Science
& Technology
in Policing

National Police Chiefs' Council

Covenant for Using Artificial Intelligence (AI) in Policing

Version 1.1

1 Introduction

The rapid growth of Artificial Intelligence (AI) within policing is unsurprising. The speed and accuracy that AI can bring to police processes make it an attractive way to deliver an effective and efficient service. However, the application of AI can be contentious[1]. Transparency and fairness must be at the heart of what we implement, to ensure a proportionate and responsible use that builds public confidence.

This Covenant outlines a set of principles that forces have agreed will define how it uses AI in its business. They were endorsed by all members of the National Police Chiefs' Council on 28 September 2023. The endorsement means that all developers and users of AI within policing must give due regard to the Covenant's principles. Whilst the implementation of these principles across policing will be an ongoing and evolving area of work, publication of our principles ensure we are acting with transparency from the outset.

1.1 What is Artificial Intelligence?

There is no definitive definition of Artificial Intelligence (Alan Turing Institute, 2021), and AI is often used to refer to related applications such as automation, neural networks, and machine learning. To bring clarity for policing, we adopt the following definitions:

Artificial intelligence (AI) refers to a machine that learns, generalizes, or infers meaning from input, thereby reproducing or surpassing human performance. An example is using image analysis to determine whether a video contains sexual activity with a child. The term AI can also be used loosely to describe a machine's ability to perform repetitive tasks without guidance.

Machine learning (ML) refers to algorithms that leverage new data to improve their ability to make predictions or decisions. ML is a widely used form of AI that has contributed to innovations such as speech recognition and fraud detection.

Advanced Data Analytics (ADA) uses subject matter expertise and techniques that are typically beyond those of traditional business intelligence to extract insights and make recommendations from complex data. The techniques vary widely, from data visualisation to complex linear models to language analytics. An example is the use of Risk Terrain Modelling to quantify environmental factors that shape risk mapping and resource deployments. A policy covering the use of ADA in policing is being written, to be owned by the Data & Analytics Board.

These three forms of compute are not independent; often an analyst will combine ADA with AI or ML to achieve the best outcome.

1.2 Policing's Use of Artificial Intelligence

Policing's use of AI is advancing quickly. All NPCC forces use data analytics and at least 15 forces have 'advanced' data analytics capabilities (NPCC, 2021). Most of our AI applications focus on organisational effectiveness and workforce planning rather than predictive analytics (see the 2021 Home Office Data Analytics Landscape Review for a broad sector analysis). This includes demand management functions such as live triage of incoming 999/111 calls and the automation of data quality assurance tasks. Many of our capabilities also utilise AI in their delivery, such as the identification algorithm in Face Recognition and the safety features within Unmanned Arial Vehicles.

There are also instances where AI is supporting decision making. For example, Avon and Somerset Constabulary use supervised machine learning to assess factors such as likelihood of reoffending, likelihood of victimisation/vulnerability, and likelihood of committing a range of specific offences. Through an app on their mobile devices, neighbourhood officers can instantly access the risk profiles for each offender registered in the force area, which are recalculated daily.

Our ambition for AI technologies, laid out in the national Digital and Science and Technology strategies, recognise: (1) the power of algorithms to achieve a 'step change' in policing efficiency; (2) the 'arms race' we face with criminals who benefit from new technologies; and (3) the need to maintain public confidence through standards, an ethical framework, and independent oversight. Individual use cases for AI are outlined in capability strategies and the NPCC Areas of Research Interest. For example, the APCC/NPCC Digital Forensic Science Strategy makes clear the infrastructure, processing, and trust requirements that are needed for effective digital forensics within policing.

The growth of our use of AI depends critically on building a specialist community. We seek to ensure continuous best practice via the Data Analytics Community of Practice, PDS's Knowledge hub, and community-led initiatives such as Police Rewired. Our communities are also active members of external networks and events, such as DataConnect21, Ordnance Survey Geospatial Hackathon, and the Government Statistical Service Methodology Symposium.

Despite its benefits, there have been concerning examples of the use of AI in policing, where models have been built on data that led them to act disproportionately against a community or race. The 2022 House of Lords Justice and Home Affairs Committee report on new technologies identified similar issues in other countries, concluding that the impact of the long-term use of AI in policing is uncertain, with limited evidence regarding the risk involved. Critically, the fear of unintended consequences and impingement of civil liberties, deservedly or not, is associated with policing's use of AI. Thus, the NPCC commits to a set of principles for guiding the use of AI in policing.

2 Principles of AI in Policing

Policing's AI principles are founded on three sets of guidance: the FAST Principles[2], the OECD AI Principles[3], and the Data Ethics Framework[4]. We apply these to policing with the intent to support an openness to scrutiny, integrity, and public confidence in our use of AI technologies.

Principle A. Lawful: All use of AI will comply with applicable laws, standards, and regulations. This includes all users of AI, ML, ADA and related data processing (e.g., where you are using national data sets as defined by the NPCC) ensuring the use is recorded centrally in the National ROPA.

Principle B. Transparent: All use of AI will be subject to 'Maximum Transparency by Default' (MTbD).

B1. Forces should ensure the public are aware of AI uses. This will typically include publishing an overview of the algorithms used and the known limitations of the training data used. The datasets will be present on the force IAR with allocated Information asset owners.

B2. Where operational or security requirements restrict the ability to share, the AI will undergo scrutiny by appropriate independent assessors (e.g., organised by the Chief Scientific Adviser).

B3. Subject to B2, all AI projects must be able to allow a third-party to: (1) investigate the algorithmic workings, use scenarios, and underlying data from an 'adversarial perspective' [5];

This might require the supplier to provide 'expert' witness/evidence of the tools' operation. All third parties will have appropriate data protection and information security policies in place.

Principle C. Explainable: The ability for any AI to provide an 'explanation' of its output will be a determining factor in its implementation.

C1. The level of explanation expected will be determined by (1) the function it performs (e.g., is it informing a high-impact decision about an individual); (2) the outputs required of it (i.e., who needs to understand what regarding the output and how was this reached).

Principle D. Responsible: All AI that affects the public will have responsible usage policies (i.e., intensions are defined before deployment so that outcomes and impact can be tracked) and procedures to ensure that users do not accept AI outputs uncritically.

D1. The ability of AI to make decisions without a human being part of that decision will be determined by the function that the AI performs.

D2. All AI that effects the public must have a human as the ultimate decision-maker.

D3. All AI will have a human or automatic means of being stopped if it displays unintended or undesired outputs.

D4. Those responsible for AI-enabled systems must proactively mitigate the risk of unintended biases or harms, during initial rollout and as they learn, change, or are redeployed.

Principle E. Accountable: All AI will have a clearly identified individual accountable for its operation and outputs.

E1. All Accountable persons and end-users will be suitably trained in the use of the relevant AI.

E2. The use of AI in policing will be subject to proper governance and oversight at the relevant organisational level.

E3. AI enabled data sets and technology systems will be governed and assured under the same frameworks as wider data processing responsibilities, linking what is used and how it is used to the appropriate IAR and ROPA.

Principle F. Robust: All data used to train, or that is analysed by, an AI will be robust and reliable enough for its intended purpose. This requires assessing, tracking and reporting on the quality of data, by way of recognising that the quality of data dictates the quality of the analysis.

F1. All AI in policing will be used only for the purpose it was designed, trained and authorised for.

F2. With regards to data usage, all data used in Police AI will be subject to a Framework outlined by a force governance board to guard against issues such as bias, unintended proxies, non-representativeness, unfairness, and untimeliness.

F3. the Government Office for Artificial Intelligence's Guidelines for AI procurement must inform contract implementation and management.

The use of AI in policing must also comply with established codes of practice, most notably the College of Policing's Code of Ethics[6], which describes the standards of accountability, fairness, honesty, integrity, leadership, objectivity, openness, respect and selflessness that is expected of all in policing.

All AI in Policing will also be subject to standard organisational technology, architectural, security and usage principles.

3 Governance of AI and the AI Principles

Chief Constables are responsible for the operational deployment of AI technologies to manage threat, harm, and risk. Together with Police and Crime Commissioners (PCC), who are responsible for holding Chief Constables to account, they ensure use is fair and lawful, balancing ethics, right to privacy, unbiased treatment and consent, with the absolute right to a fair trial.

Chief Constables and PCCs receive detailed advice from force ethics committees, with many also using specialist committees for ADA and AI projects. They are supported by the national Data & Analytics Board, which is part of the Digital, Data and Technology Coordination Committee. To embody the Policing AI Principles, these committee should, wherever possible, work in public, be independently chaired, include experts in data ethics and medical ethics, and have community representation. Their remit should include:

- the maintenance of the Covenant
- determining and disseminating the ethical framework that will govern the use of AI
- advising on the ethical impact of a new AI use before its implementation
- promoting non-discriminatory practices in the use of AI
- undertaking regular reviews of the accuracy, reliability, security, safety, performance, evidence-based decision capability and feedback ability of all AI.

A force's use of AI is informed and supported by a system of independent scrutiny, national peer support, and evidence-based guidance. The National Data and Analytics Board provides oversight, governance and support on issues of Data Quality, Data Protection and Freedom of Information, Records Management, information Sharing, Disclosure and Safeguarding, and Geographical Information portfolios. The NPCC Lead for Ethics provides oversight of ethical principles. The College of Policing are seeking to publish Authorised Professional Practice (APP) on the use of new technologies, which will guide force activity.

All forces have access to a national independent panel for assessment of complex cases. This is currently being provided by the Home Office Biometrics and Forensics Ethics Group (BFEG) with work ongoing to determine the uptake of this provision and whether it is sufficient for the needs of Chief Constables, PCCs, and their forces. BFEG provides advice and is not intended to replace the responsibility and authority of forces and PCCs to make decisions on the use of AI in their force.

3.1 Engaging the Broader AI Landscape

Policing recognises the body of work being conducted in AI across Government and commits to contribute to, and adding value to, these efforts. We will work with the:

- **Office for Artificial Intelligence**, to ensure concordance with the UK's national AI strategy.
- **Data Standards Authority (DSA),** which works to improve how the public sector manages data.
- **Central Digital and Data Office (CDDO),** whose **Data Ethics Framework (DEF)** gives principles to guide the design of appropriate data use in the public sector.

- **Centre for Data Ethics and Innovation (CDEI)** for support on developing and understanding the Ethical frameworks within which to operate AI.
- **AI Council (AIC),** whose independent members provide advice to government and high-level leadership of the AI ecosystem, to raise awareness of policing's efforts.
- **Alan Turing Institute (ATI),** for research opportunities across disciplines and generate impact, both through theoretical development and application to real-world problems.
- **Defence Science and Technology Labs (Dstl)**, via the Police Integration Hub, to identify opportunities and lessons from the use of AI within defence.
- **Industry,** via established routes, to increase their understanding of our requirements of universal AI development, and conversely, to better understand and support their trajectory of travel. Being secure by design and maintaining data sovereignty should not be an excuse for the absence of data sharing and co-working.

It will be duplicative and burdensome for each force to engage directly with such bodies. As such, the Office of the Policing Chief Scientific Adviser will maintain relationships with such bodies and should be contacted first by forces for advice and guidance. The benefit of this approach is it could act as two-way conduit to share policing insight and challenges, as well as emerging standards.

3.2 The Role of the CSA on AI

The Office of the Policing Chief Scientific Adviser (OPCSA) provides systems leadership, advice, and assurance on the use of science and technology in policing. With AI expertise sitting across government and within industry and academia, the CSA's challenge function is critical for AI. OPCSA will, *inter alia*,

(1) ensure that there is effective governance of the development and use of AI in policing, reflecting on the degree to which activity meets the expectations of the AI principles outlined here and in the UK Research Integrity Concordat;

(2) support an active programme of research examining the fairness of AI applications, using independent testers such as the National Physical Laboratory (e.g., operational testing of Face Recognition);

(3) providing accurate and up to date information on AI applications, share successful examples of good practice when using AI in policing. Where early opportunities for the responsible use of AI in police forces are identified, OPCSA will seek to support these projects and provide investment to ensure their success and ensure that their findings are shared widely;

(4) convene academia, industry, and government agencies e.g., the Home Office AI and Data Ethics team on matters related to the use of AI in policing. As part of this activity, consider development of a national data capability that could be accessed by industry and academia to test and train AI tooling in a safe environment;

(5) support the College of Policing in developing Approved Professional Practice (APP) in this area; and,

(6) promote, where operationally appropriate, an ethos of transparency and engagement with the public to maintain and promote trust and confidence.

4 References

[1] Christie (2021). AI in policing and security.
[2] Alan Turing Institute (2023). Frequently Asked Questions.
[3] OECD (2019). The OECD Artificial Intelligence (AI) Principles.
[4] UK Central Digital and Data Office (2018). Data Ethics Framework.
[5] Oswald, M., Grace, J., Urwin, S., & Barnes, G. C. (2018). Algorithmic risk assessment poicing models: Lessons from the Durham HART model and 'experimental' proportionality. *ICTL*, *27*, 223-250.
[6] College of Policing (2020). *Policing in England and Wales Future Operating Environment 2040*.