representatives, spoke of the need to broaden rights-based conversations to take account of other rights, mentioning the right to a **fair trial** and the **state's positive obligations** under the prohibition against torture and right to life.

Committee representatives also expressed concerns about lack of information which would allow them to assess the **real-world outcomes** resulting from the operationalisation of AI projects. Lack of information about outcomes negatively impacts on the Committee's ability to assess the human rights and ethical implications of deployment of AI projects, and to understand technical successes and challenges.

Interviewees noted that rights issues are frequently raised without the Committee explicitly mentioning the right(s) at stake. This makes it difficult to identify **which human rights issues** are raised by the Committee from a reading of the minutes and impacts on the utility of the minutes as a source of learning and aid to transparency. It would be helpful for the minutes to explicitly identify which human rights issues were raised by the Committee.

Interviewees expressed concern about the lack of clarity as to when projects would **return to the Committee** for discussion and advice after operationalisation. This impacts on the ability of the Committee to fulfil its remit and address human rights and other ethical issues that may only become apparent after operationalisation.

Research Question 3- *Vulnerable groups and data*: How, if at all, are the interests, views and concerns of vulnerable groups incorporated within the ethical review process?

There are significant barriers to improving the representation of vulnerable communities within the Committee. It should be noted, however, that the Committee considers the interests of the community in its **discussions around privacy, disproportionality, and safeguarding**. The research identifies three stages of community representation that need to be addressed:

- First, there is **Accessibility and Definition**. An early barrier to community engagement is that community representatives are unsure of the role of the Committee within the policing ecosystem and how it might positively impact on the community. Community representatives are more familiar with the role of scrutiny panels, for example, and see involvement in those as a more effective use of their time. It is, therefore, important to promote the work of the Committee to potential Community representatives.
- Secondly, **Capacity and Influence** should be considered. It should be communicated to potential Community representatives that their value is in their contextual and personal experiences. Community representatives discussed feeling that a lack of technological knowledge would restrict their ability to engage in discussions. However, the other stakeholder groups did not consider this to be an issue, and that contextual conversations should be encouraged.
- Finally, **Dissemination and Development** should be employed to strengthen engagement with vulnerable groups. Direct engagement with communities is essential for better representation and the inclusion of other vulnerable voices within the community. The research suggests that it may be beneficial for Community representatives who join the Committee to act as community advocates, disseminating information about the Committee to the wider community, though other approaches may also be effective.

Research Question 4- *Challenges of ethical review*: What issues and challenges have Committee members and police representatives encountered in the committee review process?

In large part, the Committee fulfils its mandate and provides effective scrutiny of AI policing technologies. There are challenges, however.

There is a need to re-evaluate **communication processes** to ensure that the Committee has visibility as to whether, for each project, its recommendations have been followed. Consideration should be given to supplementing the Committee's recommendations to the Data Analytics Lab with an explanation of the **rationale** for the recommendations. This would allow the Data Analytics Lab to contextualise the Committee's concerns and respond appropriately. Currently, there are indications that absence of this information negatively impacts the efficiency of the review process.

There is a need for a clear protocol for the **return of 'live' projects to the Committee** for consideration and review post-implementation. The protocol should set out the timescales and other 'triggers' for the return of projects to the Committee. This would mean that active projects with real-world impacts could be fed into the decision-making process for future recommendations.

The use of **technical language** was identified as a potential barrier to involvement of community representatives and may slow down discussions. Consideration should be given to finding additional time and space during meetings for the technical discussions.

Although the Committee's remit clearly requires it to consider the impact of AI projects intended for use in policing, some Data Analytics Lab and Police representatives expressed concern that the Committee occasionally sought to challenge the **operational independence** of the police. The high turnover of police workforce may impact on the ethical review process through loss of established relationships. This may also impact on the work of the Data Analytics Lab as focuses and priorities change. Both these issues may be mitigated by the regular and consistent involvement of **senior operational police officers** in the Committee meetings in order to ensure that the operational priorities and actions behind the AI tools are explained and understood.

Research Question 5- *Potential of other models to improve the Committee process*: In what ways could the use of the factor's framework (Janjeva, Calder and Oswald 2023) and matrix evaluation model (Oswald, Chambers and Paul 2023) improve the development of Responsible AI in policing?

The research suggested that a structured framework for evaluation of AI projects would help maintain accountability, enable a detailed assessment of the impact of a project on human rights, be educational, and could form the **foundation stones** for ethical review. It may also support the integration of broader rights conversations, as specified in Research Question 2. The research's technical observations indicated how user information that indicates **probability and certainty** of model outputs to the police user could reduce the risk of overreliance. Further research and testing are necessary to explore the potential for the factors framework (Janjeva, Calder and Oswald 2023) and matrix evaluation model (Oswald, Chambers and Paul 2023) to improve the development of Responsible AI in policing.

Research Question 6- *Research challenges*: What challenges emerge from the research which would need to be addressed in larger research projects investigating embedded ethics processes?

The key challenges that emerged during this research were as follows:

- There was a **low response rate** from community members for interviews, with some of those responding, expressing interest but unable to commit due to other obligations. In future, alternative research methods, such as questionnaires, should be considered to complement interviews and provide avenues for individuals in demanding and resource limited sectors to participate in the research.
- The **thematic analysis** of the meeting minutes was challenging due to the limited detail included in the minutes. Implementing standardisation and clearer communication guidelines could make this approach more sustainable in future research. This would not

only aid future research endeavours but would provide a structured feedback mechanism that could improve the Committee's review and feedback.

- Technical observations were possible through the co-operation of the Data Analytics Lab and the coordination of staff from the WMOPCC. It would have been beneficial to observe operational systems with 'live' data, but it would not have been practicable to secure **vetting** for the researchers within the timeframe of this project. This demonstrates wider issues which would need to be addressed in larger research projects investigating embedded data/AI ethics processes in policing.

5. Recommendations

For national and international strategy, policy-makers and stakeholders

This research indicates that development of policing AI that includes consultation and advice from diverse independent voices can contribute to the robustness of technology implementation, from technical, human rights, operational and community perspectives. For such advice to add value, however, it must be **incorporated fully** into the implementation and oversight processes, and not regarded as an 'add-on' or tick-box exercise.

Although it takes time and effort to construct, embed and refine such an advisory process, it can contribute to a positive culture whereby the police develop **knowledge and understanding** of the issues likely to be raised by such independent oversight, thus enabling those issues to be anticipated and considered in **planning and implementation**. There can be tension between Committee members and police staff regarding the extent to which operational decisions fall within the Committee's remit. Committee members assert however that **deployment in practice of AI outputs must be understood, to assess potential benefits, risks/harms and proportionality**. The regular involvement of **operational police officers** to discuss and explain operational priorities and actions behind the AI tools can contribute to resolving these tensions. Furthermore, attention must be paid to police responsibilities for **public safety (and how AI may support these responsibilities)** as well as to risks related to privacy, fair trial and freedom of expression.

Having a data ethics process cannot be assumed to create or improve trust in policing AI per se, especially from vulnerable groups. For this to happen, effort must be put into incorporating the **voices of the community** into the process and ensuring that the work of the committee/panel is known and that its input is respected and influential.

Steps should be considered to highlight and disseminate information about the work of the Committee and the Data Analytics Lab in a manner that allows other police authorities, ethics committees, policy-makers, oversight bodies and international stakeholders to benefit. Lessons from the Committee's experience, including those reviewed in this report, can inform **best**

practice and feed into a framework for responsible AI in policing, including a **national model** for independent advice and oversight.

Such lessons include consideration of whether review of projects fully addresses all those human rights issues that are at stake, including the **positive rights** under Articles 2 and 3 (respectively, the right to life and freedom from torture/inhuman and degrading treatment). It should be noted that the Committee's experience shows that there are often **no 'black and white' answers** to ethical, legal or technical questions and the key issue of proportionality (e.g. reconciling privacy and security priorities relevant to the assessment of the proportionality of using suspect data, and whether precision or other performance metrics should be prioritised given the nature of the project). This research indicates however that a **structured framework** such as that proposed by Janjeva et al. (2023) might improve the productiveness, robustness and objectivity of deliberations about necessity and proportionality and assist in the practical application of the proportionality test when dealing with technical issues of data analytics and AI.

Building a culture of Responsible AI in policing depends on **time, resource, commitment, knowledge and collaborative communication**. It is important for police authorities to be aware of the potential issues that may be raised by independent oversight bodies, so they can plan and prepare for those conversations, while those involved in the oversight must be aware of the operational purposes and objectives of AI projects. This requires an **openness to both teach and learn** from other groups, and investment of time and resource in relationship-building.

For the WMOPCC, the Secretariat, WMP and Committee members

Naming and role of the Committee/Terms of Reference: consideration could be given to the renaming of the Committee in order to better reflect and communicate its advisory role e.g. Advisory Committee. Furthermore, it is recommended that the Terms of Reference are reviewed in light of any accepted actions from this report.

Performance metrics: there is an ongoing need for Committee members to understand the significance of the range of performance metrics (accuracy, sensitivity, precision, specificity) for predictive models, how these metrics are assessed, why the Data Analytics Lab may have favoured one metric over another, when one metric should be preferred over another and what

that selection might mean for the impact of the tool and the circumstances in which/purposes for which the tool should be used.

Communication Strategy/Process Map: Clearer protocols should be developed for the return of 'live' projects to the Committee to improve and inform future discussions, with triggers for return agreed on a case-by-case basis. This should be built into a clear process map that informs stakeholders about how projects can be returned to the Committee, and what they should expect in terms of feedback, recommendations, and advice. The Committee, Data Analytics Lab, WMP and WMOPCC should collaboratively discuss what information needs to be shared, when, by and with whom, about the intended purpose and use of projects and the likely impacts, whether for the purpose of ethical review or to guide the development of the project.

Project Tracking: In the interests of transparency, informing public debate and improving trust, records should be maintained that not only detail the projects considered by the Committee and the advice and recommendations, but also record the impact of the advice and recommendations on the development of the project. Where appropriate, these records should be made public.

Precedent bank: The value of keeping a record of outcomes throughout the lifecycle of projects was noted throughout the research. Building a 'precedent' bank of projects reviewed by the Committee could develop knowledge of technical, operational and ethical issues raised, and their solutions/mitigations, thus making future decision-making in relation to similar projects more effective.

The Minutes: The recommendations set out in the minutes should be supplemented with an explanation of the rationale for the Committee's recommendations including statements on the human rights and other key technical, legal and social issues identified by the Committee. This would allow the Data Analytics Lab to contextualise the Committee's concerns and respond appropriately. In order to increase communications and understanding, it is recommended that time be built into the meeting agendas to allow the DAL to return to hear the recommendations and rationales for these.

Annual Report: Recognising that additional record-keeping requires additional resources, it is recommended that in the interests of transparency and as a means of informing public debate and improving public trust, the Committee should produce an annual report which summarises

its work and offers a brief commentary on human rights and other key issues addressed by its recommendations. This would additionally support public knowledge of its role and remit.

Impact of Technical Language: Both Community representatives and Committee members noted that the use of technical language can act as a barrier to participation and impede discussion at meetings. To address this, stakeholders should consider solutions, such as the timetabling of additional time for technical discussion, or members given specific responsibilities regarding the explanation of technical details to other members.

For the WMOPCC, Community representatives and Committee members

Enhanced inclusion of Community representatives: Consideration should be given to developing an engagement plan for community members to ensure they are represented in the Committee and are able to contribute appropriately. Further to this, opportunities should be sought out to improve public knowledge of the Committee, its remit, objectives, and potential benefits. As part of this, direct community engagement should be considered through practices such as informal meetings, training sessions, workshops, or engagement panels.

Creating Space for Community Voices: The research found that a concern of community representatives was that they felt they could not contribute to discussion on AI or data ethics, and that a lack of knowledge would leave them 'exposed' during meetings. Although Committee members recognised the value of community voices, and the contextualisation of ethical concerns, the cultivation of inclusive conversations and welcoming environments should be prioritised to help develop community confidence.

For academic researchers, and research funding bodies

Proportionality and policing AI: Assessment of proportionality and human rights-related risks concerning the use of policing AI in real contexts is not straightforward and can be fraught with ambiguity, particularly where law enforcement bodies may be expected to use new methods of data analytics and AI to tackle public safety risks. It is clear from the Committee's experience that a proportionality assessment that is alive to how the technical, statistical and operational details of policing AI link to the legal test is dependent on an **interdisciplinary approach**. More research

could be done to consider how the proportionality test could be better applied in 'messy' realities, and to avoid proportionality merely becoming a '***ritual incantation***'.

Challenges of AI-related research in sensitive contexts: considerable issues of **access** arise in research projects investigating data/AI and ethics processes in policing, even for researchers with existing research relationships and who hold levels of vetting. Funding bodies and national research programmes could consider ways that appropriate and systematic access could be given to researchers on a more equitable basis. However, building research relationships and knowledge of the consequences of AI in this context requires time and commitment and therefore **longer cycles of funding** would be more beneficial in this regard.

Research limitations

The researchers acknowledge a number of limitations and constraints. Two of the researchers involved in the project are members of the Committee, which could have introduced an element of bias. These researchers were primarily involved in the direction, oversight and management of the project. However, one researcher was involved in elements of data collection and recruitment as well as reflective analysis sessions to identify and refine themes for the research, and one researcher was involved substantially in the write-up and finalisation of the report. In order to reduce any potential bias in selection, interviewing and transfer, the following processes were employed:

- Embedding reflexivity in the approach to research.
- Participant selection of Committee and DAL representatives was completed by members of the research team who had no direct link with the Committee.
- Clearly defined topics were introduced to the interviews to standardise the process.
- Analysis of interviews with Committee and/or DAL representatives was completed by members of the research team who had no direct link with the Committee.

In addition, the representation of community voices was limited, primarily coming from advocacy groups that promote social inclusion and equality. This may have limited an intersectional approach to the research (see: Challenges to the intersectional research approach). For future studies, we recommend including perspectives from general community groups.

The number of systems included in the technical observations was limited. Access to a greater spread of AI systems would have provided us with a better sense of how police officers and other users interact with AI systems in use in a policing context, what information is communicated via such systems about their reliability, capabilities and limitations, and what training if any is provided about the systems.

References

Afzal, M. and Panagiotopoulos, P., 2024. 'Data in Policing: An Integrative Review', *International Journal of Public Administration*: 1–20. Available at: https://doi.org/10.1080/01900692.2024.2360586.

AI and Data Protection Risk Toolkit - OECD.AI., 2023. Available at: https://oecd.ai/en/catalogue/tools/ai-and-data-protection-risk-toolkit (Accessed: 8 July 2024).

Atewologun, D., 2018. 'Intersectionality Theory and Practice', *Oxford Research Encyclopedia of Business and Management* 1:1.

Babuta, A., 2017. 'Big Data and Policing: An Assessment of Law Enforcement Requirements, Expectations and Priorities', *Royal United Services Institute for Defence and Security Studies* (RUSI). Available at: https://static.rusi.org/201709_rusi_big_data_and_policing_babuta_web.pdf.

Babuta, A. and Oswald, M., 2020. 'Data analytics and algorithms in policing in England and Wales: Towards a new policy framework'. [Report]. *London: Royal United Services Institute (RUSI)*. Available at: https://rusi.org/sites/default/files/rusi_pub_165_2020_01_algorithmic_policing_babuta_final_web_copy.pdf (Accessed: 3 April 2024).

Babuta, A. and Oswald, M., 2021. 'Machine learning predictive algorithms and the policing of future crimes', in J.L.M. McDaniel and K.G. Pease (eds) *Predictive policing and artificial intelligence*. Abingdon, Oxford: Routledge.

Behizadeh, N., 2024. "Agential Cuts for Justice: Honoring Complexity in Research through Intersectional Design Dimensions", *The Qualitative Report*. **29** (3), pp.782-792

Bates, E.S., 2020. ‘’Impossible of Disproportionate Burden’: The UK’s Approach to the Investigatory Obligation under Articles 2 and 3 ECHR’, *European Human Rights Law Review* (5): 499–511.

Blackburn, R.A., 2018. ‘Summary of the 2018 Department of Defense Artificial Intelligence Strategy’.

Blei, D.M., Ng, A.Y. and Jordan, M.I.. 2003. ‘Latent dirichlet allocation’, *J. Mach. Learn. Res.*, 3(null), pp. 993–1022.

Borgesius, F. J., 2020. ‘Strengthening legal protection against discrimination by algorithms and artificial intelligence’, *The International Journal of Human Rights*, 24(10): 1572-1593, DOI: 10.1080/13642987.2020.1743976

Brown, S., Davidovic, J. and Hasan, A., 2021. ‘The algorithm audit: Scoring the algorithms that score us’, Big Data & Society, 8(1), pp.1-8. Available at: https://doi.org/10.1177/2053951720983865.

Brugge, D. and Kole, A., 2003. ‘A case study of community-based participatory research ethics: the healthy public housing initiative’. *Scientific Engineering Ethics*, 9(4):485–501.

Buolamwini, J., and Gebru, T., 2018. ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.’ *Proceedings of Machine Learning Research* 81:1–15, Conference on Fairness, Accountability, and Transparency.

Burrell, J., 2016. ‘How the machine ‘thinks’: Understanding opacity in machine learning algorithms’, *Big Data and Society,* 3 (1), pp 1-12.

Cash, R., Wikler, D., Saxena, A., Capron, A., 2009. *Casebook on ethical issues in international health research*. World Health Organisation, Geneva.

Castets-Renard, C., 2022. ‘Human rights and algorithmic impact assessment for predictive policing’. In Micklitz H, Pollicino O, Reichman A, Simoncini A, Sartor G, Gregorio G (eds) *Constitutional challenges in the algorithmic society*. Cambridge University Press, Cambridge.*Catt v United Kingdom* [2019] ECHR 76

Cho, S., Crenshaw, K., McCall, L., 2013. ‘Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis’, *Journal of Women in Culture and Society*, 785.

Coleman, C.H. and Bouësseau, M.C., 2008. 'How do we know that research ethics committees are really working? The neglected role of outcomes assessment in research ethics review', *BMC medical ethics*, 9:6. Available at: https://doi.org/10.1186/1472-6939-9-6.

College of Policing, 2013. *Legal framework and legislation*. Available at: https://www.college.police.uk/app/public-order-public-safety/legal-framework-and-legislation (Accessed: 7 July 2024).

College of Policing., 2016. 'Stop and Search: Authorised Professional Practice' [Online] https://www.college.police.uk/app/stop-and-search/stop-and-search. Last Accessed: 26/06/2024.

Collins, P.H., 2015. 'Intersectionality's Definitional Dilemmas', 41 *Annual Review of Sociology* 1:2.

Collins, P.H., 2000., *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment,* Hill Collins.

Colvin, M. and Cooper, J. (eds)., 2009. *Human rights in the investigation and prosecution of crime*. Oxford; New York: Oxford University Press.

Commissioner of Police of the Metropolis v DSD and another [2018] UKSC 11.

Court and Tribunals Judiciary, 2023. "Articifical Intelligence (AI) Guidance for Judicial Office Holders. [Online] Available at: https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf. Last Accessed: 22 July 2024.

Corrêa, N.K., Galvao, C., Santos, J., Del Pino, C., Pinto, E.P., Barbosa, C., Massmann, D., Manbrini, R., Galvao, L., Terem, E., de Oliveria, N., 2023. 'Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance', *Patterns,* 4(10):100857. Available at: https://doi.org/10.1016/j.patter.2023.100857.

Crenshaw, K., 1998., 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *Feminism and Politics,* University of Chicago Legal Forum.

Crenshaw, K., 1991. 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color', *Stanford Law Review,* 43 (6):1241.

Crico, C., Sanchini, V., Casli, P., and Pravettoni, G., 2020. 'Evaluating the effectiveness of clinical ethics committees: a systemic review', *Medical Health Care Philosophy,* 24 (1).

Cross, J., Pickering, K. and Hickey, M. 2014. Community-based participatory research, ethics, and institutional review boards: untying a Gordian knot. *Critical Sociology*, 1–20. doi:10.1177/0896920513512696.

Department for Media, Culture and Sport, 2021. National AI Strategy (Command Paper 525). Available at: https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf.

Devine, D., Boswell, J., and Smith, J., 2022. Managing a minister, *Institute for Government*.

Dixon, B., 2018, 'Who Needs Critical Friends? Independent Advisory Groups in the Age of the Police and Crime Commissioner', *Policing: A Journal of Policy and Practice.* 14 (3).

Durmus, M., 202. 'A brief Overview of some Ethical-AI Toolkits', Nerd For Tech, 16 June. Available at: https://medium.com/nerd-for-tech/an-brief-overview-of-some-ethical-ai-toolkits-712afe9f3b3a (Accessed: 15 July 2024).

Edwards, L. and Urquhart, L., 2016. 'Privacy in public spaces: what expectations of privacy do we have in social media intelligence?', *International Journal of Law and Information Technology*, 24 (3): 279–310. Available at: https://doi.org/10.1093/ijlit/eaw007.

Ethics guidelines for trustworthy AI | Shaping Europe's digital future, 2019. Available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (Accessed: 17 June 2024).

European Data Protection Supervisor, 2017. *Assessing the necessity of measures that limit the fundamental right to the protection of personal data: A Toolkit*. Available at: https://www.edps.europa.eu/data-protection/our-work/publications/papers/necessity-toolkit_en (Accessed: 7 July 2024).

European Data Protection Supervisor, 2019. *EDPS Guidelines on assessing the proportionality of measures that limit the fundamental rights to privacy and to the protection of personal data*. Available at: https://www.edps.europa.eu/data-protection/our-

work/publications/guidelines/assessing-proportionality-measures-limit_en (Accessed: 7 July 2024).

European Union Agency for Fundamental Rights, 2022. *Bias in Algorithms: Artificial Intelligence and Discrimination Report* (2022):25.

Ferguson, A.G., 2019. 'Predictive Policing Theory', in T.R. Lave and E.J. Miller (eds) *The Cambridge Handbook of Policing in the United States*, Cambridge: Cambridge University Press (Cambridge Law Handbooks): 491–510. Available at: https://doi.org/10.1017/9781108354721.025.

Fineman, M., 2019. 'Vulnerability and Social Justice', *Valparaiso University Law Review,* 53 (2):341-370.

Franks, P.C. and Cameron, S., 2023. Paradata: Documentation for Responsible Artificial Intelligence. Available at: https://info.aiim.org/aiim-blog/paradata-documentation-for-responsible-artificial-intelligence (Accessed: 12 February 2024).

Gan, J. and Qi, Y., 2021. 'Selection of the Optimal Number of Topics for LDA Topic Model—Taking Patent Policy Analysis as an Example', *Entropy*, 23(10), p. 1301. Available at: https://doi.org/10.3390/e23101301.

Gebru, T. et al., 2021. 'Datasheets for datasets', Communications of the ACM, 64(12), pp. 86–92. Available at: https://doi.org/10.1145/3458723.

Gill, G.N. and Joshi, F., 2024 'Environmental Public Hearings, Intersectionality and Women's Voices: Criticalities and Scrutiny from Gujarat, India', *Journal of Law and Society*, 51 (2): 163

Goddard, K., Roudsari, A. and Wyatt, J.C., 2012. 'Automation bias: a systematic review of frequency, effect mediators, and mitigators', *Journal of the American Medical Informatics Association: JAMIA*, 19(1), pp. 121–127. Available at: https://doi.org/10.1136/amiajnl-2011-000089.

Goldberg, Z., 2022. 'How to Conduct an Ethics Assessment of AI in Policing', *Proceedings of the 14th ACM Web Science Conference 2022*. New York, USA.

Golunova, V. 2023. 'Artificial Intelligence and the Right to Liberty and Security', in A. Quintavalla, and J. Temperman (eds.), *Artificial Intelligence and Human Rights* (pp. 45-60), Oxford University Press.

Gorman S., 2011. Ethics creep or governance creep? Challenges for Australian Human Research Ethics Committees (HRECs). *Monash Bioethics Review*, 29(4): 14.1–14.16.

Grace, J., 2018. 'Human Rights, Regulation and the Development of Algorithmic Police Intelligence Analysis Tools in the UK'. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3303313.

Grace, J., 2019. 'Machine Learning Technologies and Their Inherent Human Rights Issues in Criminal Justice Contexts', *SSRN Electronic Journal*. Available at: http://doi.org/10.2139/ssrn.3487454.

Grobelnik, M., Perset, K., and Russell, S., 2024. 'What is AI? Can you make a clear distinction between AI and non-AI systems?', *OECD.AI Policy Observatory*. Available at: https://oecd.ai/en/wonk/definition (Accessed: 17 June 2024).

Hagendorff, T., 2020. 'The Ethics of AI Ethics: An Evaluation of Guidelines', *Minds and Machines*, 30(1), pp. 99–120. Available at: https://doi.org/10.1007/s11023-020-09517-8.

Hamilton, M., 2019. 'The Biased Algorithm: Evidence of Disparate Impact on Hispanics' *American Criminal Law Review* 56(4).

Health Research Authority, 2023. Information for potential Research Ethics Committee Members. Available online at Information for potential Research Ethics Committee Members - Health Research Authority (hra.nhs.uk). Available at: Information for potential Research Ethics Committee Members - Health Research Authority (hra.nhs.uk). Accessed on 10 July 2024.

Hedgecoe A. M., 2012. Trust and regulatory organisations: The role of local knowledge and facework in research ethics review. *Social Studies of Science*, 42(5): 662–683. https://doi.org/10.1177/0306312712446364

Hickman, M., and Piquero, A., 2009. 'Organizational, Administrative, and Environmental Correlates of Complaints About Police Use of Force: Does Minority Representation Matter?', *Crime & Delinquency.* 55 (1): 3-162.

Ide, Y., and Beddoe, L., 2023. “Challenging perspectives: Reflexivity as a critical approach to qualitative social work research”, *Qualitative Social Work,* 24 (4):725-740.

Information Commissioner’s Office, no date a. *AI and data protection risk toolkit*. ICO. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/ai-and-data-protection-risk-toolkit/ (Accessed: 8 July 2024).

Information Commissioner’s Office, no date b. *What is a DPIA?* ICO. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/accountability-and-governance/data-protection-impact-assessments-dpias/what-is-a-dpia/ (Accessed: 8 July 2024).

Information Commissioner’s Office, A.I., 2023. ‘Definitions’. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/definitions/ (Accessed: 17 June 2024).

Janjeva, A., Calder, M. and Oswald, M., 2023. *Privacy Intrusion and National Security in the Age of AI*. CETaS Research Reports. Alan Turing Institute, https://cetas.turing.ac.uk/publications/privacy-intrusion-and-national-security-age-ai.

Justice and Home Affairs Committee. 2022. ‘Technology Rules? The advent of new technologies in the justice system’, *1st Report of Session 2021-22*.

Jenkins, R, and Purves. D., 2020. ‘Artificial Intelligence and Predictive Policing: A Roadmap for Research.’ *Ethics and Emerging Sciences Group*.

Kennedy, H.G., Mullaney, R., McKenna, P., Thompson, J.,m Timmons, D., Gill, P., O’Sullivan, O., Braham, P., Duffy, D., Kearns, A., Linehan, S., Mohan, D., Monks, S., McLoughlin, L., O’Connell, P., O’Neill, C., Wright, B., O’Reilly, K., Davoren, M., 2020. ‘A tool to evaluate proportionality and necessity in the use of restrictive practices in forensic mental health settings: the DRILL tool (Dundrum restriction, intrusion and liberty ladders)’, *BMC Psychiatry*, 20: 515. Available at: https://doi.org/10.1186/s12888-020-02912-6.

Lekhtman, A., 2021. *Should I Look at Precision & Recall OR Specificity & Sensitivity?*, *Medium*. Available at: https://towardsdatascience.com/should-i-look-at-precision-recall-or-specificity-sensitivity-3946158aace1 (Accessed: 18 July 2024).

Leslie, D., Burr, C., Aitken, M., Cowls, J., Katell, M., and Briggs, M., 2021. Artificial intelligence, human rights, democracy, and the rule of law: a primer, *The Council of Europe*.

Longstaff, A., Willer, J., Chapman, J., Czarnomski, S., Graham, J., 2015. 'Neighbourhood policing: Past, present and future'. Available at: http://www.police-foundation.org.uk/uploads/catalogerfiles/neighbourhood-policing-past-present-and-future ? a-review-of-the-literature/neighbourhood_policing_past_present_future.pdf (Accessed: 21 July 2024).

Lundy, L., 2007. 'Voice is not Enough: Conceptualising Article 12 of the United Nations Convention on the Rights of the Child', *British Educational Research Journal.* 33 (6): 927-942.

May, V.M., 2015, 'Pursuing Intersectionality, Unsettling Dominant Imaginaries', *Journal of Critical Realism,* 15 (5): 543-547.

Metropolitan Police, no date. 'How the gangs violence matrix works'. Available at: https://www.met.police.uk/police-forces/metropolitan-police/areas/about-us/about-the-met/gangs-violence-matrix/ (Accessed: 14 July 2024).

McIntosh, P., 2012. 'Reflections and Future Directions for Privilege Studies' *Journal of Social Issues*, 68:194.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Soitzer, E., Raji, I., and Gebru, T., 2019. 'Model Cards for Model Reporting', in Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 220–229. Available at: https://doi.org/10.1145/3287560.3287596.

Mittelstadt, B.D., Brent, D., Allo, P., Taddeo, M., Wachter, S., Floridi, L., 2016. 'The ethics of algorithms: Mapping the debate', *Big Data & Society*, **3**(2), p. 2053951716679679. Available at: https://doi.org/10.1177/2053951716679679.

Mkwananzi, F., Cin, F. M. and Marovah, T., 2021. 'Transformative youth development through heritage projects: connecting political creative and cultural capabilities', *International Journal of Heritage Studies*, 29(6): 581–597.

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mokander, J., and Floridi, L., 2021. 'Ethics as a Service: A Pragmatic Operationalisation of AI Ethics', *Minds and Machines*, **31**(2): 239–256. Available at: https://doi.org/10.1007/s11023-021-09563-w.

Morley, J.,Floridi, L., Kinsey, L., and Elhalal, A., 2020. 'From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices', *Science and Engineering Ethics*, 26(4), pp. 2141–2168. Available at: https://doi.org/10.1007/s11948-019-00165-5.

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., Floridi, L., 2023. 'Operationalising AI ethics: barriers, enablers and next steps', *AI & SOCIETY*, **38**(1): 411–423. Available at: https://doi.org/10.1007/s00146-021-01308-8.

Munn, L. 2023. 'The uselessness of AI ethics', *AI and Ethics*, 3(3): 869–877. Available at: https://doi.org/10.1007/s43681-022-00209-w.

Moreira, N.A., 2022. 'The compatibility of AI in Criminal System with the ECHR and ECtHR Jurisprudence' In: Marreiros, G., Martins, B., Paiva, A., Ribeiro, B., Sardinha, A., (eds*) Progress in Artificial Intelligence. EPIA 2022. Lecture Notes in Computer Sciences, vol 13566*. Springer Cham. https://doi.org/10.1007/978-3-031-16474-3_10.

Murji, K., 2011. 'Working Together: Governing and Advising the Police', *The Police Journal*. 84 (3) pp.256-271.

Murray, D., 2020. 'Using Human Rights Law to Inform States' Decisions to Deploy AI', *AJIL Unbound*, 114: 158–162. Available at: https://doi.org/10.1017/aju.2020.30.

National Police Chiefs' Council., 2023. 'Covenant for Using Artificial Intelligence (AI) in policing' [Online] Available at: https://science.police.uk/site/assets/files/4682/ai_principles_1_1_1

Ng, L.,2019. 'Feature Engineering for Design Thinking Assessment', *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), pp. 3891–3900. Available at: https://doi.org/10.1017/dsi.2019.396.

Oswald, M., Grace, J., Urwin, S., Barnes, G., 2018. 'Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality', Information & Communications Technology Law, 27(2): 223–250. Available at: https://doi.org/10.1080/13600834.2018.1458455.

Oswald, M., 2022. 'A three-pillar approach to achieving trustworthy and accountable use of AI and emerging technology in policing in England and Wales: Lessons from the West Midlands data ethics model', *European Journal of Law and Technology*, 13(1). Available at: https://ejlt.org/index.php/ejlt/article/view/883 (Accessed: 26 June 2024).

Oswald, M., Chambers, L. and Paul, A., 2023. 'Evaluating (semi)-autonomous systems in policing and national security: a new framework based on the concept of 'intelligence': A new matrix framework of evaluation and grading, based on lessons from existing processes designed to define and assess 'intelligence'', in *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*. *TAS '23: First International Symposium on Trustworthy Autonomous Systems*, Edinburgh United Kingdom: ACM, pp. 1–2. Available at: https://doi.org/10.1145/3597512.3597524.

Ovalle, A., Subramonian, A., Gautam, V., Gee, G., and Gautam, V., 2023. "Factoring the Matrix of Domination: A Critical Review and Reimagination of Intersectionality in AI Fairness", *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society,* 23:496-511

Palmiotto, F., 2021. 'The Black Box on Trial: The Impact of Algorithmic Opacity on Fair Trial Rights in Criminal Proceedings', in Ebers, M., Gamito, M.C., (eds) *Algorithmic Governance and Governance of Algorithms*.

Panda, M. and Misra, H.K. (eds), 2021. *Handbook of research on automated feature engineering and advanced applications in data science*. Hershey, PA: Engineering Science Reference.

Passi, S. and Vorvoreanu, M., 2022. 'Overreliance on AI Literature Review', *Microsoft Research*.

Perry, W.L., 2013. *Predictive policing: the role of crime forecasting in law enforcement operations*. Santa Monica, CA: RAND.

Peters, A., 2021. 'A plea for proportionality: A reply to Yun-chien Chang and Xin Dai', *International Journal of Constitutional Law*, 19(3): 1135–1145. Available at: https://doi.org/10.1093/icon/moab071.

Powell, R., & Oswald, M., 2024. *Assurance of Third-Party AI Systems for UK National Security: Research Report*. The Alan Turing Institute. https://cetas.turing.ac.uk/publications/assurance-third-party-ai-systems-uk-national-security

Purshouse, J. and Campbell, L. 2022. 'Automated facial recognition and policing: a *Bridge* too far?,' *Legal Studies* 42 DOI: 10.1017/lst.2021.22

Quattrocolo, S., Anglano, C., Canonico, M., and Guazzone, M., 2020. 'Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings,' *Global Jurist*, 20(1) DOI:10.1515/gj-2019-00

R (*on the application of Bridges*) *v Chief Constable of South Wales Police* [2020] EWCA Civ 1058.Radiya-Dixit, E. and Neff, G., 2023. 'A Sociotechnical Audit: Assessing Police Use of Facial Recognition', in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery (FAccT '23):1334–1346. Available at: https://doi.org/10.1145/3593013.3594084.

Ramshaw, A., 2019. 'The case for replicable structured full proportionality analysis in all cases concerning fundamental rights', *Legal Studies*, 39(1): 120–142. Available at: https://doi.org/10.1017/lst.2018.18.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

Rodrigues, R. 2020. 'Legal and human rights issues of AI: Gaps, challenges and vulnerabilities', *Journal of Responsible Technology*, 4. Available at: https://doi.org/10.1016/j.jrt.2020.100005.

Russell, S.J. and Norvig, P., 2016. *Artificial intelligence: a modern approach*. *Third edition, Global edition*. Prentice Hall.

Sachoulidou, A., 2023, "Going beyond the 'common suspects': to be presumed innocent in the era of algorithms, big data and artificial intelligence', *Artificial Intelligence and Law*. Available at: https://doi.org/10.1007/s10506-023-09347-w.

Santos, R.B., 2017. *Crime analysis with crime mapping. Fourth edition,* Los Angeles: SAGE Publications, Inc.

Schatz, E. and Schiffer, K. 2008. *Marginalisation, Social Inclusion and Health*, Colophon.

Schopper, D., Dawson, A., Upshur, R., Ahmad, A., Jesani, A., Ravinetto, R., 2015. 'Innovations in research ethics governance in humanitarian settings', *BMC Medical Ethics* 16 (10).

Schuppli, C.A. and Fraser, D., 2007. Factors influencing the effectiveness of research ethics committees. *Journal of medical ethics*, 33(5), pp.294-301.

Scottish Human Rights Commission. 2021. *Human Rights and New Technologies in Policing: Issues Paper for the Independent Advisory Group on Emerging Technologies in Policing*, Scottish Human Rights Commission.

Siber, R., Brandusescu, A., and Sangiambut, S., 2024. 'Who are the publics engaging in AI?', *Public Understanding of Science,* 33 (5).

Snelling, P., Macvean, A., and Lewis, R., 2023. 'Police ethics committees in England and Wale: Exploratory online and web surveys', *Policing: A journal of Policy and Practice*. 17.

Sorell, T., 2024., 'AI-related data ethics oversight in UK policing', *Policing: A journal of Policy and Practice*, 18:16

Sous, A.M., 2020. 'R v Evans: An Uneasy Precedent?', *LSE Law Review*.

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A.L., Shah, J., Tambe, M., and Teller, A., 2022. 'Artificial Intelligence and Life 2030: The One Hundred Year Study on Artificial Intelligence', *Computer Sciences*.

Stoykova, R., 2023. 'The right to a fair trial as a conceptual framework for digital evidence rules in criminal investigations' in *Computer Law and Security Review* .49 (2023).

Strathern, M., 2000. 'New accountabilities'. In Strathern M. (Ed.), *Audit cultures: Anthropological studies in accountability, ethics and the academy*: 1–18. Routledge.

Taylor, M., 2023. 'Social Network Analysis in Homicide Investigations', in *The Routledge International Handbook of Homicide Investigation*. Routledge.

Ulnicane, I., 'Intersectionality in Artificial Intelligence: Framing Concerns and Recommendations for Action', *Social Inclusion* 12 (2024).

UNICRI and INTERPOL, 2024. *Toolkit for Responsible AI Innovation in Law Enforcement (Revised February 2024).* Available at: https://unicri.it/Publication/Toolkit-for-Responsible-AI-Innovation-in-Law-Enforcement-UNICRI-INTERPOL (Accessed: 8 July 2024).

Vallance, C., 2024. 'Police worried 101 call bot would struggle with 'Brummie' accents', *BBC News*, 7 March. Available at: https://www.bbc.com/news/technology-68466369 (Accessed: 13 July 2024).

Veale, M., 2019. Algorithms in the criminal justice system. Available at: https://www.lawsociety.org.uk/topics/research/algorithm-use-in-the-criminal-justice-system-report.

Walsh, M., Stead, V., Sawyer, S. M., O'Shea, A., Watson, J. M., & Anderson, K. L. M. 2024. In Pursuit of Ethical and Inclusive Research: What Ethics Committees and Disability Researchers Can Learn From Each Other. *International Journal of Qualitative Methods*, 23. https://doi.org/10.1177/16094069241237549.

Wang, P., 2019. 'On Defining Artificial Intelligence*', Journal of Artificial General Intelligence*, 10 (2), pp. 1–37. Available at: https://doi.org/10.2478/jagi-2019-0002.

Wang, S., 2021. 'Practice AI Responsibly with Proxy Variable Detection', *GAMMA—Part of BCG X*. Available at: https://medium.com/bcggamma/practice-ai-responsibly-with-proxy-variable-detection-42c2156ad986 (Accessed: 8 July 2024).

West Midlands Police & Crime Commissioner, no date. 'Ethics Committee'. Available at: https://www.westmidlands-pcc.gov.uk/ethics-committee/ (Accessed: 26 June 2024).

West Midlands Police and Crime Commissioner, 2019. 'Terms of Reference'. Available at: https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2019/07/Ethics-Committee-Terms-of-Reference-as-at-1-April-2019.pdf. (Accessed: 26 June 2024).

Wilson, J., Hume, J., O'Donovan, C., and Smallman, M., 2023. 'Providing ethics advice in a pandemic, in theory and in practice: A taxonomy of ethics advice', *Bioethics*. 38 (3)

Wong, R.Y., Madaio, M.A. and Merrill, N. (2023) 'Seeing Like a Toolkit: How Toolkits Envision the Work of AI Ethics', *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), pp. 1–27. Available at: https://doi.org/10.1145/3579621.

Wrigley A, Dawson A. Vulnerability and Marginalized Populations. 2016 Apr 13. In: H. Barrett D, W. Ortmann L, Dawson A, et al., editors. *Public Health Ethics: Cases Spanning the Globe* Springer.

Yeung, K., Howes, A. and Pogrebna, G., 2020. 'AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing', in M.D. Dubber, F. Pasquale, and S. Das (eds) *The Oxford Handbook of Ethics of AI*. Oxford University Press, p.0. Available at: https://doi.org/10.1093/oxfordhb/9780190067397.013.5.