



**HEDDLU
GOGLEDD CYMRU
NORTH WALES
POLICE**

Response Date: 26/03/2026

2026/285 - Use of artificial intelligence and automated decision-making systems within the police force

In response to your recent request for information regarding;

- **A list of any artificial intelligence, machine learning, or algorithmic decision-making systems currently used by the force. If possible, please include the name of the system, its purpose, and when it was first introduced.**

None currently used

- **A list of any AI or algorithmic systems that were previously used by the force but are no longer in use.**

South Wales Police (SWP) live facial recognition
A5 Object Detection

- **Copies of any policies, guidance, or internal documents relating to the use of AI or automated decision-making tools within the force.**

Please see attached our AI Usage Policy

Additionally, our Live Facial Recognition Policy is available on our website;
[north-wales-police-live-facial-recognition---policy-doc.pdf](#)

- **Copies of any impact assessments relating to these systems.**

North Wales Police do not hold a Force impact assessment.

- **Any reports, evaluations, or testing that have been carried out to assess the accuracy, reliability, or potential bias of these systems.**

Please see NIST report; [Face Recognition Vendor Test \(FRVT\) | NIST](#)

- **Details of any partnerships with external organisations (for example technology companies, universities, or other agencies) relating to the development or use of AI systems.**

Mutual aid from South Wales Police for live facial recognition
Assistance from National Highways with object detection in cars
See attached report on public perceptions of AI
Dataiku supply the MLOps platform that we are using to build our own ML solutions

- **Copies of any strategies, roadmaps, or plans relating to future adoption of AI technologies within the force.**

MLOps platform Dataiku DSS has just been implemented and the first few models are being prioritised

Pencadlys yr Heddlu,
Glan-y-Don, Bae Colwyn LL29 8AW
www.heddlugogleddcymru.police.uk

Police Headquarters,
Glan-y-Don, Colwyn Bay LL29 8AW
www.northwales.police.uk

THIS INFORMATION HAS BEEN PROVIDED IN RESPONSE TO A REQUEST
UNDER THE FREEDOM OF INFORMATION ACT 2000, AND IS CORRECT AS AT 25/03/2026



**HEDDLU
GOGLEDD CYMRU
NORTH WALES
POLICE**

AI Usage Policy

Governance:	Strategic Management Board		
Document Type:	Policy		
Policy Owner:	Head of Corporate Services		
Department:	Corporate Services Business Intelligence		
Policy Writer:	Advanced Analytics and AI Lead: ...		
Policy Number:	037	Version:	1.1
Effective Date:	25/09/2023		
Recommended Review Date:	25/09/2024		

POLICY

CHANGE HISTORY				
Version No	Author	Changes	Ratification	Date
1.1	...	Updated to reflect changes of names of boards and change name from PAMLAI to AI		

CONTENTS

1. WHY IS THIS POLICY REQUIRED?	3
2. WHO SHOULD USE THIS POLICY?	3
3. WHAT SHOULD I CONSIDER WHEN USING THIS POLICY?	3
3.1 The AI Ethics Code	4
<i>The entire ethics code can be found in Appendix D but the main headings are summarised here to give context:</i>	<i>4</i>
Summary of Code of Ethics	4
3.2 The AI Ticklist can be found in Appendix E	4
3.3 The AI Use Case and Guidance can be found Appendix F.....	4
4. ROLES AND RESPONSIBILITIES	5
4.1 AI Ethics Process Initiator	5
4.2 AI Facilitator	5
4.3 AI Ethics Group (AIG)	5
4.4 Assurance Board.....	5
4.5 Senior Management Board	5
5. THE PROCESS	6
5.1 AI Process: Initiation	6
5.1 AI Process: Guidance	7
5.2 AI Process: Oversight, Authorisation, or Escalation	7
5.3 AI Process: Authorisation or escalation	7
5.4 AI Process: Authorisation	8
6. DECLARATION & LEGALITIES.....	8
7. APPENDICIES	10
7.1 Appendix A – Authorisation Levels	10
7.2 Appendix B - Authorisation Matrix.....	11
7.3 Appendix C – Detailed Process Map	12
7.4 Appendix D – AI Ethics Code	13
7.5 Appendix E – AI Ticklist.....	13
7.6 Appendix F – AI Guidance.....	Error! Bookmark not defined.
7.7 Appendix G – Sample Registration Form	13
7.8 Appendix H – Terminology	13

1. WHY IS THIS POLICY REQUIRED?

This policy will

- 1.1 Reduce the number of risks associated with **Predictive Analytics**, **Machine Learning**, **Artificial Intelligence**, and related technologies (hereinafter referred to as **AI** technology).
- 1.2 Lower the high level of risk associated with the use of **AI** technology.
- 1.3 Minimise the danger of large-scale reputational damage experienced by others for not having an appropriate policy in place.
- 1.4 Ensure that any AI technology in use is created, sourced, and used ethically.
- 1.5 Ensure that any AI technology used will be *benevolent*.
- 1.6 Ensure that there is *accountability* for the
 - Creation of AI technology.
 - Purchase of AI technology.
 - Use of AI technology.
 - Maintenance of AI technology
- 1.7 Ensure that any AI technology used is *robust*.
- 1.8 Help to avoid biases inherent in AI work.
- 1.9 Help to avoid injecting biases into AI work.
- 1.10 Ensure compliance with the NPCC guidelines (ALGOCARE).
- 1.11 Facilitate compliance with UK National, and European legislation.

2. WHO SHOULD USE THIS POLICY?

- 2.1 This policy should be used by NWP officers, special constables, police community support officers, police staff, volunteers, and anyone under contract to NWP (e.g. agency staff). considering the creation, purchase, rental, use or deployment of AI technology, including the use of AI technology from the public domain.
- 2.2 This policy includes:
 - Direction, together with relevant information outlining how the force should approach the use of AI technologies.
 - Procedures for different types of AI work.
 - The levels of authorisation for different types and categories of AI work
 - Links to relevant documents and guidance.

3. WHAT SHOULD I CONSIDER WHEN USING THIS POLICY?

It is intended that the process outlined in this AI Ethics policy should start at the initiation phase of any programme or project that will make use of AI technology, or as soon thereafter as it is ascertained that a project will involve the use of AI technology.

The terms **piece of work**, **proposal** and **project** will be used to refer to the project, programme, or part thereof, for which this policy is being consulted.

The term **process** will refer to the **activities outlined herein**.

This policy should be used in conjunction with the AI ethics code, tick list, and guidance notes:

3.1 The AI Ethics Code

The entire ethics code can be found in Appendix D but the main headings are summarised here to give context:

Summary of Code of Ethics

*When considering the initial, or continued, use of **AI** (Predictive Analytics, Machine Learning, Artificial Intelligence) or other data reliant technologies, in order to maximise the benefits and minimise the associated risks, it is imperative that it not just be ethical, but that it be seen to be ethical i.e. that it be demonstrably trustworthy.*

*To be trustworthy a AI technology must have three qualities, it must be: **BENEVOLENT**, **ACCOUNTABLE**, and **ROBUST**.*

To demonstrate these qualities there is a minimum set of principles that must be upheld.

TRUSTWORTHINESS

To be trustworthy a AI technology must be benevolent, accountable, and robust:

1. Benevolent

- a. Seeks to do good (BENEVOLENCE: Purpose)
- b. Is free from prejudice (BENEVOLENCE: Fairness)
- c. Respects privacy (BENEVOLENCE: Privacy)

2. Accountable

- a. Adheres to the law and regulations pertaining to its use (ACCOUNTABILITY: Lawfulness)
- b. Must not deceive, deliberately or otherwise (ACCOUNTABILITY: Honesty)
- c. Is subject to continued external oversight (ACCOUNTABILITY: Scrutiny)
- d. The method for decision making can be explained (ACCOUNTABILITY: Transparency)

3. Robust

- a. Is reliable (ROBUSTNESS: Reliability)
- b. Keeps data secure (ROBUSTNESS: Security)
- c. Corrects bias (ROBUSTNESS: Unbiased)

The principles derived from the qualities required for trustworthiness can be further broken into several tenets to test the ethics of any piece of work connected to AI technology. The breakdown can be found in the Appendix D.

3.2 The AI Ticklist can be found in Appendix E

3.3 The AI Use Case and Guidance can be found Appendix F

This policy, the AI ethics code, Ticklist, and guidance notes will be available in both English and Welsh versions.

4. ROLES AND RESPONSIBILITIES

The authorisation hierarchy can be seen in Appendix A

4.1 AI Ethics Process Initiator	The person or group who initiate the AI Ethics policy process with a proposal for a project or programme that involves collaborating on, using, creating, renting, or purchasing some element of AI technology
4.2 AI Facilitator	Carries out first line checks that the process initiator has considered the ethicality and full ramifications of the proposed work and can authorise some work (up to level 2). The AI Facilitator will be able to speak Welsh at level 3 or above.
4.3 AI Ethics Group (AIG)	Has oversight of the work carried out by the initiator and facilitator, and can authorise some work (up to level 3)
4.4 Assurance Board	Has oversight of the work carried out by the initiator and facilitator, and AIG and can authorise some work (up to level 4)
4.5 Senior Management Board	Overall authorisation of high risk, complex, and collaborative projects (level 5)

AI projects generally fall into one of four categories, each being associated with different levels of risk. The level required to authorise each type of work is shown in the diagram below, along with a list of criteria that define what fits into each category:

Hierarchy of Authorising Bodies
Lowest → Highest

Risk	Accountability Authority		Project Initiator	PAMLAI Project Worker	AI Group	Assurance Board	SLT
Low	Ad hoc	Small scale	A				
		and exploratory					
		and temporary					
		and simple					
		and only certified data from the datawarehouse					
	Operationalised	and non-operational					
		and not shared beyond author					
		or Procedural assistance required					
		or Data is uncertified or from outside data warehouse		A			
		or Shared with manager or team mates					
Medium	Operationalised	Any operational use			A		
		or Requires monitoring					
		or Non trivial scale					
		or Results shared beyond manager team					
		or Some complexity					
	Acquisitions	or Escalated by previous hierarchy					
		Third (3rd) party services				A	
		or Third (3rd) party products					
		or Third (3rd) party software					
		or Proprietary test data					
High	Collaborative	or Any reputational risk					
		Significant reputational risk					A
		or Third party data collaboration					
		or Third party analytical collaboration					
		or legal or financial risk					
		or high complexity					
		or large scale					

5. THE PROCESS

At the outset of every programme or project a self-assessment must be made by the originator of the piece of work as to whether it will use any AI technology. If it will then, the AI Ethics policy process (hereafter referred to as the process) will need to be followed. Likewise, if a publically available tool uses AI technology then the process should also be followed.

The process begins with an assessment, by the project's originator, of the category of work being undertaken. The authorisation hierarchy matrix should then be used to determine the level of authorisation required. Assistance may be sought from the AI facilitator in this assessment.

The request will flow through the authorising bodies as in Appendix A. the process is designed to be iterative to enable feedback to be incorporated in order to facilitate the authorisation of acceptable proposals

If authorisation is required or desired, then the piece of work will be registered using the AI registration form (see Appendix G)

The process should be followed for each part of the proposed project that uses a different AI technology or that uses the same types of AI technology for different tasks.

e.g. a large programme involving text analysis for summarising statements, text analysis for sentiment analysis, facial recognition should assess the ethicality of each technology and use combination i.e. 3 items to consider in this example.

The AI Ethics Policy Process begins once it is noticed that there will be some element or use of AI technology involved in any given project.

The actions for each role in the process are listed below. A detailed visual process map can be found in Appendix C

5.1 AI Process: Initiation

Role	Action Required
5.1.1 Proposal Initiator	Use the AI Ethics code and Ticklist to determine whether the work should proceed, and if so, then register the proposed work using the online form (see Appendix G)
5.1.2 Proposal Initiator	Use the AI Authorisation Matrix to determine whether authorisation is required.
5.1.3 Proposal Initiator	Ask the AI Facilitator for guidance on how to use the Ticklist if desired.
5.1.4 Proposal Initiator	If guidance is desired, submit a request to the AIG using the online form with a partially completed AI Ticklist.
5.1.5 Proposal Initiator	Use the AI Authorisation Matrix to determine the required authorisation level.

5.1.6	Proposal Initiator	Submit the completed AI Ticklist to the AIG.
5.1.7	Proposal Initiator	Await response
5.1.8	Proposal Initiator	Amend the proposal as required by a more senior authorising body.
5.1.9	Proposal Initiator	If authorised, proceed with the proposal under the oversight of the AIG

5.1 AI Process: **Guidance**

Role	Action Required
5.1.10 AI Project Facilitator	Upon receipt of a request provide guidance to the proposal initiator.
5.1.11 AI Project Facilitator	Use the AI Ethics code and Tick-list to determine whether the work should proceed.
5.1.12 AI Project Facilitator	Use the AI Authorisation Matrix (Appendix B) to determine what level of authorisation is required or refer to the AI facilitator.
5.1.13 AI Project Facilitator	Recommend required changes to the project proposal.

5.2 AI Process: **Oversight, Authorisation, or Escalation**

Role	Action Required
5.2.1 AIG (Artificial Intelligence Group)	Each member of the AIG will interpret the AI Ethics code and Ticklist in the light of their own domain of expertise (legal, technical, operational, compliance etc) and the AI guidance notes to determine whether the request should proceed.
5.2.2 AIG	Where there is difficulty in determining whether the request should proceed, the AI Board should gather the opinions of appropriate experts, and use the insights gleaned to interpret the AI Ethics code and Ticklist to determine whether the request should proceed.
5.2.3 AIG	Where necessary request amendments from the Proposal Initiator
5.2.4 AIG	If, based on the AI Authorisation Matrix (Appendix B)or situational judgement, a higher level of authorisation is required or desired, forward the proposal, Ticklist and notes to the Assurance Board

5.3 AI Process: **Authorisation or escalation**

Role	Action Required
5.3.1 AB (Assurance Board)	In the light of the information gleaned from the AIG investigation, AI Ticklist, Ethics Code, and Guidance Notes, determine whether the work should proceed.
5.3.2 AB	Where there is difficulty in determining whether the work should proceed, request AIG gather further information from appropriate experts and resubmit once the fresh information is gathered
5.3.3 AB	Where necessary request amendments from the Proposal Initiator
5.3.4 AB	If the proposal is fundamentally unsound, end the proposal

5.3.5 AB	If based on the Authorisation Matrix, or situational judgement a higher level of authorisation is required or desired, forward the proposal, Ticklist and notes to the Senior Management Board (SMB)
----------	--

5.4 AI Process: Authorisation

Role	Action Required
5.4.1 SMB (Senior Management Board)	In the light of the information gleaned from the AB, AIG investigation, AI Ticklist, Ethics Code, and Guidance Notes, determine whether the work should proceed.
5.4.2 SMB	Where there is difficulty in determining whether the work should proceed, request AIG gather further information from appropriate experts for resubmission once the fresh information is gathered
5.4.3 SMB	Where necessary request amendments from the Proposal Initiator
5.4.4 SMB	If the proposal is fundamentally unsound, or undesirable, then reject the proposal
5.4.5 SMB	Authorise the work

6. DECLARATION & LEGALITIES

- 6.1 In line with all Force policies, the overarching purpose of this document is to directly support the PCC police and crime plan objectives. Overall, the intention of this policy is to make North Wales the safest place in the UK.
- 6.2 In the writing of this policy cognisance has been taken of the college of policing code of ethics (2014) and proposed revisions (2023).
- 6.3 North Wales Police policies will be written in accordance with the approved corporate format and published on the Force Intranet, allowing access to staff and public, where appropriate, on the pages of the public facing Internet site under the Force publication scheme and Freedom of Information Act 2000.
- 6.4 The following main legal requirements and regulatory have been identified within this policy:
- Equality Act 2010
 - Human Rights Act 1998
 - The Welsh Language (Wales) Measure 2011 and the Welsh Language Standards for the Chief Constable
 - Data Protection Act 2018 (schedule 8) (UK GDPR)
 - Freedom of Information Act 2000
 - Health and Safety Act 1974
 - The General Data Protection Regulation 2016/679
 - Data Protection and Digital Information (No. 2) Bill March 2023
 - CDDO Algorithmic Transparency Recording Standard
 - European Convention on Human Rights (ECHR)
 - EU Artificial Intelligence (AI) Act 2021
- 6.5 Recommendations from following bodies have been consulted for this policy:
- Centre for Artificial (CAIDP) and Office for Artificial Intelligence
 - CAIDP, UNESCO, ELI CDDO, CDEI recommendations
 - European Law institute (ELI)
 - Central Digital and Data Office (CDDO)

- Centre for Data Ethics and Innovation (CDEI)
- Privacy international
- National Institute of Standards and Technology (NIST)
- Royal United Services Institute (RUSI)
- National Police Chiefs Council (NPCC)
- College of Policing (CoP)

6.6 This policy has been written giving due regard to the above legislation and has considered the risk of unfair and/or disproportionate impacts on individuals or groups (actual or perceived) and has done so via an equality impact assessment (EIA) and a Welsh Language Impact Assessment (WLIA).

6.7 The area of AI ethics laws, regulations, and recommendations is very active, and changes are expected to be frequent. Changing legislation and deeper understanding of the topic will necessitate regular reviews of this policy document to ensure continued compliance and best practise.

7. APPENDICIES

7.1 Appendix A – Authorisation Levels

The proposal will be considered by the following bodies in the order shown below to the level required according to the Authorisation Matrix (Appendix B). The steps in yellow are only required in some cases.



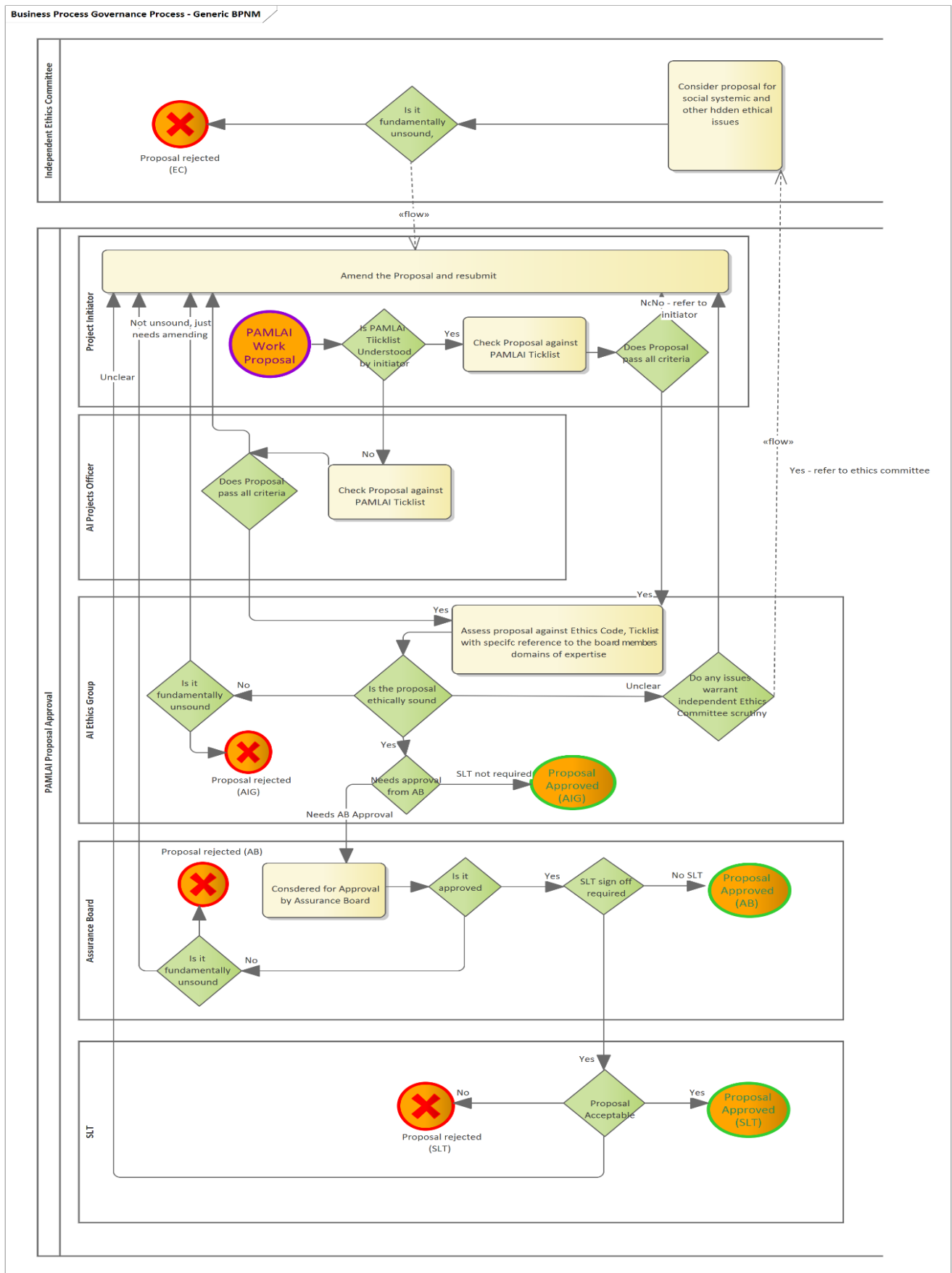
7.2 Appendix B - Authorisation Matrix

Hierarchy of Authorising Bodies

Lowest → Highest

Risk	Accountability Authority		Project Initiator	PAMLAI Project Worker	AI Group	Assurance Board	SLT
Low	Ad hoc	Small scale	A				
		and exploratory					
		and temporary					
		and simple					
		and only certified data from the datawarehouse					
		and non-operational		A			
		and not shared beyond author					
		or Procedural assistance required					
		or Data is uncertified or from outside data warehouse					
		or Shared with manager or team mates					
	Operationalised	Any operational use			A		
		or Requires monitoring					
		or Non trivial scale					
		or Results shared beyond manager team					
		or Some complexity					
		or Escalated by previous hierarchy					
	Acquisitions	Third (3rd) party services				A	
		or Third (3rd) party products					
		or Third (3rd) party software					
		or Proprietary test data					
		or Any reputational risk					
	Collaborative	Significant reputational risk					A
		or Third party data collaboration					
		or Third party analytical collaboration					
		or legal or financial risk					
		or high complexity					
High		or large scale					

7.3 Appendix C – Detailed Process Map



7.4 Appendix D – AI Ethics Code

Embedded File



PAMLAI Ethics Code
(PEC) v0.4 - English .

7.5 Appendix E – AI Ticklist

Embedded File



PAMLAI-Ticklist
(PET) v0.3 - English.c

7.6 Appendix G – Sample Registration Form



Example of
registration form - fill

7.7 Appendix H – Terminology

The following abbreviations are used

Term	Meaning
AI technology	Predictive Analytics', Machine Learning, Artificial Intelligence, and related technologies
SMB	Senior Management Board (highest authorising body)
AB	Assurance Board – policy based decision making board
AIG	Artificial Intelligence group - Deals with AI matters including the ethicality of proposals
PF	AI facilitator – AI / Ethics expert capable of providing guidance to those using the AI Ticklist
PI	Proposal initiator – the individual or group making the proposal

AI, Ethics & Policing in North Wales: Establishing Citizens' Perspectives

Technical Project Report (Final Draft)

January 2025

Dr. Bethan Loftus
Dr. Alexander Laffer
Prof. Martina Feilzer
Prof. Vian Bakir
Prof. Andrew McStay



Contents

1. Introduction	3
1.1. Aim	
1.2. Focus	
1.3. Context	
2. Artificial Intelligence for Policing Purposes	5
2.1. Intelligent Facial Recognition	
2.2. Predictive Policing and Data Analytics	
3. Legal Context	8
4. Unpacking Concerns: A Review of the Literature	10
5. Research Methodologies and Approach.....	15
6. Findings	21
6.1 Citizen Focus Groups.....	21
6.2 National Survey	29
6.3 Stakeholder Interviews.....	39
6.4 Interviews with Police Personnel	45
7. Preliminary Recommendations	47
8. References	49

1. Introduction

This report presents key findings and recommendations from a research study entitled: *AI, Ethics & Policing in North Wales: Establishing Citizens' Perspectives*. The research was initiated by personnel within North Wales Police, and was conducted by a team of academic researchers from Bangor University.¹ Funded by a Bangor University Impact and Innovation Award (BU-IIA), the project set out to explore the future use of AI-technologies in policing. In particular, it sought to: *identify and explore the benefits, challenges and ethical concerns raised through the police use of new computational technologies, most notably those involving claims to detect and predict criminality through risk-modelling- often labelled **Predictive Policing**, and those based on **Intelligent Facial Recognition*** (Fussey & Murray, 2019; Hamilton 2021).

To explore this, we started with the following overarching research question:

With respect to civic life outside of major cities, what ethical issues are raised through the police use of Live and Retrospective Facial Recognition as well as other profiling technologies, and what are the views of the public, police and other relevant stakeholders with respect to the deployment of such technologies?

The research comprised a mixed-methods approach, with the findings emerging from the collection and analysis of four data sets:

- North Wales Citizen **Focus Group Discussions**
- A national demographically representative **Survey**
- **Interviews** with expert stakeholders from regulatory and civic organisations
- **Interviews** with serving police officers and employees of North Wales Police

The research was viewed as particularly important and timely because it attended to three areas that currently require more attention, both by police organisations, policy makers, and academics:

- Public understanding of, and feelings towards, the real-world implementation of AI technologies in the context of policing (Ada Lovelace Institute, 2019; Hintz et al., 2022). Ongoing debates are concerned with how AI draws upon and interacts with sensitive data – such as biometrics – and how it is used within policing contexts where it can have life-changing consequences.
- Whether the use of AI – and public receptiveness towards it – differs outside of major urban centres that are densely populated. There is a need to explore AI within rural settings, such as villages and small towns, where the application and attendant concerns surrounding AI technology may differ from those applying to urban environments.
- Although there is an established body of research on Retrospective Facial Recognition (RFR), we still have little understanding of how Live Facial Recognition (LFR) may be deployed by those who use it, how it is understood by those on the receiving end, and the potential ethical dilemmas for its implementation. This

¹ Principal Investigator: Dr. Bethan Loftus, Reader in Criminology and Criminal Justice, Bangor University. Email: b.loftus@bangor.ac.uk.

project sought to better comprehend these concerns and, in so doing, supporting the development of up-to-date ethical guidance for the use of such technologies.

There are proclaimed benefits for the use of AI-technologies and predictive methods in policing. Such technologies can, for instance, significantly enhance a range of police activities – such as enabling the rapid location of missing and vulnerable people, and the identification of suspects. While digital rights groups have argued that the deployment of AI and allied surveillance technologies prompts human rights abuses and scope for creation of complex historically oriented profiles of citizens (Woodhams 2021), the Biometrics and Surveillance Commissioner has argued that there will always be circumstances in which the use of Facial Recognition – and AI more broadly – should be deployed, giving the example of a recent terrorist attack in New York.² Yet, as the Scottish Biometric Commissioner made clear, the decision to utilise AI technologies should be informed by a ‘consideration of what *should* be done, as opposed to merely considerations of what *can* be done’ (2018).

Therefore, at the heart of our research is a consideration of both the proclaimed benefits of AI-technologies within policing against its potential negative and unethical qualities. If the benefits of AI in policing contexts are clear – and there is a certain level of inevitability to its routine deployment across police organisations – then what might an ethical and acceptable implementation of this technology look like to reduce the risk of inadvertent harm done? What governance, oversight, and consultation processes ought to be put into place to ensure that AI is used correctly, ethically, and fairly?

This report unfolds in the following manner: In Section 2, we begin by defining the key terms to be used in our discussion, and we introduce an overview of the technologies under scrutiny. This includes providing a brief coverage of the ways that these have been used and could potentially be used. In Section 3, we then provide a background to the current legal context in England and Wales, with the important caveat that this is a dynamic and fast-changing area with legal challenges to the technology – such as the high-profile *R (Bridges) v South Wales Police*, as well as changes in regulation. For example, at the time of conducting our research, the Data Protection and Digital Information Bill was being read by parliament which could have impacts on the regulatory responsibility for these technologies. We provide a contrast with Scotland, where a recent code of Practice by the Scottish Biometrics Commissioner has been approved as Scottish Law. In Section 4, we present a brief coverage of the main concerns emerging from recent academic literature and relevant stakeholder reports on police adoption and implementation of AI-technologies. Section 5 provides an overview of the mixed-method approach that we adopted during the empirical investigation, including an insight into the survey questions and design fictions presented to the focus group participants. Section 6 presents the key findings based on the thematic analysis of the datasets, before outlining several recommendations for consideration by North Wales Police.

² Stated at the annual CRISP lecture at Sterling University (see also Fussey and Webster 2023).

2. AI for policing purposes

2.1 Artificial Intelligence

The focus of the project is on **Artificial Intelligence** (AI), which refers not to 'a specific technology', but rather, 'a collection of interrelated technologies used to solve problems and perform tasks that, when humans do them, requires thinking' and that can handle 'uncertainty' and 'ambiguity' (Walsh et al., 2019, pp. 14-15).

The implementation of AI is often considered desirable because it promises to make, 'human decisions smarter, more efficient and more rational' when data and information is 'becoming increasingly complex, voluminous and fast changing' (Chan, 2021).

Although the concept of Artificial Intelligence has gained currency in recent times, Babuta and Oswald (2019, p. 4) prefer the term **Machine Learning**, arguing that AI is 'ambiguous' and 'poorly defined'. Machine learning better encapsulates the technologies employed for policing purposes, including 'facial recognition and video analysis; mobile phone data extraction; social media intelligence analysis; predictive crime mapping; and individual risk assessment.' However, as machine learning can be viewed as a subset or element of artificial intelligence (Walsh et al., 2019, p. 14), in this report we use the more familiar and broader term of Artificial Intelligence (AI).

Following Veale and Brass (2019), it may be useful to categorise AI uses based on purpose, i.e. as **automation systems** which use computation to improve the efficiency or increase scope of operations, and **augmentation systems** which are used to improve human decision making.

It is also important to recognise that the introduction of AI represents not just a technological shift but a cultural one (Crawford & Calo, 2016, p. 313), which will have far reaching, and potential unforeseen, consequences on how people and institutions behave. For example, the successful implementation of AI is largely built on access to high volumes of (sometimes highly personal) data, leading to what van Dijck (2014, p. 198), building on Mayer-Schönberger & Cukier (2013), describes as datafication: 'the transformation of social action into online quantified data', which is 'gradually becoming normalised in science and society' affording **real-time tracking** and **predictive analysis**. Later in this report, we will explore some of the concerns surrounding this shift.

In our research, we specifically focus on two forms of Artificial Intelligence that can be mobilised for policing and security purposes, that we shall describe in the following two sections:

- a) Intelligent Facial Recognition
- b) Predictive Policing

2.2 Intelligent Facial Recognition

Intelligent Facial Recognition refers to the use of biometric technologies that draw on algorithms to identify people – usually by their physical features or characteristics – in a video or a static image (Fussey & Murray, 2019). This might be used for: **verification**, performing a one-to-one comparison of a new and stored image; **identification**, a one-to-many comparison of a new image to a database of images; and potentially for

categorisation of individuals, for example into gender or ethnic categories, based on physical characteristics, (Madiaga & Mildebrath, 2021, p. 7), with arguably increasing levels of concern associated with each process. Furthermore, there is potential for the technology to classify facial expression or emotional states, which we term **emotional AI**.

We focus here on **identification** as most salient to the current police activities in which NWP may employ intelligent facial recognition. Identification, and similarly related processes, can be done through a range of different technologies, but are generally achieved through the automated creation of a **faceprint** – a unique map of someone’s face made by measuring the distances between key facial features. Figure X below provides a simplified illustration of what this process might look like, with a faceprint created for a new image and which can then be compared to an existing database of stored faceprints:

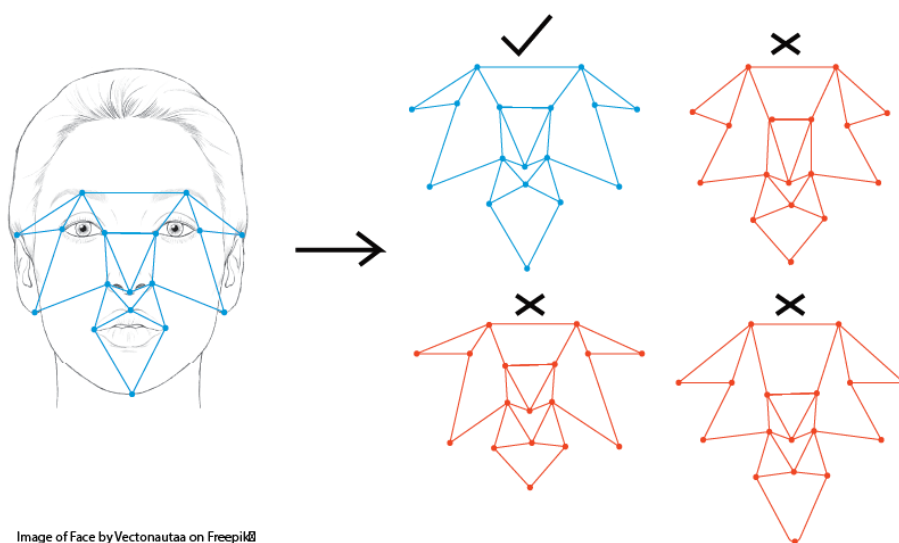


Image of Face by Vectonautaa on Freepik®

Figure 1: Illustration of facial recognition process.

Most uses of intelligent facial recognition for identification purposes can be characterised under two forms:

(1) When facial recognition is performed on previously saved images (such as police custody photographs or CCTV footage), it is known as **Retrospective Facial Recognition**. This compares a prior collected images (such as CCTV footage) of an individual against a database of stored images. The AI provides a likelihood of a match based on a comparison of faceprints.

(2) **Live Facial Recognition** makes a live reading of people traversing in front of a camera, and compares images to a specific and pre-determined list of those sought by the police (Bradford et al., 2020). It involves cameras being strategically placed in public locations for set periods of time to record footage of people as they pass by in real-time. As soon as the footage is recorded, software identifies any faces in the video and produces a unique ‘faceprint’ to be compared to other faceprints from a specific watchlist of people the police are searching for. An officer working with the software checks the match and decides whether to advise officers in the vicinity of the camera to approach the identified person. Only footage containing potential matches to people on the watchlist are stored and all other faces are pixelated. Any footage containing no potential matches is deleted.

There is an additional variation of facial recognition, **Operator Initiated Facial Recognition** (OIFR) (Radiya-Dixit, 2022). In these instances, an officer actively uses a mobile camera to collect an image of a specific individual which is then sent immediately for facial recognition processing.

While access to Retrospective Facial Recognition is possible for many police forces through the centralised National Police Database, which stores custody images of those arrested by police, a few forces have begun to trial Live Facial Recognition³, often without accompanying meaningful public engagement and debate (Ada Lovelace Institute, 2022; Hintz et al., 2022), and have been challenged based on their breaching of privacy, data protection and equality laws.

While the focus of regulation and challenges to the technology is often directed at Live Facial Recognition, there has been more recent concerns that the distinction between Retrospective and Live Facial Recognition is not meaningful as both can have 'equally negative impact on fundamental rights' (Madiaga & Mildebrath, 2021, p. 34). Indeed in 2012 a High Court ruled that the automatic retention of images by the police was unlawful under Art 8 of the ECHR, the right to private life, and that any such images should be destroyed.⁴ There is some ambiguity of whether OIFR should be considered as Live or Retrospective. While we might want to consider it a targeted form of live facial recognition, they could be classed as retrospective depending on the delay in activation. It would be of concern if an arbitrary wait time was instigated, so that they could be treated as retrospective in terms of regulation.

While the proposed EU Artificial Intelligence Act argues that the delay in 'post' or retrospective systems mitigates the risk of using facial recognition, examples from the Georgetown Law investigation into Facial Recognition (Garvie, 2019) demonstrate how the delay can actually provide an opportunity for misuse, for example, replacing collected images where no hits were found with photos of similar looking celebrities from the internet, and using the results as the basis for identifying a suspect.

2.2 Predictive Policing

Through a synthesis of research, Meijer & Wessels (2019, p. 1033) provide a useful definition of predictive policing as:

'...the collection and analysis of data about previous crimes for identification and statistical prediction of individuals or geospatial areas with an increased probability of criminal activity to help developing policing intervention and prevention strategies and tactics.'

Put simply, police forces may use AI to forecast who might be involved in (or affected by) criminal activity or where it might occur by checking for connections and patterns across large amounts of data, more than could be analysed by a person. This might include information such as a person's age, gender, and postcode, as well as details on individuals' and groups' prior association with crime (as offenders or victims) and broader historical crime statistics. The chief purpose being to find targets for intervention, whether through the deployment of officers to a particular location or to make a risk assessment on an

³ A list of potential mobilisations is given by Big Brother Watch, in the section 'UK Facial recognition in detail' on their website: <https://bigbrotherwatch.org.uk/campaigns/stop-facial-recognition/>

⁴ [2012] EWHC 1681 (Admin), RMC and (2) FJ - and - Commissioner of Police of the Metropolis

individual's likelihood to commit a crime and so decide what punitive or rehabilitation measures to employ. It is important to re-iterate that these are predictions, based on analysis of probability and do not provide certainties (Podoletz, 2022; Strikwerda 2020). Actuarial practices and risk assessments predicting individual risk of future/further offending have been used extensively in the probation service and youth justice system, yet predictive accuracy has been shown repeatedly to be low (Vaswani and Merone 2014).

These types of predictive processes are often supplied through platforms, such as those provided by Palantir and Geolitica (previously known as PredPol) – the latter was implemented by Kent Constabulary between 2013 and 2018, for example, although it is no longer in use as it had shown a limited effect on crime reduction. Until recently, the actual use of predictive mapping software by UK police forces has been limited, often just short term trials without full-scale deployment (Babuta & Oswald, 2019, p. 6) and 'a relatively small number of UK police forces (14%) were using computational or algorithmic data analysis or decision-making in relation to the analysis of intelligence' (Oswald et al., 2018, p. 224). However, this is a dynamic landscape, with Hamilton (2021) pointing to the development of the National Analytics Solution (NAS), led by West Midlands Police on behalf of the Home Office, a 'new, scalable and flexible analytics capability for UK law enforcement using advanced analytics to deliver insights to partners on agreed high priority operational and organisational issues', which has the potential to merge datasets across different departments but also across multiple policing agencies.

As with facial recognition 'the use of algorithms to make predictions about future crime and offending raises considerable legal and ethical questions, particularly concerning the risk of bias and discrimination.' (Babuta and Oswald, 2019, p. 5).

3. Legal Context

It is not within the scope of this report to provide in-depth coverage of the legislation governing the use of AI-technology as well as those that cover biometric data use. For a good overview of regulation in the European context Madiaga and Mildebrath's (2021) report for the European Parliament research service is recommended, and more specifically for England and Wales, the recent Ryder Review (2022), commissioned by the Ada Lovelace institute, provides coverage in relation to biometric data, and so is directly relevant to Intelligent Facial recognition.

However, it is worth providing a brief overview of *some* key pieces of legislation, both to foreground the main concerns, from a legal basis, raised in the deployment of Facial Recognition and Predictive Policing technologies, and to present the current jigsaw of guidance that ought to be navigated by police forces considering the use of these technologies. These are:

Human Rights Act (1998)

Particularly:

- Article 5: Right to liberty and security
- Article 8: Respect for your private and family life, home and correspondence
- Article 11: Freedom of assembly and association
- Article 14: Protection from discrimination in respect of these rights and freedoms

Data Protection Act and General Data Protection Regulation

Personal data collection and use must follow these principles:

- used fairly, lawfully and transparently
- used for specified, explicit purposes
- used in a way that is adequate, relevant and limited to only what is necessary
- accurate and, where necessary, kept up to date
- kept for no longer than is necessary
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage

Equality Act 2010

- In particular, section 149 public sector equality duty

Investigatory Powers Act 2016

- Notably, the general duties it sets out in relation to privacy and storage of information

Protection of Freedoms Act 2012

- Particularly, provision in respect of the destruction, retention, and use of biometric data.

This complexity in multiple legislation is mirrored in regulatory responsibility, with ongoing and shifting responsibilities between regulators, such as the Information Commissioner's Office (ICO) and the Biometrics and Surveillance Camera Commissioner (BSCC). This will be explored later in the report which includes discussion of an interview carried out with the BSCC for the project.

While there is a 'lack of explicit legal basis' (Radiya-Dixit, 2022, p. 25) in England and Wales, we might look to the recent publication by the Scottish Biometrics Commissioner of a Code of Practice becoming Scottish law via a Scottish Statutory Instrument (SSI). There have been calls for specific legislation to regulate the use of facial recognition technology in England and Wales (Fussey & Murray, 2019).

We also might consider the forthcoming EU Artificial Intelligence Act as a piece of legislation with the purpose of defining and constraining the use of AI-technologies that takes a risk-based approach. Currently, this categorises 'realtime' remote biometric identification (RBI, which includes Live Facial Recognition) as high risk, proposing that they should be banned because they 'are particularly intrusive, [and] severely interfere with the rights and freedoms of the persons concerned.' (Madiaga and Mildebrath, 2021, p. 31) Although certain exceptions are in place, including, deploying the technology to search for potential victims of crime; prevention of a substantial and imminent threat to life, such as a terrorist attack; or in cases involving suspects referred to in the European Arrest Warrant Framework Decision. There is also acknowledgement that EU Member states have a degree of control over national security issues, which the use of RBI systems might impact on.

However, we must be aware of the concerns of some scholars (cf. Bradford et al. 2020) that the use of predictive policing technologies might potentially erode human rights that are currently enshrined in law.

4. Unpacking Concerns: A Review of the Literature

In this section we address some of the key concerns that emerge from the literature on AI, policing, and surveillance, including insights from recent reports that have addressed the usage of intelligent facial recognition and predictive policing. In this section, we aim to address and potentially temper what Chan (2021; chapter 1) calls the 'promise of AI', the potential to make 'human decisions smarter, more efficient and more rational' when 'information has become increasingly complex, voluminous and fast changing'.

Techno-solutionism and the 'rush' towards technology

There is a concern that broader discourses of techno-solutionism – i.e., where current issues in systems and processes may be resolved with ever more powerful processing and technological advances – are driving the adoption of algorithmic and AI-based solutions. As Chan (2021) warns, we should resist overly optimistic perspectives of technology if they gloss over or obscure potential harms.

A lot of research assumes that police are rushing to use this technology (see Best 2010). This assumption is arguably born out by rates of adoption (particularly in the United States, (Garvie et al., 2016) as well as the purchase and use of surveillance and predictive policing platforms by police forces (Nilsson, 2018).

This is particularly pertinent when considering the cost of technology which needs to be balanced against other resourcing needs. There is a concern in the literature that budgeting pressures are driving this adoption, particularly due to austerity measures (Babuta and Oswald, 2019; Ratcliffe and Kikuchi, 2019; Hamilton, 2021), raising concerns that these solutions would not be necessary if adequate resources were already in place.

More broadly, there are concerns connected to the unreflexive and uncritical use of data. Hintz et al. (2022) raise an important point that data should be seen as social data and understood contextually. Big Data is not objective, nor simply a neutral collection of naturally occurring data, and its algorithmic analysis reflects the ideologies of those who commission and construct the algorithm.

Privacy

As might be expected, privacy is a chief concern in much of the literature on algorithmic systems both in terms of the intrusion into people's activities but also in terms of the use of data (Radiya-Dixit, 2022) and the impact it may have on people's anonymity (Podoletz 2022), the potential loss of control of personal sensitive information (Kindt 2013) as well as the potential to track people (Campisi 2013). Bradford et al. (2020, pp. 1515-16), suggests it is the 'the primary factor motivating acceptance or rejection of police use of LFR is concern about privacy'.

We might follow Smith and Miller (2022) in their distinction between *informational privacy* (the possession of information about people by other persons and organisations), and *observational privacy* (of movements, relationships and behaviours) by surveillance systems. Arguably, this is heightened through the use of remote biometric identification, where everyone passing in front of a camera is scanned even if only individuals are being searched for (EDRi, 2023).

Privacy is connected with virtues and ideas of autonomy, with intrusions into the private sphere impacting on peoples' freedom and right to decide what to think and do, through the act of observation (Smith and Miller, 2022, p. 172).

Accuracy and discrimination

In both academic and civic literature, there are concerns around the accuracy of the predictions made by these technologies, with an emphasis on the potential for discrimination (*Biometrics and Surveillance Camera Commissioner*, n.d.; Buolamwini & Gebru, 2018; Centre for Data Ethics and Innovation, 2020; Hintz et al., 2022; Madiaga & Mildebrath, 2021; Podoletz, 2022; Radiya-Dixit, 2022) which has legal repercussions.

While accuracy rates are improving, it is worth highlighting a headline figure from research performed by Big Brother Watch (reported in Hintz et al. 2022) that claimed there was a 95% mismatch rate in terms of people stopped by police after being identified though LFR actually being someone wanted by the police.

Errors in algorithms, particularly triggered by a lack of representation within training data (Buolamwini & Gebru, 2018; *Facial Recognition Technology*, 2020), are identified as causing AI-systems to be less accurate in the treatment of non-white and non-male faces, with Buolamwini & Gebru highlighting that all the systems they evaluated were based on binary distinctions of sex, failing to address the complexities of gender and transgender identities.

Further, bias may be introduced into data, even if it is representative, through the process of labelling it for training purpose, again leading to algorithmic discrimination (Bolukabasi et al., 2016; Casikan et al. 2017). As Buolamwini & Gebru (2018) argue, 'monitoring phenotypic (observable physical characteristics) and demographic accuracy of these systems...is necessary to protect citizens' rights and keep vendors and law enforcement accountable to the public.'

The Centre for Data Ethics and Innovation (2020) claim that these issues of bias are not intrinsic or inevitable features of these algorithms and, indeed, may exist in manual processes. However, they do recognise that 'any bias or inaccuracy takes on greater significance in the context of policing' (p.23). This is further problematised if the systems are developed in a blurred Public-Private partnership, with potentially less or even restricted scrutiny of training datasets and algorithm testing (*Biometrics and Forensics Ethics Group*, 2021, p.6)

It is important to note that while it is possible to argue the above issues could be resolved through improvements to the technology, we need to question the use of the technology in terms of wider practices around surveillance of particular groups. As Radiya-Dixit (2022) argues, improvements in these systems may actually deepen existing patterns of discrimination by rendering marginalised communities 'hypervisible' and increasingly subject to further surveillance. There have been calls for controls when putting people with

protected characteristics on watchlists (College of Policing 2021, p.21). However, any use of the technology then needs to be considered in light of prior policing and data collection practices as well as current police cultures.

Police Cultures

Commenting on algorithmic technologies, Houwing (2020, p. 175) argues that 'as long as technology that is built to identify, profile and analyse people in order to treat them differently is deployed by people in societies that suffer from systemic inequality, the technology will most likely reinforce and exacerbate those inequalities'. It is, therefore, vital that any potential implementation of new technologies recognises any dysfunctional or systemic issues within those organisations that seek to use them. For instance, the use of facial recognition systems examined across 100 American police departments have reinforced existing practices within conventional policing, with African-American individuals more likely than any other ethnicity to be stopped and subjected to facial recognition searches (Garvie et al. 2016).

While this should be addressed internally to provide the foundation for equitable and fair use of the technology, it is important to stress that this is also a *perceptual* issue, with increased scrutiny on officer behaviour and wider systemic issues. Undesirable experiences of, and scandals in, policing can make their way into the public consciousness and have the potential to influence perceptions about the police use of AI and related technologies.

Where these two elements combine is at community level interactions between the police and citizens, with judgments about individual police officers being a more important predictor of acceptance of technology use rather than concerns about crime. As Bradford and colleagues explain, because of the strong effect of individual interactions with police officers (whether via direct interaction or vicarious events), 'the mundane activity of everyday police work ('low' policing) can have important implications for the public acceptance or rejection of novel technologies working towards the 'high' end of policing.' (Bradford et al., 2020, p. 1517).

This also raises important questions around the need for police organisations to engage in public consultation and participant recruitment when exploring public views on the use of technology. People engaging directly with police and providing feedback through police instigated processes will likely have a positive, or at least neutral, response within community level policing.

Transparency

The theme of transparency extends across:

- 1) The development of algorithmic technologies and the training of systems (for example, Buolamwini & Gebru, 2018)
- 2) The procurement of such technologies (UN High Commissioner for Human Rights, 2020)
- 3) The implementation and use of the technology (Hintz et al., 2022; Radiya-Dixit, 2022)

Central to the first issue is **intelligibility**, with even technological experts potentially struggling to fully explain how results are arrived at due to the number of variables and complex process. This is known as the 'black box problem' (Chan 2021) and needs to be considered in tandem with the Ada Lovelace Institute (2019) which conducted a national survey (n=4109) of public attitudes towards Facial Recognition. The Institute found that while general awareness of facial recognition technology is high, knowledge about what exactly it is, and how it works, was very low. This suggests that more needs to be done to render this technology transparent and better inform and educate the public. This is an important factor in the consideration of informed consent: the level of understanding required by those citizens subject to its usage. For example, civil rights organisations argue that a lack of transparency about AI and predictive policing systems prevents a meaningful, well-informed public debate, and undermines processes of consent.

These issues are heightened when the technology is procured and there is any degree of privatised corporate secrecy, resulting in private companies failing to be fully accountable for the claims they make about the accuracy of facial recognition products or the implications for individuals' rights leading to further reduction in opportunities for citizens to seek redress (UN High Commissioner for Human Rights, 2020).

Transparency is a key prerequisite for processes of accountability. If those using the technology – and indeed those subjected to it – lack clear understanding of its operation, capabilities, and processes, then it limits the ability to: (i) put appropriate structures and mechanisms in place to restrict misuse, (ii) understand when it has been misused or produced errors, and, (iii) understand who is responsible for undesirable issues arising from its use (Smith and Miller, 2022). It is for these reasons that some scholars argue that transparency and accountability are central to the proper and lawful use of this technology (Brader, 2022; Smuha et al., 2021). Equally, it is important to note that transparency does not always result in effective accountability (Radiya-Dixit, 2022) with both aspects requiring attention. More generally, shifting to algorithmic and AI-based decision-making processes means there is a lack of clarity as to **who is accountable and who a citizen can go to** in order to find an explanation for any decisions made on the basis of such systems (Information Commissioner's Office, 2021).

There are, then, calls for transparency in relation to the deployment and practical use of these technologies (*Biometrics and Surveillance Camera Commissioner*, 2022.; Hintz et al., 2022; Radiya-Dixit, 2022). Concerns have been raised about the need to produce robust and clear guidelines on practice and lines of accountability to ensure that mobilisation is non-discriminatory. The latter were found to be lacking by Veale (2019) in their report with the Law Society. As argued for by the Centre for Data Ethics and Innovation (2020), part of the UK government's Department of Science, Innovation and Technology: 'At a minimum, we would expect police forces to be transparent about how they use live FRT, including where it is deployed, the basis by which people are included on watchlists, and how deployments are approved.' The Ada Lovelace institute (2022, p.4) also found that, 'people want to see a stronger legal framework with independent oversight and minimum standards to prevent harm, create accountability and transparency.'

Yet, meaningful transparency may be difficult to achieve. As Oswald et al., (2018) suggest, 'It is becoming increasingly difficult to explain to non-computer scientists and non-

statisticians how a machine learning forecasting model arrives at its outcomes, and the potential for misunderstanding, and even intentional misrepresentation is vast.” Hamilton (2021) suggests some methods for improving transparency:

- Publicly acknowledging which predictive systems are being used.
- Providing opportunities for representatives of the developer to engage in discussion with members of the public.
- Making at least parts of any algorithm open source.

Pre-crime

There are concerns that certain sections of the public and suspects might be treated more negatively based on an assessment by algorithmic technology, which has the potential to discriminate. As Ferguson (2017) highlights, police may behave more coercively when individuals have been identified as ‘high risk’. This formal attribution essentially becomes a ‘status offense...[where] the **person is punished for his or her status and not for any action on their part.**’ (Hamilton 2021). This becomes even more problematic if assessments are made on factors other than actual offending and based instead on hypothetical future offending – akin to the idea of pre-crime as popularised in the film *Minority Report* (Hamilton 2015).

5. Research Methodologies and Approach

Our research set out to identify and explore with police, civic and public stakeholders the benefits, challenges and ethical concerns raised through the police use of new computational technologies in policing, notably those involving claims to detect and predict criminality through risk-modelling and those based on Intelligent Facial Recognition.

To explore these concerns, a mixed-methods approach was adopted. Alongside engagement with current documentation and guidance supplied by North Wales Police we collected data from four points:

1. Focus group discussions, utilising a narrative approach to familiarise participants with examples of predictive policing technologies and intelligent facial recognition.
2. A national demographically representative survey ($n=2155$) including an additional 155 respondents from North Wales.
3. Semi-structured interviews with expert stakeholders from civic organisations, governmental departments and regulators.
4. Semi-structured interviews with NWP police staff.

Focus Group Discussions

We recruited five focus groups ($n=23$) of participants from the North Wales area using a professional recruitment company. Some of these groups were purposefully sampled to better reflect demographic attitudes towards the use of new technology – for instance, those older groups (65+) younger citizens (18-34). We also aimed to ensure that the perspectives of those people who could be potentially over-impacted by the implementation of this technology – such as those from a minority ethnic background – were included too. The groups comprised the following:

Group	Demographic Sampling	# of participants
A	UK Ethnic Minority	5
B	Older (65+)	5
C	Mixed	5
D	Mixed	4
E	Younger (18-34)	4

Focus groups were conducted digitally in Spring 2023 using video meeting software to facilitate access and reach to participants across North Wales. Participants were provided with a short introduction to AI and relevant technologies (Intelligent Facial Recognition and Predictive Policing) with brief examples, before being presented with short narratives, building on the design fiction principle (Bleecker, 2009; Jensen & Vistisen, 2017) of diegetic (storytelling) prototypes – i.e., designed objects or technologies that exist within a fictional world. The latter has been shown to better facilitate the understanding and familiarisation of new technologies, and enable participants to consider the personal and social impacts of these. The selection of scenarios was based, in part, on initial discussions held with NWP

about the potential and relevant uses of the technology, and further informed by engagement with scholarship, as well as developments and trends in technology to create prototypes that were similar to currently available or emerging technology.

Groups were then asked to respond to a binary question that sought to gauge initial responses to the technology, and which could subsequently prompt discussion. Drawing on Contravision (Mancini et al., 2010), the narrative provided positive and negative examples and outcomes of using the technology, which has been shown to encourage a wider range of responses among participants and deeper reflection on the potential impacts of technology. The below table provides a brief and indicative overview of each scenario presented to participants:

Scenario	Technology: General	Technology: Specific	Short description of narrative
1	Intelligent Facial Recognition	Retrospective Facial Recognition	A customer at a garage overhears a discussion between police officers and staff about the use of RFR to identify suspects using CCTV of a break-in at the garage. Police discuss its success in previous cases, while the garage staff member questions its accuracy based on the mistake questioning of an acquaintance.
2	Intelligent Facial Recognition	Live Facial Recognition	An individual walks through the area outside a sporting stadium during the use of live facial recognition. They witness the successful apprehending of someone on the police watch list. They also witness a parent, with their child, mistakenly stopped after being identified by the facial recognition software.

3	Intelligent Facial Recognition	Officer-operated facial recognition	A during a traffic stop, an officer asks to take a photo of the driver using a mobile camera in order to run the photo through facial recognition software and ensure they are not a person of interest to police. This facilitates a quick resolution to the stop. The passenger in the car subsequently finds a social media post about a prior stop which led to the mistaken arrest of a driver.
4	Predictive Policing	Location-based	A person walks past a night club and expresses concern about previous crime in the area. On this occasion they notice an increase police presence intervening in potential disturbances. The police explain they are using predictive policing to help identify where and when to most effectively deploy officers. As the person is leaving the area they also witness the police moving on a couple of young men from a minority group based on pre-conceived associations with criminality

5	Predictive Policing	Individual-based; risk scoring	An interested citizen visits their local police to ask about the use of predictive policing. The officer talks them through the use of a new system they are implementing designed to assign risk scores for potential re-offending. The citizen is shown a promotional leaflet produced by the developer of the technology with a range of potential benefits. However, the citizen then raises a number of concerns with the technology.
6	Predictive Policing	Text-mining	After reporting the theft of a car, an individual is asked if they would allow the police to include the report detail into a new text analyse system designed to find patterns across police data. This leads to potential recovery of the individual's car through the automated linking in intelligence. However, the individual then overhears a complaint by a garage owner who has been repeatedly identified by the system despite the police agreeing there has been no

			intentional criminality.
--	--	--	--------------------------

These ‘design fictions’ were successful and prompted much thought and reflection by the participants. The data was coded using Nvivo, a qualitative analysis software, combining a deductive approach (drawing on concepts and ideas from the literature) and inductive approach (with codes emerging from the data itself). These were used to generate the themes that are discussed in Section 7.1.

Survey

The national survey was developed by building upon the initial findings emerging from the focus groups to identify areas and themes for enquiry. These were as follows:

Question #	Topic/technology	Sub-question theme	Style of response
1	Attitudes towards police	Trust in police (general)	Likert-scale
		Trust in police (local)	Likert-scale
		Police are positive (national)	Likert-scale
		Police are positive (local)	Likert-scale
		Police are fair and impartial	Likert-scale
		Police make decisions based on law	Likert-scale
2	Retrospective Facial Recognition	Privacy	Likert-scale
		Geographical context	Likert-scale
		Human-in-the-loop	Likert-scale
		Scope Creep	Likert-scale
		Accuracy	Likert-scale
		Necessity	Likert-scale
3	Live Facial Recognition	Privacy	Likert-scale
		Geographical context	Likert-scale
		Human-in-the-loop	Likert-scale
		Chilling effects	Likert-scale
		Consent	Likert-scale
		Safety	Likert-scale
4	Operator instigated facial-recognition	Privacy	Likert-scale
		Geographical context	Likert-scale
		Human-in-the-loop	Likert-scale
		Scope-creep	Likert-scale
		Consent	Likert-scale
		Nothing-to-hide	Likert-scale

5	Predictive Policing: Risk Scoring	Discrimination	Likert-scale
		Geographic context	Likert-scale
		Pre-crime	Likert-scale
		Accuracy	Likert-scale
		Transparency	Likert-scale
		Procurement	Likert-scale
6	Text mining	Discrimination	Likert-scale
		Geographic context	Likert-scale
		Pre-crime	Likert-scale
		Human-in-the-loop	Likert-scale
		Scope-creep	Likert-scale
		Chilling effects	Likert-scale
7	Proportionality	19 short examples of potential police uses of AI-technology	Tick box if use is acceptable.
8	Police cultures	AI-use to reduce prejudice	Likert-scale
		Trust towards police use of AI	Likert-scale
		Police use of AI positive (National)	Likert-scale
		Police use of AI positive (North Wales)	Likert-scale
		AI in policing good (local area)	Likert-scale
		Focus of AI (White collar crime)	Likert-scale
9	AI-technology	Benefits	Open text comments
10	AI-technology	Concerns	Open text comments

The Likert-scale is a five-point indicator with the following potential responses: 'Strongly agree', 'Tend to agree', 'Neither agree nor disagree', 'Tend to disagree', and 'Strongly disagree'. The full survey is included in the appendix.

The survey was delivered by a professional survey company (in accordance with ISO 20252 and ISO 27001 (international standards for market research quality and information data security) as part of a bi-weekly omnibus to 2005 demographically representative participants, with an additional 150 participants from Wales.

Expert stakeholders

As a counterpoint to the citizen perspective, we also interviewed 4 elite, expert stakeholders to provide further insight into issues that might impact on the implementation and use of AI-technologies in policing. Interviews were conducted digitally, with each lasting around 60 minutes. These were as follows:

1. The Biometrics and Surveillance Camera Commissioner (BSCC)

2. A policy advisor for the Centre for Data Ethics and Innovation, part of the Department for Science, Innovation and Technology (CDEI)
3. A representative of Open Rights Group, a human rights advocacy group, focussing on digital rights. (ORG)
4. A representative of a global non-profit, with a focus on digital civil rights. (NP)

Semi-structured interviews with NWP Staff

We also undertook in-depth interviews with a sample of strategically-positioned NWP police officers and staff ($n=5$) to ascertain their understandings of deploying AI technologies in policing, and to capture their normative views of the potential benefits and ethical dimensions of using such approaches in the local communities of North Wales.

6. Findings

The above complementary methodologies provided a uniquely rich, primary dataset and resulted in a number of novel and important findings. In the next section, we present the key themes from the research.

6.1 Citizen Focus groups

Tensions: Safety vs. Surveillance

'I wouldn't know what the right solution is. I think it's good when it comes to, you know, if it means that we get more crime off the street, then I'm all for it. But then, at the same time... the more technology you use, the less freedom that you have.' (Barry FGA)

The discussion in the focus groups was often characterised by tensions in opinion. As might be expected, there were competing opinions between groups, with participants expressing either a strong pro-perspective on the use of AI/predictive policing or a distinctly anti-perspectives on the use of the technology. However, there was also tension at the individual level, with participants holding changing, even competing, views on the technology. For example, many participants were uncomfortable about the potential for children to be included in live facial recognition surveillance, but they still believed that it should be used without being able to justify this surveillance of minors.

This reveals the underlying complexity of the topic and the need to consider the contextualised use of the technology – that is, in specific contexts of usage. It is also the first indication of a key finding of the project: that citizens perhaps lack the fundamental critical awareness and important understanding of the technology and this can impact their ability to consent and make clear judgements.

This is perhaps most clearly expressed in the tension between safety and surveillance. While there were concerns expressed around surveillance – such as scale and the legacy of storing data – and what this might allow police to do, in general, participants were less concerned

about the impacts and potential erosion of privacy. Rather, for many participants, personal and societal safety is an overriding factor, even if AI-use impacts on aspects of human freedoms. Again, context and prior experience of crime and policing becomes important, with participants suggesting that their previous experiences of crime (directly or indirectly) meant they are more inclined to accept the use of the technology. It is important to be aware of personal impacts while not using this as a justification to overrule broader concerns about the use of the technology:

"I was just thinking how if something did happen to say me or my loved ones in a crowd, and my first instinct would also be to see if there is, I mean, reach out to police and understand if there's any like, do we have a recording of this? Do we have a record of people who came in? And I feel like, although it is again infringing on human liberty and freedom, I think that there is still a sort of security that I would also fall back to if something happened. And I suppose in a society, we are asked to trust the police, and I probably would." (Alisha FGX)

At the same time, this must be caveated with the fact that privacy was the most common response when *survey participants* were asked what their greatest concern was about the use of the technology. This disparity between focus group responses and national survey responses demonstrates the nuances in individual viewpoints. Equally, the datasets suggest that, while people are concerned about infringements on freedoms and privacy, if forced to choose, they would champion security measures.

Personal Awareness and Proximity

To a degree, people are more comfortable with the technology if it does not impact them directly, or the impact is minimised by the police. In other words, a little inconvenience and intrusion was deemed to be tolerable if it supported broader concerns of safety and security. By extension, many participants in the study revealed that they would prefer the technology to be discrete and unobtrusive - for example, body worn cameras automatically performing facial recognition, rather than an officer very overtly and actively taking a photograph of a person. This is a theme captured in prior research on emergent technology, with people being more averse to visible and obvious security technology (Laffer, 2022). Interestingly, this meant that many participants seemed more comfortable with the mass surveillance potential of Scenario 1 (retrospective facial recognition) and Scenario 2 (Live Facial recognition) and unnerved by the up-close, targeted, and personal interaction of the officer-instigated facial recognition of Scenario 3.

Preventative vs. Predictive

"It's you're innocent until you're proven guilty. And it's straight in with a, "I'm looking for stuff to try and find to, just to, not so much to verify what you've said, but to try and connect you to something that might have occurred down the line, or something in the past." (Ben FGC)

The most disliked aspect of the technologies was the risk-scoring, with participants concerned in particular about the predictive elements. Hesitancy coalesced around the potential for inaccuracy but, also, in the ability of the technology to frame certain people as

suspicious – most notably, when they were (yet) to commit a crime or reoffend. In its most egregious form, this was viewed as a means of legitimising the targeting of specific groups and forcing them into the stigmatising label of – and path towards – criminality. As one participant stated:

"It should be innocent until proven guilty. But the way it's becoming, it's like guilty until proven innocent." (Ryan FGC)

While none of the technologies were completely uncriticised, those used to support and enhance current policing practices (such as Facial Recognition and location-based predictive policing) were viewed more positively. There was a belief that the police would be performing these actions as part of their routine and ordinary work – such as, attempting to locate suspects, select certain locations in which to conduct patrol – and that the AI and predictive policing systems were simply *helping* officers with these duties and not radically changing existing practices.

Trust and Police Cultures

Across all datasets, there was a reasonable level of trust in the technology itself, although this is heavily tied into accuracy. Participants discussed the need for a very high guarantee/promise of accuracy – some participants even wanted 100% accuracy – but it was equally acknowledged that this would not be achievable. There was concern that those who are involved both in designing and using technologies may contaminate the systems with their own biases and prejudices that inevitably impact on accuracy and effectiveness. This led to participants discussing the – real or imagined – culture of policing. One female respondent stated the following:

'Statistically... it is more likely the police will be using that for stuff that maybe they shouldn't be... That's why they're being put under special measures. That's why in Wales, we have a particularly bad instance with police, especially against women and domestic violence victims.' (Jessica FGE)

There was a perception of systemic issues and biases within the police force that may be replicated through the creation and deployment of the technologies, and there was a broad concern that the police may not already be acting with the best interests towards certain sections of the public. However, this sentiment should be balanced against other participants who believed that AI and associated technologies are neutral and the positive assistance it can provide police officers working in challenging contexts to keep the public safe.

There is evidence to suggest that the recent stream of national scandals exposing the criminal offences committed by serving police officers from across numerous UK forces – and, with it, the spectre of a dysfunctional culture within policing – has influenced

perceptions of the fair and equitable use of technologies by the police. This was particularly the case for North Wales Police, with concerns raised about the use of the technology by individual officers, as well as structural and hierarchical issues:

I wouldn't trust them at all. ... you've only got to see what's been on the news over the last few weeks with the police. And North Wales Police, I don't think their record is particularly good. This data, I mean, like, ... one of the other guys said, you know, once it's out there you don't know who's got access to it. (Oliver FGB)

In the context of North Wales, such sentiments would suggest that concerns around existing police culture ought to be discussed and, ideally, resolved before AI and associated technologies are introduced that might reproduce these issues at scale, or that can be subject to misuse by police.

There was, however, an example of a participant with protected characteristics who was very supportive of, and defended, the police use of technology. This was largely because of their own positive interactions with the police and specific policing groups that were set up to support their community. This indicates that respectful and fruitful interactions citizens have with police officers on a day-to-day level are an indicator of acceptance of new technologies (see also Bradford et al 2020).

Nothing to Hide

'I've got nothing to hide... in theory, I don't have a problem with the police having information on me. What I do have a problem is, is who runs the police, and they might not feel the same way now as they do in 5 years, when they've still got access to my data. But, the circumstances which they can treat and use it has changed.'

There was a strong participant response to the (now clichéd) notion that the technology is acceptable to use if a person has “nothing to hide” – in other words, that individuals have no reason to fear or oppose surveillance programs unless they are themselves engaged in illicit activities. This broad maxim – nothing to hide, nothing to fear – was widely cited across all datasets. However, when the research began to invite participants to reflect on the *personal* impacts of AI and associated technological policing, as well as introducing some of the potential negative affordances of the technologies, then respondents did shift in their perspective. This reveals a gap, or dissonance, in citizens’ critical understanding of the technologies, which needs to be addressed, rather than uncritically accepting the nothing-to-hide-nothing-to-fear argument. Indeed, the broad adoption of this phrase by law enforcement was heavily criticised in the expert and regulator stakeholder interviews.

Where this phrase becomes particularly problematic is when it is combined with resignation that the use of the technology is inevitable. The below comments by participants reflect this:

Yeah, if you got nothing to hide, there's nothing to worry about. And it's everywhere, and you can't avoid it. (Kiera FGC)

No, but it's not face recognition, but it comes back to the point of if you've got nothing to... Even when you do have nothing to hide, and you've done nothing wrong, the system can still find you guilty. (Ben FGC)

As noted earlier, there is a lack of understanding that the adoption and deployment of AI and predictive policing modules are the product of *decisions* and *choices* made by the police. In an article written by the Open Rights Group (ORG 2014), it was argued that the nothing-to-hide-nothing-to-fear narrative that has come to dominate debates about the desirability of surveillance and associated technologies, has several features:

- It sets up an unrealistic 'Us' versus 'Them' dichotomy. It encourages a complete trust in state powers – that you will never face wrongful suspicion or misuse of powers, for only the guilty are affected by mass surveillance.
- It encourages people to embrace their own innocence, to look inwards, and not to look at how other people have been treated or targeted.
- It also introduces the vague threat that if you have not behaved – either now or in the past – that you *do* have something to fear. Not something to challenge, or criticise, but to fear.

Agency and Chilling Effects

In the existing literature, there are general concerns that the use of AI and other technologies for surveillance might have a 'chilling effect' on people's behaviours, particular in relation to the exercise of their fundamental rights of expression and protest. The term is used to describe a situation where a person's speech or behaviour is suppressed because of a fear of penalisation. However, based on the focus group discussions, a broader view of what this 'chilling' entails is required to understand the potential impact of this technology on behaviour.

Some participants did express unease at the use of FR in surveillance, explaining that they would avoid those geographical areas subject to LFR, and also advise others to avoid these areas if they belonged to groups where the accuracy was known to be lower or in question. There was some understanding that this might impact on the general ability for people to go about their ordinary lives – although this was heavily caveated by the desire for enhanced security. For example, one participant spoke about their identity as a Transgender person and wondered how this might influence the effectiveness of the technology. However, they indicated that they would still welcome the use of the technology if it improved safety and security in the area.

Indeed, chilling effects are not always recognised as such. For example, participants suggested that the surveillance of mass populations – i.e., crowds – could be beneficial as it may encourage people to behave better and be less antisocial. What needs to be

considered is the perception that this technology is designed to encourage chilling effects, and not specifically to identify and catch suspects of crime and disorder.

In addition, regard ought to be given to considering the chilling effect of these technologies on the behaviour of *individual police officers*. As Jones (2022) reminds us, police officers increasingly orientate themselves towards future audiences – particularly when their interactions are being recorded – and there were some suggestions that surveillance – whether directed specifically at officers or not – would shape and influence interactions with the public. Further work is needed to examine how AI and allied technologies are internalised by officers and how they may reshape the police use of discretion and commonsense.

It's a bit of a reach. I think, Tim, hit on it really well earlier on. It's when people become reliant on something like this to do their job, they lose the ability to think critically themselves and assess the situation. (Ben FGA)

All of this suggests the need for a nuanced and granular approach to understanding the impacts of this technology on citizen (and police) behaviour, and we have only begun to access this in the focus group discussions.

Accuracy and Human in the Loop:

Accuracy and an over-reliance on technology becomes the point at which people questioned the use of AI and predictive policing technologies, and the ethics of normalising such methods within policing. Earlier, we noted the concern of a Transgender respondent who believed that their identity could pose accuracy problems for the technology. Other comments made by participants also reflect this concern with accuracy:

Accuracy is a big concern as well. It's like... The way that we live in now, it's like Kiera said, we live in a very strange world. I could go out this afternoon and literally have cosmetic surgery done locally that would change my appearance. It would change the width of my lips, it would change the shape of my forehead. Botox and fillers, you could do this, and this is a relatively inexpensive way to [move around]. It's temporary, it doesn't last forever. But you could manipulate how you look quite easily. This is without things like makeup, where you're going to stand out because you've put too much fake tan on, or whatever. This is just stuff that you can blend into a crowd because, let's be honest, we can see... [...]... My sister-in-law, there are fifty girls like her in in my town. They all look the same, they sound the same. They act the same. They've got the same hairstyle, same hair dye, everything. You can't say that there is much difference in the accuracy there, to pick one out of the crowd. (Ryan, FG4)

There are interesting questions around accuracy, and I know some of the big companies are developing ways of, you know, mitigating the fact of people wearing masks and things like that. So that might be something interesting to consider (Tim, FG4)

The technology is shown to be less accurate if, for example, depending on your gender, depending on your ethnicity. So do you think people should be more worried if they're from particular communities? [...] you could be innocent, but if your face is recognized on them, you know, you've got a hell of a fight to win, haven't you? So it's a case of potentially people just sort of accepting what the machine says without using their own discretion. (Rachel FG3)

I just wanted to add to what has been said in terms of whether AI could also be used as a political racial tool. I feel, you know, that could pose a threat, because if the police are looking out for a certain type of people or based on ... If there is any discrimination that's being propelled by this, it might be a threat to humanity in general. (Alisha, FGA)

Discussion also highlighted a concern around errors in the system which then shift the burden onto individuals to show they are not who the system says they are, and its broader implications for undermining the presumption of innocence. This may disproportionately impact already vulnerable communities if the system misidentifies them at higher rates.

A related point concerns the need to be careful not to succumb to discourses and perceptions around techno-solutionism – that is, desire to think about and deploy AI and related technologies as *the* solution. Despite the misgivings around police culture noted earlier, the majority of participants equally felt that the operationalisation of AI and predictive policing systems should always have a human-in-the-loop. This refers to the need for people – rather than machines – being involved in setting up the systems, tuning and testing so the decision-making improves, as well as checking the decisions made. In the context of our research, the idea of the human-in-the-loop also referred specifically to the need to have a certain degree of human supervision in fields where errors can happen. As the next set of comments indicate:

There's always got to be a safety net of a human at the end of it that, ultimately making the decision because the algorithm is going to crunch the numbers, and going to say, "We give you a score of 9 out of 10 for potential reoffending". But the person, there's somebody at the end of that screen, going to say, "9 out of 10 for reoffending? This is a little old lady that's forgot her pension, and created a scene in the post office and thumped some postmaster with an umbrella. This isn't somebody that's masterminding a drugs operation for North Wales." And they [the human] can see, "No, it doesn't look right, that doesn't make sense. We've got to re-look at this and re-enter the data," or whatever. So, there's always going to be some kind of human check involved (Ben, FGC).

I don't think AI is capable of being completely flawless. So you still need that element of real person to do the job (Ryan, FGC)

It was also strongly felt that AI and predictive models should never *replace* the people involved in creating, deploying, and reviewing such systems, as indicated in the following comments:

AI should not be used as an austerity measure. Its potential to erode or detract from human agency and decision making needs to be balanced ...[concerns about] AI providing a digital response to analogue issues. (Ben, FCG)

Although there was broad agreement that any deployment in AI technologies should always include a human-in-the-loop, it was equally felt that any personnel involved in the machine learning system should be experts and well trained on the technologies.

Mission Creep and Data Protection

There were common concerns across the participant groups – albeit at varying degrees of strengths – about the potential for the stated purpose for deploying AI and related technological practices within policing ‘creeping’ into non-law enforcement imperatives. Concern coalesced around the potential for AI-driven systems to be used in more pernicious ways, from inferring political beliefs and automatically detecting and punishing trivial infractions.

There's concerns that, you know, once that technology is out there, they might start using it for additional means [...] we do see like the gangs matrix, for example, you know, it didn't just pick up individuals who had a propensity to violence, it will pick those vulnerable to violence, but kind of encapsulating them all as gang nominals so you couldn't really tell who was the vulnerable person, or like who was of serious concern (Ben, FGC).

Overall, however, there was broad consensus that it was reasonable – even desirable in certain circumstances – for the police to deploy machine-learning technologies if this was *primarily* used for preventing and responding to crime and harmful behaviours. Research participants’ reflection on the ‘creep’ issue entered discussion about the collection of, access to, and storage/retention of the personalised and biometric data that are (potentially) gleaned from police practices. There was very little public understanding of the legal onus placed upon police organisations in relation to using and securing the data. The key data protection issues highlighted by participants included:

- The governance of AI and predictive policing systems, including how police organisations will be scrutinised and held accountability for any (potentially) over-zealous intrusion.
- Some respondents voiced their concern about the governance of so-called watchlists and escalation processes.
- How the speed and scale of the automatic collection of personal and biometric data will be monitored, and how clear justification for its use will be communicated. It was broadly felt that there was little transparency with respect to data subjects’ rights. The processing of children’s and vulnerable adults’ data was highlighted as a particular issue.
- Participants were unclear of how terms such as ‘necessity’ and ‘proportionality’ were considered when the police make key decisions to deploy the technology.

- Another theme was the general lack of choice and control for individuals who may find themselves caught up in the data collection process.
- Some participants also alluded to the potential for wider, unanticipated impacts for individuals and their communities stemming from the police use of machine-learning technologies.

6.2 National Survey with Citizens

The data gleaned from the National Survey demonstrated a broad and varied range of public opinion towards the police use of AI, Facial Recognition and predictive modelling. The greatest areas of difference were *age-related* and *geographical location*, with *ethnicity* also playing an important – if not always straightforward – role in the results. Key highlights are included in the following table and explored further in the discussion below.

Table 1. Attitudes towards the police

Survey Questions N=2155	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Trust in Police, England and Wales	52%	24%	23%
Trust in police in local Area	57%	22%	21%
Police as a positive force, England and Wales	53%	19%	28%
Police as a positive force, local area	53%	19%	27%
Police make fair and impartial decisions	50%	24%	26%
Police make decisions based on the law and not personal opinion	52%	24%	24%

In general, just over half of the respondents agreed with the questions relating to trust in the police. However, as in most surveys on this topic levels of trust were lowest amongst young survey respondents (18-24 year olds) and increased with age bar a dip amongst 45-54 year olds. There were concerns, though, around general trust in the police culture and behaviour, and perceptions are potentially influenced by recent high-profile scandals involving officer misconduct. This could potentially have a negative impact upon how the public view the police use of AI, as alluded to in the open answer boxes – with occasional responses suggesting the technology would be best used internally within the police.

Participants were asked a range of questions around their trust in the police force and their belief that it is a positive force. There were higher levels of trust in local police forces (57%). However only 52% of people in Wales agreed or strongly agreed with the statement (London was the lowest at 49%). Age was an indicator, with positivity rising as participants got older.

Table 2. Attitudes towards the use of Facial recognition technology

Survey Questions N=2155	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Retrospective Facial Recognition Technology			
Use of retrospective FR technology is an invasion of privacy	39%	31%	30%
Comfortable with the use of retrospective FR in rural context	58%	14%	28%
Police officer should make the decision on identification	70%	7%	23%
Concern about the use of retrospective FR beyond identification of suspects for active cases	51%	16%	34%
Concern that FR could lead to discrimination	51%	15%	34%
Use of AI and FR is necessary for the police to be effective	46%	15%	39%
Live Facial Recognition Technology			
Use of live FR technology would be an invasion of my privacy	50%	25%	25%
Live FR is most useful in densely populated areas.	66%	8%	27%
Live FR needs to be checked and verified by a person before officers are allowed to act	66%	6%	28%
If I saw live FR in use, I would be more self-conscious of my behaviour even if not doing anything illegal	50%	21%	29%
I would be comfortable entering an area in which live FR is used	47%	21%	32%
Use of live FR at sporting events - everyone's image is recorded and processed - this	63%	10%	27%

surveillance is acceptable if it improves safety			
Operator Instigated Facial Recognition			
Traffic stop: I would not want a police officer to make photo and use FR - this would be an invasion of privacy	45%	27%	27%
Ability to check identity is most useful in rural areas	52%	12%	36%
Concern that police officers rely on FR technology rather than their experience, training,... when identifying people	56%	14%	30%
Concern that a photo taken and processed by FR would be used for other policing activities	52%	19%	29%
I would feel able to tell a police officer that I did not consent to them taking my photo and using FR to check my identity	41%	24%	35%
If I had not done anything wrong I would have no concern allowing an officer to take a photo to in FR	48%	21%	30%

Retrospective Facial recognition

- Generally, there was a significant level of concern about privacy - with 39% of respondents agreeing that privacy could be a problem in the context of AI in policing. Table 3 and 4 below provide an indication of differences in responses to specific questions by age, ethnicity, and gender. The number of respondents for some ethnic minority groups were too small for meaningful inclusion.
- Concern about privacy was skewed slightly by age. Younger age groups had more concern, whereas only 28% of those 75+ saw the use of RFR as an invasion of privacy. Ethnicity was the biggest indicator of difference here. While only 38% of white respondents expressed concern about privacy, it was the case that 51% of non-white respondents agreed that the use of RFR and its potential invasion of privacy would concern them. In terms of the geographical context and use in rural areas, older age was a positive marker that the technology should be used.

- Echoing the group discussions, most respondents agreed that there should be a human-in-the-loop of any AI and machine learning system – and, particularly when making a final decision on any identifications. Only 7% of participants disagreed with the statement. Interestingly, agreement increased with age, rising from 51% for 18-24 year-olds to 81% of 75+. It is unclear why this is the case and might indicate a greater level of trust in technology or a lower level of trust in police officers for younger respondents. Ethnicity was an important indicator, with 72% of white respondents in agreement and non-white participants at 61%. These variations in responses warrant further investigation.
- Interestingly, more respondents agreed that scope creep was more of a concern (51%) than the invasion of privacy (39%). Ethnicity was a key marker of difference with much greater levels of agreement (65%) – as opposed to 50% of white participants. This was particularly high for Black or Black British participants (70%) but relatively low for Asian or Asian British respondents (55%).
- Concerns around accuracy were similar to scope creep (also 51%). The youngest respondents were least concerned about accuracy and discrimination (44%), potentially revealing greater trust in the technology. As might be expected, ethnicity interplayed with trust in the accuracy of the technology, with 57% of people from non-white and specifically, 64% of Black or Black British backgrounds agreeing that they were concerned about accuracy and potential discrimination versus 50% of white respondents.
- Of potential concern for the police is that only 46% of respondents agreed that RFR is necessary, despite the fact that its current use was explicitly mentioned in the framing description. There were interesting trends in relation to age: agreement shifts 18-24 (32%) rises to peak at 35-44 (50%), then drops at 45-54 (43%) before gradually rising to 75+ (52%). While disagreement is a steady trend down from 22% to 9%. Social class was a strong marker here, with the most professional occupations (class AB) where 52% of respondents agreed that RFR was necessary whereas other groups averaged around 40%.

Table 3: Survey responses to selected highlights for RFR by ethnicity, age, and class

Respondents, n=2,155	Use of retrospective FR technology is an invasion of privacy			Police officer should make the decision on identification on retrospective FR			Facial Recognition Systems have been shown to have varying levels of accuracy when identifying people of different ethnicities and genders. I would be concerned that the use of Facial Recognition could lead to discrimination.		
	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Overall	39%	31%	30%	70%	7%	23%	51%	15%	34%
Gender: Female	39%	29%	32%	72%	4%	24%	52%	12%	36%
Gender: Male	40%	32%	28%	68%	9%	22%	49%	19%	32%
Ethnicity: White	38%	32%	30%	72%	6%	22%	50%	16%	34%
Ethnicity: Black	53%	27%	20%	67%	10%	23%	64%	9%	27%
Ethnicity: Asian	52%	21%	27%	60%	12%	28%	52%	12%	36%
Age: 18-24	43%	24%	32%	51%	15%	34%	44%	15%	41%
Age: 25-34	44%	29%	27%	66%	9%	26%	57%	12%	31%
Age: 35-44	42%	27%	30%	63%	7%	29%	54%	11%	35%
Age: 45-54	43%	29%	27%	72%	5%	23%	49%	15%	36%
Age: 55+	36%	32%	31%	76%	6%	18%	48%	21%	32%

55-64									
Age: 65-74	35%	35%	30%	83%	3%	14%	49%	19%	32%
Age: 75+	28%	40%	32%	81%	4%	15%	51%	16%	34%
Class: AB - Higher + intermediate managerial, administrative + professional occupations	36%	37%	26%	72%	9%	19%	57%	15%	28%
Class: C1 - Supervisory, clerical, + junior managerial, administrative + professional occupations	40%	31%	30%	68%	6%	25%	51%	14%	34%
Class: C2 - Skilled manual occupations	39%	29%	31%	73%	6%	21%	47%	17%	36%
Class: DE - Semi- skilled and unskilled manual occupations; unemployed +	43%	25%	32%	68%	6%	25%	46%	16%	38%

lowest grade occupations									
Wales	47%	26%	27%	67%	10%	24%	55%	16%	29%

In sum: a significant number of people do not see RFR as a *necessity* for policing, with a relatively notable proportion of respondents actively disagreeing that it is. Issues around scope creep and accuracy are of greater concern – more so than unease around privacy – with ethnicity a general marker of increased concern. However, there was much variation within ethnic groups – e.g., people from a Chinese background seemed much more positive than others. Throughout all the questions, respondents from Wales expressed among the highest levels of concern (alongside Scotland and London).

It is also worth noting that fear of crime likely influenced responses to the use of AI and machine-learning in policing. Feelings of insecurity and public safety were not just apparent in the elderly respondents. Those aged 35-44 also revealed concern about crime in their communities.

Table 4: Survey responses to selected highlights for LFR by ethnicity, age, and class

Respondents, n=2,155	Use of Live FR technology is an invasion of privacy			Police officer should make the decision on identification on Live FR			Chilling effect: Police using live FR in a public area would impact on behaviour		
	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Overall	50%	25%	25%	66%	6%	28%	50%	21%	29%
Gender: Female	49%	23%	28%	68%	3%	29%	54%	19%	27%
Gender: Male	50%	27%	23%	63%	9%	28%	46%	24%	30%
Ethnicity: White	48%	26%	26%	66%	6%	29%	48%	23%	29%
Ethnicity: Black	66%	22%	11%	74%	9%	17%	70%	9%	21%
Ethnicity: Asian	61%	15%	24%	65%	9%	26%	57%	10%	33%
Age: 18-24	53%	18%	28%	48%	13%	29%	56%	13%	31%
Age: 25-34	56%	19%	24%	68%	5%	27%	60%	12%	28%
Age: 35-44	49%	25%	27%	56%	7%	31%	52%	17%	31%
Age: 45-54	51%	24%	25%	63%	6%	31%	49%	22%	29%
Age: 55+	46%	26%	28%	72%	4%	24%	47%	28%	25%

55-64									
Age: 65-74	46%	33%	21%	72%	5%	24%	44%	29%	27%
Age: 75+	44%	34%	22%	84%	1%	15%	38%	32%	30%
Class: AB - Higher + intermediate managerial, administrative + professional occupations	45%	31%	24%	72%	5%	22%	48%	24%	28%
Class: C1 - Supervisory, clerical, + junior managerial, administrative + professional occupations	50%	25%	25%	62%	8%	30%	50%	21%	29%
Class: C2 - Skilled manual occupations	51%	22%	27%	64%	5%	31%	52%	20%	28%
Class: DE - Semi- skilled and unskilled manual occupations; unemployed +	53%	22%	26%	65%	5%	30%	51%	20%	29%

lowest grade occupations									
Wales	52%	20%	27%	65%	7%	28%	54%	23%	23%

Q3 Live Facial Recognition

- LFR elicited much higher levels of concern around privacy than RFR, but only 50% of respondents nevertheless agreed it is a problem for them. Ethnicity was a key indicator (white respondents scored 48% versus non-white groups at 62%). Within the ethnic groups, and echoing other patterns, Asian or Asian British (61%) and Black or Black British (66%) respondents' concerns were higher.
- There was a relatively high belief that LFR is most useful in densely populated areas. This is strongly marked by age which tended to trend upwards in terms of agreement: 18-24 (53%), 25+ (65%), 35+ (58%), 75+ (78%).
- As with RFR, there was also high agreement that any use of LFR should involve a human-in-the-loop. Interestingly – and in significant contrast to RFR – Black or Black British had the highest agreement for a human-in-the-loop (74%).
- With respect to concerns around the chilling effects of LFR, there was a significant disparity between respondents based on ethnicity. For example, white groups agreed that this was a concern at a lower rate (48%) compared with Asian British (57%), Black and Black British (70%) respondents. These disparities in Ethnicity should be an area for consideration in any future use of LFR and RFR and reflect some unease in specific groups of the population which should be explored further in research.
- Welsh people are relatively more concerned about LFR use, than RFR. The data shows that Wales continues to occur as one of the UK regions with high levels of concern. mission creep is relatively higher in terms of concern than privacy. Within this context, ethnicity is major differentiator, but it is important to attend to the variations between and within groups.
- The impact of gender on concerns around public safety and consent for the technology, as well as the need for a human in the loop and other issues showed a mixed picture and is another variable worthy of further exploration.
- Social class is a less significant indicator for acceptance of LFR, but it does appear to be more acceptable at the higher socio-economic levels.
- In questions that directly address discrimination, we are seeing higher disparities between Wales and closer regions (e.g., North West). However, trust in data sharing with private companies was higher, although this is also still relatively low.

Table 5: Attitudes towards Predictive Policing

Survey Questions N=2155	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Concern about predictive policing reinforcing discrimination	52%	14%	34%
Locally, police deployment decisions are better supported by local police knowledge rather than predictive policing	55%	8%	36%
Concern individuals will be identified as offenders/re-offenders before they have engaged in new criminal activity	50%	13%	38%
Confident that police data is good enough to be used for predictive policing	36%	16%	48%
Policer officers should be able to fully explain to those under arrest how AI has produced informed that informed police decisions	68%	5%	27%
Comfortable with police sharing data with private companies for data collection and analysis	27%	40%	33%

- There was broad agreement that text analysis and associated machine-learning would be a good use of police resources, particularly when it came to finding hitherto unnoticed offending patterns – such as domestic violence.
- General levels of acceptability were low, with 19% of respondents finding little acceptability with the majority of scenarios. Two of the scenarios were, however, rated as over 50% acceptable. The first scenario covered the deployment of Live Facial Recognition outside a football stadium, prior to the beginning of a match, to identify people on a watchlist who were banned from attending (64%). The second scenario incorporated a non-punitive element in which CCTV was used to assist in the swift location of a missing persons using (52%).

- However, as soon as the prospect of using LFR in public environments was introduced - as opposed to targeted events - acceptability fell to 33%. If incorporating children into any recordings lowers acceptability further (28%). However, this figure does rise for non-punitive scenarios, such as young missing persons (38%).
- Those who took part in the survey were also less keen for RFR to be used on home footage (41%) than commercial CCTV (44%).
- Predictive policing is also rated rather low in terms of its acceptability, with individual risk-scoring viewed as largely problematic. However, these sentiments shifted depending on geographical location and if it had the potential to directly help victims of crime.
- Police-worn body cameras were deemed much more preferable to officers using overt mobile camera when taking photos for FR. This finding mirrors the focus groups.

Table 6: Attitudes towards Text Mining

Survey Questions N=2155	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Concern that use of AI on written police reports would reflect and reproduce existing biases that negatively impact certain groups in society	47%	17%	36%
Locally, confident that text analysis would a good use of police resource to support officers investigating crimes	43%	13%	44%
Concern that people identified as connected to a case by AI would be treated as suspicious regardless of prior behaviour or relationship to crime	53%	12%	35%
Concern that police officers rely on AI suggestions without understanding how the AI arrived at them,	58%	11%	31%

leading to potential misjudgements and wrong people being questioned			
Comfortable for the police to use social media for text analysis – including those of suspects, associates and broader networks	44%	20%	37%
Concerned about talking to the police if I knew that everything I said would be recorded, stored in intelligence logs and analysed in future	43%	22%	35%

Table 7: Proportionality of AI Use

Use of AI in scenario	Use seen as acceptable in percentage
Live Facial Recognition is used outside a football stadium before a match to identify people on a watchlist banned from attending.	64%
Live Facial Recognition is used during a protest to help police identify protestors suspected of previous criminal activity that the police think is relevant.	50%
A camera used for Live Facial Recognition is placed on a busy street and captures footage of adults as they pass by. This footage is processed and faceprints created to compare to a watchlist of people suspected of criminal activity.	33%
A camera used for Live Facial Recognition is placed on a busy street and captures footage of adults and children as they pass by. This footage is processed and faceprints created to compare to a watchlist of people suspected of criminal activity.	28%
A railway station CCTV system is updated to allow for automatic delivery of footage to a missing persons database. A facial recognition system then assists officers to quickly scan the video footage to search for missing persons.	38%
A railway station CCTV system is updated to allow for automatic delivery of footage to a police serious crimes database. A facial recognition system then assists officers to quickly scan the video footage to search for suspects of serious crime.	52%
A railway station CCTV system is updated to allow for automatic delivery of footage to a police database of local residents subject to anti-social behaviour orders banning them from the station. A facial recognition system then assists officers to quickly scan the video footage and search for anyone not allowed in the area.	46%
Retrospective facial recognition, where facial recognition is performed on previously saved images, is used on CCTV footage of a burglary recorded in a petrol station.	43%

Retrospective facial recognition, where facial recognition is performed on previously saved images, is used on CCTV footage of a burglary recorded in a petrol station.	44%
Retrospective facial recognition, where facial recognition is performed on previously saved images, is used on home footage of a burglary recorded in a private home.	41%
An AI-algorithm uses historical crime statistics to identify probable locations for future crimes. Based on this guidance, police increase their presence in identified crime 'hot spots'.	33%
An AI-algorithm is used to identify future potential victims of crime. They are contacted by community support officers to provide guidance on how to stay safe.	26%
An AI-algorithm is used to predict an individual's likelihood of re-offending. This is used by police to quickly identify individuals to be investigated when subsequent crimes are committed.	22%
In order to support police ability to fight crime, AI text analysis can be performed on internal police records, such as written intelligence logs.	30%
In order to support police ability to fight crime, AI text analysis can be performed on social media data if it belongs to a suspect in a case. This might include a social media profile or messages.	30%
In order to support police ability to fight crime, AI text analysis can be performed on social media data if it belongs to either a suspect in a case or anyone in their network of connections (e.g., their Friends List on Facebook). This might include social media profiles or messages.	26%
A police officer stops someone suspected of criminal behaviour. Their body camera collects footage of the person. Facial recognition technology is used to identify whether they are a known suspect.	45%
A police officer stops someone during a routine traffic stop. Their body camera collects footage of the person. Facial recognition technology is used to identify whether they are a known suspect.	35%
A police officer stops someone during a routine traffic stop. They take a photograph of the person using a mobile camera. Facial recognition technology is used to identify whether they are a known suspect.	30%
None of the above	19%

Table 8: Police culture and AI

Survey Questions N=2155	Proportion Agree (Strongly and Tend to)	Proportion Disagree (Strongly and Tend to)	Neither Agree or Disagree
Comfortable that the use of AI will improve police outcomes by reducing prejudice	43%	23%	34%
Given criticism I am less trusting of the police use of AI	43%	16%	40%
AI in policing is overall a good idea	47%	18%	34%
For NWP in particular, AI in policing is overall a good idea	32%	18%	50%
For the local area where I live, AI in policing is overall a good idea	44%	20%	36%
Benefits from AI in policing but white collar crime is being overlooked	57%	10%	33%

Other Highlights/Observations:

- Members of the public do nevertheless see the utility of this technology if it will help with resourcing, such as over-worked police officers. There was also general agreement that Facial Recognition technologies, and associated machine-learning, can benefit investigations – e.g., could be used in an emergency to identify a lost vulnerable person, speeding up investigations, and promptly identifying a ‘wanted’ suspect..
- Ethnicity was not a straightforward indicator of perspective on the use of AI technologies, reinforcing the need not to homogenise groups. Generally, attitudes were relatively in line when comparing White respondents and the Net results of Non-White respondents. However, Chinese respondents tended to be very positive about the police and AI. Black and Black British respondents were positive in general about the police, but less so when it came to considering their decision making.
- Gender is relatively balanced across most statements, other than statement 6. Women (57%) were much more comfortable with use of LFR for helping to locate missing people (including children) than men (48%).

- In relation to age, there was an almost universal trend upwards in approval from younger people to older populations. There were a few minor deviations from this linear pattern but, overall, the trend does not change.
- Broadly speaking, social class trends downwards in approval, from the highest socio-economic groups (AB) to those at the lower end of the scale (DE)

6.3 Elite Stakeholder Interviews – Key Highlights

Part of our research incorporated interviews with relevant elite stakeholders and others in the regulatory environment. We are only providing a thematic overview here. As one might expect, there were different perspectives on desirability of the police deployment of AI and allied machine-learning technologies. While there were certainly a number of positive statements made for incorporating FR and predictive models into policing practices, there were also strong concerns raised.

There were tensions and differences of opinion between the organisations we interviewed. Some raised concerns about whether the police, as a state institution, could be trusted to use any AI capability ethically and appropriately, although there was an understanding that the police would have the desire – perhaps because it was imposed on them through external regulatory and legislative systems or political pressure – to find out how to do so.

Some of the interviewees were more rehearsed with FR, rather than predictive policing. However, the act of claiming to predict crime and offenders was, in principle, dystopian. One respondent suggested that predictive policing was simply another guise of (racial) profiling.

We have provided the views of these stakeholder by the organisation they represented with their consent.

Centre for Data and Ethics Innovation

- The ethical implications of its use came to the fore in a way that differed from the group discussions with ordinary members of the public. However, these were couched more in terms of how to make AI work ethically, rather than the question of whether it *should* be done.
- The chief preoccupation was on any potential algorithmic bias that may produce and reproduce systemic inequalities along axes of social class, ethnicity, and age. However, it was suggested that this can be partly mitigated through ensuring the human-in-the-loop.
- It was accepted that, due to the marked difference between the 43 police forces of England and Wales, there would be difficulties in any attempt to standardise approaches towards deploying AI technologies.
- Echoing some of the sentiments expressed in the group discussions, concerns were raised about how to ensure public consent to the use of AI and predictive models. It was explained that transparency and consent go hand-in-hand.

- There was a pressing need for police forces to better communicate what software is being used (or will be used), by who, and for what purpose. The theme of data security and storage is also relevant here.

The Open Rights Group (ORG)

The ORG were most critical:

- LFR was considered to be far less desirable than RFR. It was believed that LFR had a higher potential for inaccuracy, and there was also an allied concern about how it can influence the police to behave more aggressively based on its recommendations.
- There was much to indicate that technological developments in policing only serve to re-marginalize those communities who are already targeted by the police.
- While many respondents – across all the datasets – noted the importance of ensuring a ‘human-in-the-loop’, the ORG found that any police involvement in designing, overseeing, and reviewing AI and machine-based learning methods as a negative element. It was suggested that the toxic aspects of police culture and prejudice that have long been implicated in the over-policing of Black and ethnic minority groups may extend into AI policing practices.
- The newness of – and, with it, the unfamiliarity with – the technology was also a cause for concern. There was a degree of scepticism around the ability of those who will operate the systems to truly understand both the capabilities and implications of such technologies.
- It was suggested that the police should not purchase or deploy machine-learning technologies until these existing issues (and potential for bias) are solved.

Biometrics and Surveillance Camera Commissioner

The following key themes emerged during interviews with the BSCC:

- **Legislation**
 - Concern was raised around new legislation being framed in terms of data protection, which was not deemed sufficient.
 - The potential tension between centralised guidance versus localized sensitivity analysis.
- **Procurement of the Technology**
 - The need to engage citizens in pre-procurement and set clear procurement guidance, with appropriate review and revision processes.
 - Current guidance on procurement is insufficient, leading to unethical – and unsecure – purchasing. Interviewees pointed to the recent scandal involving the UK Health Security Agency use of video surveillance technology provided by the controversial Chinese firm Hangzhou Hikvision Digital Technology

- It was understood that different police forces ought to be afforded a level of autonomy and discretion, although the difficulties of achieving this across multi-force operations was acknowledged.
- **Accountability**
 - A potential tension around knowledge of accountability and governance structures. The Police and Crime Commissioners plays an important role, but only 4 of 39 police forces surveyed by the office of the BSCC mentioned the PCC when asked about being held accountable for their surveillance.
- **Importance of citizen buy-in and engagement**
 - A pressing need to involve the public in the conversation, including the need to educate people on the technologies currently being used, and those that may be deployed in the future.
 - It is important to ask challenging and difficult questions of the police in order to arrive at balanced views.
 - In that regard, it is paramount that the 'nothing-to-hide-nothing-to-fear' argument – that features in police discourses and justification – is avoided.
- **Predictive policing is too risky**
 - There are too many shifting variables, increasing the risk of inaccuracy.
 - Resources and attention in technology could instead be put elsewhere
 - Risks more prejudice/bias and erosion of public trust: 'I think some of the predictive policing stuff that we've seen both here and in the U.S., it is the science of the sorting hat'.
- **Recognises the risks of LFR, but is an advocate**
 - The potential to draw on 'old' data sets is concerning as this can lead to higher rates of inaccuracy and bias.
 - Suggests searching through online sex abuse images/videos is useful for AI-driven technology – largely to spare officers from doing this work.
 - Can be most helpful during serious events (active terror attack), but also for missing people (particularly abducted children). There may, then, be a marked difference between policing and law enforcement. It may make sense to – at least initially – use these technologies in non-law enforcement settings (e.g., safeguarding and missing persons).
- **Proportionality and Necessity**
 - The level of intrusion and collateral impact needs to be considered. It is finding a balance between the desired outcome and the prospect of what is actually collected.
 - One issue is that some police do not see capturing and storing images of the public as intrusive. Points to problems in current practice around storing images. If an officers, or another member of staff, cannot point to an authority under which they are currently retaining it, then the default position must be to delete the image(s). Police have been instructed to bulk delete, but this

has not been done and that is eroding public trust in the introduction of new technology.

- It is common to hear the police state, “it may help” or “it will be useful”. But a focus on the word *necessary* is important because there is a general understanding that this needs to be more than just helpful or more than speculative. It often requires the need to have excluded all the less intrusive options.

- **Mission Creep**

- Any deployment must be bound by what has been agreed.
- One issue around the CoP guidance on the potential use of LFR to identify witnesses.
- Concern around connections with other (public/private) databases.

Non-Governmental Organisations

NGOs were most strongly against the use of the AI technologies within policing, as well as the adoption of predictive models to inform routine police work.

- **Banning Biometric Surveillance?**

- The interviewee suggests that aspects of biometric surveillance are being used for purposes that are already banned/illegal. In their view, remote biometric identification in publicly accessible spaces is particularly problematic, and illegal according to GDPR. They also claimed that such a sentiment was supported the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) opinion as well on the AI act.
- There was a pressing importance for clarity on interference of rights. As they put it: “They're confusing being identified with being subject to a process of identification. And the interference with your rights comes from the moment that you enter the kind of surveilled space, your image is captured, biometric template is created, matching process is done...” I don't think when they're talking about proportionality that they're accurately measuring the interference. They're only focusing on false positives”.
- Chilling effects are concerning, although the way in which surveillance technologies are internalised by citizens and prompt changes in thoughts and behaviours cannot be easily measured.

- **Transparency at the basic level**

- The public sector ought to embark on a broad drive to better inform citizens about which organisations are using AI, as well as a clear explanation of what they are doing with it. This should be concerned to address basic questions, such as:
 - What is the system? Who is responsible? What is it designed to do? What is the objective? What kinds of data was it trained on?

- The UK could look to models currently being rolled out in other countries: e.g., the Algorithmic Register Initiatives in Helsinki and Amsterdam. The AI registries were launched to detail how the government in each city uses algorithms to deliver services.
- If the positive discourse about the successes of AI is to be believed, then the public organisations using it should be happy to provide further detail of how it is used, and its outcomes.
- Disclosure of AI and other surveillance technologies is through brief promotional materials that often underscore the necessity of the methods for responding to the most serious and emotive crimes (e.g., terrorism) or when it goes wrong, as seen in the unlawful ruling against South Wales Police.
- The lack of transparency means there are deep issues around the legitimacy of the systems currently being bought. This lack of scrutiny and understanding within the procurement process means – at least theoretically – that AI systems are being designed “based on 19th Century racism”.
- The dominant ‘arms race’ narrative – that often also includes the idea that the police adoption of AI and machine-based learning is necessary to keep up with criminals – should be rejected, or at least fact checked.
- The interviewee was also concerned that ‘necessity’ was being used as a proxy for austerity measures – a potential issue for rurality.
- Like-for-like comparisons of technologies is not always appropriate:
 - There has been an argument that Retrospective FR is better (‘less bad’) than Live FR. It can be worse in some cases than the live, given its opportunities for future issues. It can weaponize historical footage against people in a way that Live cannot.
 - Worries around a permanent AI infrastructure within (police) organisations that can be turned on (and, arguably, off) without notice or attention.
- Current practices are worrying – a predisposition to simply ban the technology *after* it has been used, rather than fully assessing it in advance.

Other Tensions Between the Organisations:

Privacy and consent

- CDEI - comfortable with issues of privacy and consent and could be led by public attitudes. Having the potential for citizens to *not* provide public authorities with their consent or opt-out is considered to be a possible model.
- However, this is considered problematic by Non-Profit groups and the ORG. Public views may run counter to fundamental human rights, and are very unlikely to take into account marginalised and over-policed communities. There has not been enough public consultation, and so it is likely that many people do not understand the technology.

Procurement

- There seems to be slightly more concern here than in the group discussions. CDEI focussed on responsibility and accountability. BSSC were clear that there needs to be guidelines and transparency as current tech procurement practices are leading to problematic purchases. Procurement is outdated and becoming a global security concern (again pointing to camera systems).

Predictive policing

- BSCC - predictive policing is akin to the 'science of the sorting hat'.
- CDEI - transparency is a major issue and how to interpret/use outputs of the system.
- ORG - predictive policing is another way of profiling.

Governance

- Concerns over governance and the fluid and fragmented regulatory landscape (BSCC) and, even current attempts at regulation are potentially not going far enough.
- UK will inevitably be impacted by the EU Artificial Intelligence Act, when it comes into force.

Proportionality

- There needs to be clearer guidance on what proportionality is within this context. ORG - this is not made clear on current documents, although the BSSC provides some insight.
- For NP - proportionality is never sufficient. For the rare cases where the technology might be helpful, 'those very, very, very, few cases where this bar is met, you're going to destroy anonymity in public spaces? If it's about the Live, then you're going to scan everybody who goes through those spaces, infringe on their rights, undermine people's freedom of assembly. Everything. That seems totally disproportionate [...] You know, it's not necessary. It's disproportionate.

Police data practices

- Criticism from all organisations (ORG, NP, BSCC). These need to be sorted before appropriate implementation can even be considered.
- Before procurement – images/data that should be deleted but police do not see retaining images as an issue (BSCC)

Future projections

- Concern about the way that AI/FR/PP is moving. For example, the CoP discuss using the technology to identify *witnesses*, not just potential suspects. Concerns over mission creep and the desire to use existing databases for the new purpose. BSCC raised the question of where the police are getting the probe images from?

6.4. Interviews with Police Personnel

The empirical research also included semi-structured interviews with strategically positioned staff within NWP to ascertain their understandings of deploying AI technologies in policing, and to capture their normative views of the benefits and ethical dimensions of using such approaches in local communities. The interviews elicited a range of thoughtful and nuanced views among NWP personnel. In what follows, we present several insights from the officers and staff we spoke with, exploring their hopes and concerns about the role of this technology.

- Some interviewees believed that AI technologies and processes had the potential to be an important (and neutral) tool, even a ‘game-changer’, in the fight against crime. If the technical capabilities promised by AI systems came into fruition, a system capable of pointing the police in the right direction could be extremely relevant. Indeed, all interviewees heralded ‘speed’ in identifying major suspects as a major breakthrough and as a real benefit for policing.
- Nonetheless, such confidence was caveated around the premise of the technology being ‘fit for purpose’ and working effectively. An often-cited example was the grainy or distorted CCTV images that can frustrate police investigations.
- Many respondents discussed the ‘inevitability’ of police organisations adopting AI-based technologies. It was often stated that other – once novel – technological methods for policing were subjected to heated debate are now deployed in the North Wales area on a daily basis – including Automatic Number Plate recognition (ANPR), Body-Worn Video (BWV), mobile fingerprint devices and drones. Developments in the use of Facial Recognition were viewed simply as the ‘next step’ in this regard. To this end, interviewees showed confidence in potential future uses of FR for policing.
- The above sentiments were, however, often accompanied with the *arms-race-narrative* noted in the stakeholder interviews. The NWP respondents championed the idea that the police adoption of AI and machine-based learning is necessary to

keep up with increasingly tech-savvy criminals who cross borders (virtually and physically).

- However, several respondents felt that they, personally, did not have enough understanding of the mechanics behind AI and machine-based learning - both generally and in the context of policing. For example, when asked, a number of interviewees were unsure of how to articulate or define what an algorithm is, how exactly FR captures faces, or the theory behind predictive/risk-based policing models. At the same time, all respondents agreed that any deployment of AI in policing ought to be managed, overseen, and reviewed by a knowledgeable *person* - in other words, they were hesitant to trust in a technology that is largely operated by non-human actors. Thus, the human-in-the-loop theme was a key theme of interviews with NWP personnel.
- One respondent who has responsibility for a largely rural policing area within North Wales questioned whether the size of the Local Policing Unit (LPU) and the types of criminal activity they often face would lend itself to the use of AI systems. This participant stated that the community officers who covered the area tend to know the members of the public they often interact with, and this may therefore reduce the value and relevance of AI technologies, particularly FR.
- Similarly, another respondent, whilst revealing optimism with the potential use of FR for large events and crowds, also queried its appropriateness in a rural setting due to the vast majority of interactions with the public occurring at a much smaller and local scale.
- Although respondents generally believed that technologies for policing had become 'smarter', concern was expressed about suspects evading FR by changing their appearance, the way they walked (gait) and other altered characteristics.
- While interviewees commonly articulated faith in FR systems, and other technologies for machine-based learning, to enhance policing, they all stated that a key challenge for NWP would be to prove its worth to the public. Everyone suggested that the public in NW deserved to be briefed early to explain the purposes of the AI systems (e.g., LFR) so it is not deemed invasive and counter to the community policing ideals of NWP.
- In a similar manner to the group discussions, interviews with police personnel also revealed concern around: accuracy, the due privacy infringements, consent, bias against certain populations, and areas where its use would be unethical. Particular trepidation was expressed in relation to data collection and the need to follow due process. There were also particular concerns related to data management and how footage is stored in a scenario of *constant collection*.
- An allied concern was proportionality, in particular, if the purpose of use is not clear. It was felt that any covert or opaque deployment of FR might impact public confidence in police work. Interviewees could also see the potential issue of mission creep noted during the elite interviews, group discussion and survey data.

7. Preliminary Recommendations

It is important to note that the below are not an exhaustive list of recommendations for implementation. Rather, we highlight arising issues from the research and make tentative suggestions that can potentially guide future use of AI in policing.

- Transparency
 - Context specific understandability (CDEI)
- Trustworthiness
 - Ensure transparency is at the heart of any system.
 - Concerns around police culture, behaviour and bias. This should be addressed before the technology is introduced. (ORG)
- Data protection practices need to be exemplary → for existing and future identifiers.
- Human-in-the-loop
 - However, the human in the loop should be the *right* person, with the *right* training and understanding.
 - ORG – use of social media text mining is worrying – “content can be taken out of context and crafting like narratives out of it that do not necessarily add up to an accurate picture of a person. And it's essentially racial profiling.”
- Danger of over-reliance on ‘respectable’ citizen voices – perhaps focus on concerns that are raised and not on any perceived blanket support.
 - The need for critical literacy in the technology, and a better awareness and understanding of how the technology works.
 - Impact on marginalised voices – danger of homogenising different groups:
 - Suggest further demographic investigation and interaction with different communities within North Wales.
 - Need to engage with those already in the system or associated with crime → what are their views?
- Necessity
 - Is AI and predictive systems *essential* to the work of NWP? This should be weighed up in relation to the many issues identified, the concerns voiced by ordinary people and experts in the field, and the desire for traditional community policing methods.
 - However, this should also be balanced by BSCC assertion that there is a potential accusation that police forces in England and Wales are not fulfilling their duty if they do *not* use the available technologies in certain situations.
- There could be a consideration of binding agreements in place to limit mission/scope creep and to not change the parameters of use:
 - Real concerns about the insertion of finding ‘witnesses’ to crimes using LFR in the CoP guidance.
- There is a need to go beyond a simple ‘check-list’ for individual uses of technology. Rather, there needs to be general consideration of its impact across uses – to ensure

there is no unintended targeting/bias and that it is applied to areas of concern as identified by citizenry.

- Clear scenarios of use could be used to communicate the technologies to the public. CDEI have a public database of Algorithmic uses by public sector bodies.
- Until these elements are taken into consideration (and, ideally, addressed), caution should be exercised against the implementation of this technology. It is currently open to legal challenge, based on the confused and changing legal and regulatory backdrop, but there are further resourcing and ethical implications that require further mitigation. This is born out in the focus group and survey findings. Positive support may be given by some members of the public but even this support is contingent on safeguards and comes with important reservations as set out above. Public attitudes towards AI technology use in policing paint a more complicated picture, and this notwithstanding the assumption that those who may be most potentially impacted will be the most negative.
- As the (now removed) Surveillance Camera commissioner argued (*Facing the Camera: Good Practice and Guidance*, n.d.), the highest ethical and legal standards are required in creating watchlists.
- Public acceptance will be difficult to secure unless there are reassurances around police cultures to enhance trust in the police. This is, in addition, against a backdrop of limited levels of understanding.
- Residual concerns persist around current data practices – for example, the storage of custody photos on the police national database, which have been ruled to be in unlawful. → address concerns around not changing or only paying lip service to guidance.
- Should be a much fuller and genuine engagement with citizens, but particular care ought to be taken to ensure that minority voices are heard.
- As suggested by stakeholder participants, perhaps it would be best to allow this technology to be used by other emergency services. Although, this should not be taken as endorsement of technological solutionism. There are a number of issues raised in the literature, and by participants, that more directly feed into questions of what *should* happen rather than what *could* happen.
- Human-in-the-loop is a strong theme across all datasets → what levels of expertise and training are NWP going to be able to draw upon to ensure users are effective (and appropriate)? For example, it is particularly important to address concerns around the potential of the technologies to erode or detract from human agency and decision making, and to uncritically provide a digital response to analogue issues.
- Should NWP embark on a public information campaign, or similar, to highlight the strengths, weaknesses and realities of deploying AI, FR and smarter predictive

policing models? One point for clarification is whether NWP are using AI as an austerity measure.

- Legitimacy is central to citizen approval, but it can be viewed negatively if it emerges only from passive acceptance. It would be problematic to have high-levels of AI deployment in a context where there is a low-level of public understanding and consent.

2. References/Indicative Bibliography

Akhgar, B., et al., (2022). Accountability principles for artificial intelligence (AP4AI) in the internet security domain. AP4AI Project.

Babuta, A. and Oswald, M. (2020). Data analytics and algorithms in policing in England and Wales. RUSI.

Bayerl, P.S., et al. (2021). Artificial intelligence in law enforcement: cross-country comparison of citizens in Greece, Italy and Spain. S. N.-R. Intelligence: Springer Nature, 11769

Bradford, B., et al. (2020). Live facial recognition: trust and legitimacy as predictors of public support for police use of new technology. SocArXiv. doi:10.31235/osf.io/n3pwa.

Carrasco, M., Mills, S., Whybrew, A. and Jura, A. (2019). The Citizen's Perspective on the Use of AI in Government. BCG Digital Government Benchmarking.

Committee of the Regions for Artificial Intelligence in Europe Brussels, (2018). *AI Communication from the Commission to the EU Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the regions on Artificial Intelligence for Europe*. Brussels: Europe. Available from: AI Communication (europa.eu).

Dencik, L., Hintz, A., and Cable, J., (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big data & society*, 3 (2).

Di Vaio, A., Hassan, R., and Alavoine, C. (2022). Data intelligence and analytics: A bibliometric analysis of human-Artificial intelligence in public sector decision-making effectiveness. *Technological Forecasting and Social Change* (174)

European Digital Rights and Edinburgh International Justice Initiative (2021). *The rise and rise of biometric mass surveillance in the EU report*. Brussels: EDRi and EIJI.

Feldstein, S. (2019). The global expansion of AI surveillance. Carnegie Endowment for International Peace.

Fussey, P., Davies, B., and Innes, M. (2021). 'Assisted' facial recognition and the reinvention of suspicion and discretion in digital policing. *The British journal of criminology*, 61 (2), 325-344.

Fussey, P. and Sandhu, A. (2022). Surveillance arbitration in the era of digital policing. *Theoretical criminology*, 26 (1), 3-22.

Getting the future right: artificial intelligence and fundamental rights (2021). *European Union Agency for Fundamental Rights*.

Gurinskaya, A., 2020. Predicting citizens' support for surveillance cameras. Does police legitimacy matter? *International journal of comparative and applied criminal justice*, 44 (1-2), 63-83.

Heaven, W. (2020). Predictive policing algorithms are racist. They need to be dismantled. MIT Technology Review.

Jahankhani, H., et al. (2020). *Policing in the era of AI and smart societies*. Champ: Springer.

Lockey, S., Gillespie, N., and Curtis, C. (2020) Trust in artificial intelligence: Australian insights. The University of Queensland and KPMG Australia.

McGuire, M., 2020. The laughing policebot: automation and the end of policing. *Policing and society*, 31 (1), 20-36.

Moses, B.L. and Chan, J. (2016) Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and society*, 28 (7), 806-822.

Oswald, M., et al. (2018). Algorithmic risk assessment policing models: lessons from the Durham HART model and 'experimental' proportionality. *Information & communications technology law*, 27 (2), 223-250.

Rovatsos, M., Mittelstadt, B., and Koene, A. (2019). *Bias in algorithmic decision making*. The University of Edinburgh: Centre for Data Ethics and Innovation. UK Government.

Solove, D. (2011) *Nothing to hide: the false tradeoff between privacy and security chapter in NOTHING TO HIDE: THE FALSE TRADEOFF BETWEEN PRIVACY AND SECURITY*. New Haven and London: Yale University Press.

Turner, E., Medina, J., and Brown, G. (2019). Dashing hopes? The predictive accuracy of domestic abuse risk assessment by police. *The British journal of criminology*, 59 (5), 1013-1034.

Urquhart, L. and Miranda, D., 2022. Policing faces: the present and future of intelligent facial surveillance. *Information & communications technology law*, 31 (2), 194-219.