



HM Inspectorate  
of Probation

## Artificial Intelligence in Probation

Elizabeth McClory-Tiarks, Claire Paterson-Young,  
Michael Maher and Michele Sevegnani

---

HM Inspectorate of Probation

Academic Insights 2026/06

JULY 2026

# Contents

---

- Foreword..... 3
- 1. Introduction ..... 5
- 2. AI in probation: uses, opportunities, and risks ..... 6
  - 2.1 Uses of AI in probation ..... 6
  - 2.2 Opportunities and possibilities of AI ..... 8
  - 2.3 Risks and challenges of using AI..... 9
  - 2.4 Toward a bespoke responsible AI checklist ..... 12
- 3. Conclusion..... 14
- References ..... 15

## Foreword

---

HM Inspectorate of Probation is committed to reviewing, developing and promoting the evidence base for high-quality Probation and youth justice services. *Academic Insights* are aimed at all those with an interest in the evidence base. We commission leading academics to present their views on specific topics, assisting with informed debate and aiding understanding of what helps and what hinders Probation and youth justice services.

This report was kindly produced by Elizabeth McClory-Tiarks, Claire Paterson-Young, Michael Maher and Michele Sevegnani. It explores the current and potential uses of Artificial Intelligence (AI) within the Probation Service, highlighting the opportunities presented and the challenges raised for practice, decision-making, and service delivery. The direction of travel is clearly one of increasing experimentation, with AI-driven tools having been proposed in the areas of information retrieval, transcription and summarisation, risk assessment, sentence planning, resource allocation, compliance monitoring, and early identification of reoffending risks.

The potential for AI to reduce costs whilst increasing speed and efficiency in the criminal justice process is attractive in the current environment. More specifically, AI tools provide opportunities to automate tasks and reduce workload, freeing up practitioners to focus upon the uniquely human aspects of their role, notably developing good relationships with people on probation. At the same time, as is set out in this paper, the introduction of AI into probation raises ethical, legal, and professional challenges. Attention needs to be given to risks and challenges in relation to (i) bias, discrimination and privacy, (ii) transparency and overreliance, and (iii) due process and accountability. It is essential that each proposed use of AI is carefully assessed, avoiding assumptions about the benefits within current working environments and infrastructures, and also considering the broader risks/benefits, recognising that there can be implications for other related or linked tasks or processes.

To maximise benefits and minimise risks, the value of developing a bespoke Responsible AI framework is clearly demonstrated, promoting the ethical, lawful, and accountable use of AI within probation. Such a framework can provide a mandatory decision-making tool prior to AI adoption and a learning/review framework that enables adaptation, evaluation, and withdrawal where harms emerge, providing safeguards and protections for both practitioners and people on probation.



**Dr Robin Moore**

Head of Data Analysis and Research

## Author profiles

**Dr Elizabeth McClory-Tiarks** (BA DipLaw MA PhD) is a Lecturer in Criminal Justice at Newcastle University. Elizabeth has over 20 years' experience working in the field of criminal justice. Elizabeth practised as a criminal barrister before undertaking a PhD in Law exploring restorative justice and sentencing, and then moved into academia. Her research interests are in sentencing, theories of punishment, and AI and criminal justice.

**Professor Claire Paterson-Young** (BA MSc PhD) is a Professor of Social Justice at the University of Northampton. Her research interests are social justice, youth studies, participatory research, child sexual exploitation, social impact measurement, and AI in the criminal justice system. Claire has over 15 years practice and management experience in safeguarding, child sexual exploitation, trafficking, sexual violence, youth and restorative justice. Claire is Chair of the University of Northampton Research Ethics Committee and a serving member of the West Midlands Police and Crime Commissioner Ethics Committee. She is a trustee of the National Association for Youth Justice (NAYJ) and Research Affiliate at the Vulnerability & Policing Futures Research Centre. Claire is currently an Associate Editor for the *Journal of Child and Family Studies*, International Advisory Board Member for the *YOUNG Journal*, and Editorial Board Member for the *Journal of Qualitative Criminal Justice and Criminology*.

**Dr Michael Maher BA MA PhD** is an early-career, post-doctoral researcher at the University of Northampton. He has published work on participatory research, organisation theory, and the third sector. He has also contributed to a significant body of work on the impact of AI in high-risk sectors, including criminal justice. A large part of his work has involved ensuring research accessibility, developing videos and booklets that highlight key findings from research for public audiences.

**Dr Michele Sevegnani** is a Senior Lecturer in Computing Science at the University of Glasgow. His research focuses on formal methods for stochastic and AI-enabled systems, with applications to safety-critical and socio-technical contexts. His current interests and activities include the responsible use of probabilistic AI in policing, criminal justice and law enforcement, and the development of approaches for analysing, explaining and governing AI systems in high-stakes settings.

The views expressed in this publication do not necessarily reflect the policy position of HM Inspectorate of Probation

# 1. Introduction

---

The use of Artificial Intelligence (AI) in the criminal justice system has grown substantially in recent years, with examples of AI found at almost every stage of the criminal process.<sup>1</sup> This Academic Insights paper examines the current and potential future adoption of AI in the Probation Service, considering the benefits that AI can bring, as well as the risks and challenges presented by its use in this context.

The potential for AI to reduce costs, whilst increasing speed and efficiency in the criminal justice process is particularly appealing in the current environment – the system has been described as being ‘in crisis’ (Leveson Review Part 2, 2026). This includes a backlog of criminal cases (Ministry of Justice, 2026c), the Probation Service experiencing strain, and prisons becoming increasingly crowded and difficult to manage (HM Inspectorate of Prisons, October 2025; and December 2025). Furthermore, the Probation Service is to be further tested following the recommendations in the Independent Sentencing Review (2025) and the implementation of the *Sentencing Act 2026* which increases the use of community-based sentences. It is acknowledged in the Independent Sentencing Review (2025, p.121) that ‘the Probation Service is likely to be managing an increasing number of high-risk cases in the community as a result of the Review’s recommendations’.

This paper will consider the risks and benefits of AI tools within this context of significant and increasing pressures on the Probation Service. In this paper, AI encompasses a range of distinct technologies, including:

- embedding-based semantic retrieval
- statistical prediction models
- generative systems (e.g. large language models (LLMs)).

These differ significantly in capability, autonomy, and risk profile, and are distinguished where relevant. While the paper focuses upon the Probation Service, the risks and benefits can also be considered in the context of youth justice services (see Youth Justice Board, 2026, for a set of key principles).

---

<sup>1</sup> See for example this interactive map of AI Tools in the Criminal Justice System in England and Wales developed by the PROBABLE Futures research team found here: <https://probablefutures.github.io/ai-mapping/tools/?cjstages=5>.

## 2. AI in probation: uses, opportunities, and risks

### 2.1 Uses of AI in probation

#### Current use

Within the Probation Service, AI-driven tools have been proposed for use in risk assessment, sentence planning, resource allocation, compliance monitoring, and early identification of reoffending risks. The *AI Action Plan For Justice* (Ministry of Justice, 2025a) was released with a statement from Lord Timpson:

*“we will harness the power of AI to transform the public’s experience, making their interactions with the justice system simpler, faster, and more tailored to their needs. This plan focuses on three priorities: strengthening our foundations, embedding AI across justice services, and investing in the people who will deliver this transformation...”*

These proposed developments align with broader trends towards digitisation, data-driven decision-making, and performance management in public services (see [Academic Insights 2025/03](#) for a review of responsible and ethical digitisation in probation).

There is currently limited use of AI within the Probation Service, but nDelius is one example where AI has been integrated and is intended to reduce the administrative workload of probation staff. NDelius is a browser-based probation case management system that stores information relating to people on probation, including pre-sentence reports, community orders, and custodial sentences pre- and post-release. Its primary function is to allow probation practitioners to access and review case information. As much of the information needed by practitioners exists in large volumes of unstructured text, the platform has integrated a LLM-based semantic search function, designed by Mixedbread AI (Ministry of Justice, 2025b). This tool helps practitioners locate semantically similar case notes based on the user’s search criteria, even when different wording or misspellings are used. The National Inspection of the Probation Service (HM Inspectorate of Probation, 2025) further noted a move toward semantic search and other AI-based information retrieval tools for a similar platform ‘EQUIP’, which provides policies and practice guidance to over 9,000 users a month.

The perceived reduction in administrative burden has been a central theme of the integration of AI within the Probation Service. Tools such as Microsoft’s Co-Pilot and OpenAI’s ChatGPT Enterprise have been trialled through a pilot programme, supported by training resources and role-specific guidance to help staff use them effectively (UK Parliament, 2025a; Ministry of Justice, 2025a). In addition, Justice Transcribe, an in-house AI transcription and meeting summarisation tool was introduced (UK Parliament, 2025b) and, between 07 October 2025 to 02 June 2026, the system summarised over 800,000 meetings with a reported 50% reduction in note-taking time (Justice AI Unit, n.d.; Ministry of Justice, 2026a).

The current Offender Assessment System (OASys) does not incorporate AI but is useful to consider as a contrast with emerging AI systems, highlighting the distinction between actuarial/statistical tools and contemporary machine learning approaches to decision support. Initially introduced in 2002, OASys is a risk and needs assessment tool and is used, amongst other things, to estimate an individual’s likelihood of reoffending (Prison and Probation Services, 2002). It is a structured actuarial assessment framework, rather than an AI-driven system, and incorporates statistical scoring components (such as the Offender Group

Reconviction Scale (OGRS), recently renamed as the All Reoffending Predictor (ARP)) and draws predictions from national data (Moore, 2015). Weighted scoring is applied after collecting relevant data, allowing practitioners to evaluate a range of risk factors to produce an assessment of an individual's reoffending risk. In 2024, it was noted that as offending patterns change the tools need revalidation to correct model miscalibrations (Emeagi et al., 2024; Craik et al., 2024).

Importantly, OASys is scheduled to be replaced during 2026 by Assessing Risks, Needs and Strengths (ARNS) (Ministry of Justice, 2023). The new system responds in part to pressures within the Probation Service to reduce workload, alongside policies such as 'Probation Reset' which prioritise supervision during the periods where reoffending risk is at its highest (HM Inspectorate of Probation, 2025). ARNS produces risk assessments, risk management plans, and pre- and post-sentence plans, drawing on actuarial reoffending data. It also introduces automated features, such as text message appointment reminders to improve compliance and tools to streamline the administrative processes to increase efficiency within the Probation Service. It is believed that ARNS will also provide the digital infrastructure for the further integration of AI (Ministry of Justice, 2025a).

While AI use within probation services remains limited, the potential for expansion is clearly well recognised. The Ministry of Justice's *AI Action Plan for Justice* (2025a) emphasises reducing staff workload as a key objective, though its implementation remains modest in scope and confined to administrative innovations. As the infrastructure develops with the introduction of ARNS and pressures to reduce workloads continue to be felt, it is likely that further AI and digital transformation within probation practice will occur. This will be explored further in the next section, which looks ahead to potential future uses of AI in probation.

### **Potential future uses**

The *AI Action Plan for Justice* sets out a vision for the possible beneficial uses of AI, including the following in probation:

- automating routine tasks to enable staff to focus on meaningful interactions
- resource planning
- personalisation of rehabilitation plans
- AI-enabled predictive tools for risk assessments.

In conjunction with this, recommendations from two recent reviews provide some insight into how AI might be adopted in probation services in the near future. The Independent Sentencing Review (ISR) (May 2025) and the Leveson Review (Independent Review of the Criminal Courts, July 2025 and February 2026) both include in their recommendations proposals to use AI to increase the speed and efficiency with which criminal cases are dealt with. The more specific recommendations made in the ISR relating to probation cover three areas:

- Firstly, tools aimed at reducing the administrative burden on probation officers – specifically mentioned is AI-enabled notetaking as a way to enable officers to spend less time manually taking notes and more time 'building relationships'. As outlined above, one such tool – 'Justice Transcribe' – has now been rolled out across the Probation Service (Justice AI Unit, n.d.; Ministry of Justice, 2026a).
- Secondly, it is proposed that AI tools be used to support monitoring and risk analysis of people on probation, such as using AI-enabled risk and needs assessment tools that can determine the suitability of an individual for community-based sentences, by analysing data such as behavioural patterns, compliance history and other factors. The ISR (2025,

p.136) suggests that this could be combined with facial recognition to track an individual's movements in public spaces. This is with a view to these tools enabling probation officers to detect warning signs, adjust supervision levels and intervene when needed. This could enable more intense supervision and 'secure environments outside prison for offenders' (ISR, 2025, p.125).

- Thirdly, the ISR proposes use of AI tools to personalise probation services to people's differing needs, risk levels and in different situations: 'mobile-based rehabilitation apps and tools could be used to enable an offender to connect with the Probation Service and access information more easily, all in one place on their phone' (ISR, 2025, p 125). It is not clear whether this would involve AI, although 'biometric check-ins' using facial biometrics on a mobile phone are discussed; and the ISR recommends consideration of sensor-based monitoring technologies (e.g. location tracking or device-based behavioural indicators) and remote check-ins, referencing a pilot scheme in a UK prison (ISR, 2025, p.138).

More generally, the ISR recommends further collaboration with industry on research and development of AI tools. One suggestion in the ISR is the development of an AI assistant, which would fulfil certain functions, such as 'offering low-risk individuals basic support, generating personalised rehabilitation plans and analysing compliance data to predict breaches. ... [the tool] could connect people with local services such as housing, employment, education or mental health support' (ISR, 2025, p.135). The inspiration for this is drawn from a similar tool currently being trialled for prison leavers (SherlockAI) and an AI chatbot developed by academics in the US for supervising people released on parole (University of Cincinnati, 2023).

The Leveson Review does not make any specific recommendations for the use of AI in the Probation Service but makes recommendations for use of AI in the criminal justice system more generally, which may have implications for AI adoption in probation, and may also impact the work of probation services through the way that information at earlier points in the criminal justice system is processed and used. For example, the Leveson Review recommends that the police and CPS use AI summarisation tools in the preparation of case files and case summaries, which could have implications for the amount of information available to a probation practitioner preparing a pre-sentence report. The Leveson Review also proposes the use of AI translation tools by police and in courts, which may mean that a probation practitioner at court would also find themselves reliant on AI translation tools, rather than a human translator, to communicate with an offender at court.

## 2.2 Opportunities and possibilities of AI

The main claims made by AI companies and proponents of AI are centred around the use of AI tools to increase efficiency and reduce costs. These claims are particularly attractive in the current criminal justice context of scarce resources and the need to increase the speed with which cases are dealt with to tackle backlogs. The direction of travel is one of increasing experimentation with, and adoption of, AI tools. To date, there has been particularly notable experimentation with AI tools in policing, including the use of tools for auto transcription and summarisation, video footage searches, and live facial recognition (PROBABLE Futures, 2026).

The potential benefits of AI for the Probation Service in particular are considered by the ISR, which identifies the need to both prioritise resources and prioritise opportunities to build relationships with people on probation to help with their rehabilitation. AI and other technological tools are seen as a way to reduce the amount of time that probation practitioners spend on 'paperwork and process' (ISR, 2025, p.121). Similarly, the Leveson Review notes that

'tools that automate tasks reduce the administrative burden on staff, therefore increasing their capacity and accelerating case throughput' (Leveson Review, 2025). AI tools therefore appear to offer the opportunity to automate tasks which take probation practitioners away from the uniquely human aspect of their role, developing good relationships with people on probation which is important for rehabilitation. AI tools are also proposed as a way to improve the accuracy of risk assessments and support case management that is more tailored to the individual (Whitmore, 2025), as outlined in the previous section. This would save practitioners time, allow for better resource allocation, and would enable probation to meet future demands on the service.

It is important to note that the potential benefits of AI tools need to be assessed in terms of how realistic they are in the current context (i.e. of large workloads and limited time and resources) and whether they can perform as claimed within current working environments and infrastructures. It may be that an AI tool cannot achieve certain time or efficiency savings in the current environment, but with additional resources directed to, for example, updating of systems, an overall benefit can be achieved.

Each proposed use-case also needs careful assessment; assumptions should not be made about increases in efficiency or other benefits. For example, the assumption that automating administrative tasks such as note-taking will lead to probation practitioners building better relationships with people on probation (ISR, 2025). This may be the case, or it may be that manual notetaking offers practitioners more engagement with those supervised and is part of the process of building relationships. It also cannot be assumed more generally that the administrative burden will always be reduced by the introduction of AI. In some cases, it might *increase* the administrative burden of the user, where it creates (for example) a need to carefully check tool outputs (Calder, McClory-Tiarks et al., forthcoming).

It could also be that the administrative burden is decreased for the user, but increased elsewhere. For example, Hertfordshire Constabulary trialled the use of LLMs to generate witness statements, which reduced the administrative burden on the police officers taking the statement, but increased the burden on the Disclosure officer and the CPS, as they had to disclose all versions of the statements produced by officers using the AI tool (Walley and Glasspoole-Bird, 2025). This particular example highlights the need to consider the risks/benefits of introducing an AI tool more broadly, rather than a narrow focus on the risks/benefits relating to the discrete task it might be introduced to assist with. It should also be asked what the implications are for other related or linked tasks later in the criminal justice process. This indicates a need for caution with the agency-specific procurement process recommended by the Leveson Review: 'Given the different budgetary pressures, priorities and operational requirements across the criminal justice system, I again recommend each agency either develops or procures its own tailored suite of embedded AI tools to ensure relevance and usability' (Leveson Review, 2025, p. 199).

### 2.3 Risks and challenges of using AI

The introduction of AI into probation raises ethical, legal, and professional challenges (Council of Europe, 2024a; 2024b). Probation operates at the intersection of punishment, rehabilitation, public protection, and human rights (Canton, 2013), and relies on professional judgement, relational engagement, and contextual understanding (Canton, 2013; Robinson et al., 2014). Decisions made by probation practitioners affect liberty, family life, employment, and long-term social outcomes for individuals under supervision (Canton, 2013). Embedding AI into this

environment risks reshaping how risk, responsibility, and accountability are understood and enacted.

This section provides a critical overview of the key risks and challenges associated with implementing AI within the Probation Service, with particular emphasis on:

- bias, discrimination and privacy
- lack of transparency and overreliance on technology
- due process and accountability.

### **Bias, discrimination and privacy**

There is the potential for biased or inaccurate outputs arising from multiple sources, including data quality, labelling practices, model design choices, opaque algorithms, and the reproduction of existing inequalities within the criminal justice system (Babuta, 2017; Babuta and Oswald, 2021; Perry, 2013; Santos, 2017). Individuals under probation supervision already experience intrusions into their private lives. A risk of AI use in probation is its potential to perpetuate and amplify existing biases embedded in historical data (Joseph, 2025). AI systems are typically trained on large datasets derived from past criminal justice decisions, including arrests, convictions, sentencing outcomes, and compliance records. These datasets reflect longstanding structural inequalities related to race, gender, class, disability, and geography (Lammy, 2017; Arowosegbe, 2023). When AI systems learn from biased data, they may reproduce discriminatory patterns while presenting them as neutral or objective (Joseph, 2025; Council of Europe, 2024b). Specific examples are as follows:

- Individuals from racialised communities are disproportionately policed, arrested, and sentenced in the UK. As a result, AI tools trained on these data may systematically assign higher risk scores to Black and minority ethnic individuals, regardless of actual behaviour or rehabilitation progress (Howard and Borenstein, 2018). Evidence from international contexts illustrates these risks, with the COMPAS risk assessment tool used in the United States being shown to overestimate reoffending risk for Black defendants and underestimate it for White defendants (Angwin et al., 2022). While the UK probation system differs structurally, similar tools such as offender assessment systems and reoffending prediction models face analogous challenges if bias is not explicitly addressed.
- Potential bias impacts on women as AI systems designed around male-dominated datasets may fail to account for all differences, leading to inappropriate supervision levels or interventions (Corston, 2007; Kelani, 2024).
- Socioeconomic bias may further disadvantage individuals from deprived communities, where higher levels of police contact and enforcement skew risk predictions.

Crucially, AI-driven bias can become self-reinforcing where higher risk classifications lead to increased surveillance and enforcement, which in turn generate more data confirming perceived risk. This feedback loop risks embedding discrimination deeper into probation practice.

There is also the risk of function creep, where data collected for one purpose is repurposed for others, such as policing or immigration enforcement (Brayne, 2021). This can undermine trust between individuals and probation practitioners, making engagement and rehabilitation more difficult (Brayne, 2021). From a legal perspective, extensive data processing engages rights under the UK General Data Protection Regulation (UK GDPR) and the European Convention on Human Rights, particularly Article 8 (right to private and family life). Ensuring compliance

requires robust safeguards, transparency, and clear limits on data use – requirements that may be difficult to meet when using complex AI systems.

### **Lack of transparency and over-reliance**

Machine learning models, particularly complex ones such as neural networks, often function as 'black boxes', meaning that their internal reasoning processes are not readily interpretable, even by their developers (Burrell, 2016). In probation, transparency is required in explaining decisions to courts, individuals under supervision, victims, and inspectors. Individuals subject to probation supervision have the right to understand how decisions affecting their liberty and privacy are made, and when AI systems produce risk scores or recommendations without clear explanations, this undermines procedural fairness and trust (Rudin, Wang, and Coker, 2020; Grimmelikhuijsen, and Meijer, 2022).

The problem is not merely technical but institutional. Commercial AI systems may be protected by intellectual property rights, preventing full disclosure of how decisions are generated (Popa and Paterson-Young, 2025). Even internally developed systems have complexities that frontline staff may be unable to understand. As a result, probation practitioners may be asked to rely on outputs that they cannot fully explain or challenge but are expected to defend. This lack of transparency also hampers effective oversight as inspectors, courts, and policymakers can struggle to evaluate whether AI systems are operating fairly, accurately, and lawfully (Popa and Paterson-Young, 2025). In a system where legitimacy and public confidence are already fragile, opaque decision-making risks further erosion of trust (Crewe, 2011).

Another risk is the overreliance on AI-generated outputs at the expense of professional judgement and critical thinking (Weaver et al., 2026; Council of Europe, 2024b). Probation work involves nuanced assessments of behaviour, motivation, relationships, and social context (Canton, 2013; Robinson et al., 2014). While AI can support decision-making, there are risks that practitioners may defer to algorithmic recommendations (Cobbe 2019; Edwards and Veale 2017; Zerilli et al., 2019). This phenomenon, sometimes referred to as 'automation bias', occurs when users place undue trust in automated systems, assuming they are more accurate or objective than human judgement (Cobbe 2019; Edwards and Veale 2017; Zerilli et al., 2019; Young et al., 2021). This risk could be heightened in a context of high workloads and increasing pressures, leading to inappropriate supervision levels, unsuitable interventions, or excessive enforcement actions.

Overreliance on AI poses risks in narrowing the scope of probation practice. If risk scores become the dominant lens through which individuals are viewed, strengths-based and rehabilitative approaches may be sidelined. The relational and supportive aspects of probation, widely recognised as central to desistance, risk being overshadowed by technocratic risk management (McNeill, 2012). Furthermore, practitioners may feel constrained by AI outputs, particularly in high-risk cases where deviation from algorithmic recommendations could expose them to organisational or reputational risk.

### **Due process and accountability**

Individuals subject to probation have the right to due process, meaning that they can understand, challenge, and seek redress for decisions affecting their rights and freedoms (Gelsthorpe, 2013). When decisions are informed by opaque algorithms, individuals may be unable to access meaningful explanations or contest the basis of those decisions. Risk scores or recommendations may be treated as authoritative, even when based on flawed data or inappropriate assumptions which undermines principles of fairness and equality before the law.

There are concerns that AI may shift decision-making power away from courts and practitioners towards technical systems, reducing opportunities for individualised consideration (Dancy and Zalnieriute, 2025). In probation, where personal circumstances and change over time are critical, AI systems that are static or inflexible may fail to reflect progress or desistance. Individuals with limited digital literacy, language barriers, or cognitive impairments may be particularly disadvantaged in navigating AI systems – exacerbating existing inequalities (Calder, McClory-Tiarks et al., forthcoming). Although there is evidence that such systems can help enhance digital literacy (Parra-Valencia and Massey, 2023), the potential role they could play for individuals supported by probation requires further investigation.

AI does not only affect the rights of individuals under supervision but may also have implications for due process within probation as a profession. Professional judgement, reasoning and discretionary decision-making are core components of probation identity and effectiveness (Canton, 2013; Robinson et al., 2014). Although systems are not yet established in practice within probation, evidence from other public sector settings suggests that decision-supporting technologies can influence professional judgment (Selten et al., 2023). When AI systems produce recommendations that implicitly or explicitly guide practice, practitioners may feel pressured to conform, even where they believe the output is inappropriate (Cobbe 2019; Edwards and Veale 2017; Zerilli et al., 2019; Young et al., 2021). Notably, studies have found that professionals may lack the technical knowledge required to fully understand or critically scrutinise complex outputs (Eubanks, 2018), which can create tensions between professional ethics and organisational expectations. Challenging AI outputs may thus require technical expertise that probation practitioners do not possess, placing them at a disadvantage in organisational decision-making. This raises questions about procedural fairness within institutions and the capacity of probation services to retain professional autonomy.

A final challenge concerns accountability. If AI systems cause harm through biased decisions, errors, or misuse then it can be unclear who is responsible. Responsibility may be diffused across developers, data providers, policymakers, and frontline users – creating significant ethical and legal uncertainty. In other words, if a harmful decision is influenced by an AI system:

- is the practitioner accountable for following it?
- is the organisation responsible for deploying it?
- or does liability lie with the developers who designed the system?

This diffusion of responsibility risks creating accountability gaps where harms are acknowledged but not adequately addressed. Clear governance frameworks, audit mechanisms, and lines of responsibility are essential but currently underdeveloped in many AI initiatives.

## 2.4 Toward a bespoke responsible AI checklist

The use of AI within the criminal justice system is expanding rapidly, with policing leading the development of governance frameworks to ensure ethical, lawful, and proportionate use. The National Police Chiefs' Council (NPCC) *Responsible AI Checklist* (2025) represents an example of good practice, offering a structured mechanism for assessing risks, legality, ethics, and accountability prior to and during AI deployment. It provides a strong and credible model for ensuring ethical, lawful, and accountable use of AI in policing, while its emphasis on governance, transparency, human oversight, and continuous review represents sector-leading good practice within UK criminal justice.









However, probation differs fundamentally from policing. While policing AI is often oriented towards intelligence, detection, and operational decision-making, probation AI could directly shape supervision, rehabilitation and risk escalation with sustained impacts on liberty, family life, and desistance. As such, probation requires a bespoke responsible AI framework, rather than direct adoption of policing models. It should adapt the principles and strengths of the NPCC *Responsible AI Checklist* (2025) and pay attention to the Council of Europe recommendation and explanatory memorandum on the use of AI by prison and probation services (Council of Europe, 2024a; 2024b), while also embedding the principles within HMPPS governance structures, HM Inspectorate of Probation expectations, and frontline probation practice. A probation responsible AI checklist should be published for full transparency and function as a:


- mandatory decision-making tool prior to AI adoption
- live governance document throughout the lifecycle of the AI system
- professional safeguard for practitioners
- rights-protective mechanism for individuals under supervision
- learning framework that enables adaptation, evaluation, and withdrawal where harms emerge.

As outlined earlier, the Leveson Review makes a recommendation that criminal justice agencies should develop or procure their own AI tools, to ensure they are tailored to the needs and tasks of each agency. This has implications for the Probation Service in making decisions about allocation of resources and whether AI tools offer sufficiently significant benefits to warrant apportioning funds in that direction, rather than another, for example staffing. This agency-specific approach to procurement of AI may mean that a bespoke responsible AI Checklist for probation is even more important to support decision-making in this area.

## Need for a Responsible AI Framework


To maximise benefits and minimise risks, a bespoke Responsible AI framework ensures the ethical, lawful, and accountable use of AI within probation.

 <p><b>Ethical Safeguards</b></p>  <p>Ensure AI systems are fair, unbiased and respect the rights and dignity of people on probation.</p>	 <p><b>Human Oversight</b></p>  <p>Maintain meaningful human control and professional judgement at every stage.</p>	 <p><b>Accountability</b></p>  <p>Clear governance, documentation and accountability for AI decisions and outcomes.</p>	 <p><b>Ongoing Evaluation</b></p>  <p>Continuously monitor, evaluate and adapt AI tools to ensure they remain safe and effective.</p>
---	---	--	---



**This framework provides:**

- ✓ A mandatory **decision-making tool** prior to AI adoption.
- ✓ A **learning and review framework** that enables adaptation, evaluation, and withdrawal where harms emerge.



**Safeguarding practitioners and people on probation.**

### 3. Conclusion

---

The use of AI in the Probation Service is currently limited to assistance with administrative tasks, but the expansion of AI beyond this is underway. Government policy and independent reviews of sentencing and the criminal justice system recommend that AI must be further incorporated into the criminal justice process, including probation, as part of measures to ease the burden on the system. In particular, there are recommendations for a move towards using AI in risk assessment, compliance monitoring, and early identification of reoffending risks.

This paper has considered the benefits which may be achieved in increasing efficiency and enabling better use of probation practitioners' time to deal with tasks only a human can do. These benefits are particularly appealing in the current context of large caseloads and limited resources. However, the same context which makes AI tools attractive may also undermine the effectiveness of some tools, if they cannot function as claimed within current working environments and infrastructures. It is important that the use-case for any proposed AI tool is considered carefully, to ensure that there is overall benefit in the existing context. This should extend to consideration of how the integration of a tool in the Probation Service may impact at other stages of the criminal justice process.

This paper has outlined the risks arising from an increased use of AI, particularly risk assessments and LLMs, in the Probation Service, detailing concerns about: bias, discrimination and privacy; lack of transparency and over-reliance on AI; and issues around due process and accountability. Taking into account these risks, as well as the potential benefits if a proper assessment is carried out to enable the appropriate integration of AI, this paper recommends a bespoke responsible AI framework for Probation, which could be adapted from the NPCC *Responsible AI Checklist* (2025) and which should then be published. This could help to ensure the best quality decision-making in a high-stakes environment, maximising benefits and minimising risks.

## References

---

- Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2022). 'Machine bias', in Martin, K. (ed.) *Ethics of Data and Analytics*. Auerbach Publications, pp. 254-264.
- Arowosegbe, J.O. (2023). 'Data bias, intelligent systems and criminal justice outcomes', *International Journal of Law and Information Technology*, 31(1), pp. 22-45.
- Babuta, A. 2017. *Big Data and Policing: An Assessment of Law Enforcement Requirements, Expectations and Priorities*. Royal United Services Institute for Defence and Security Studies (RUSI).
- Babuta, A. and Oswald, M. 2021. 'Machine learning predictive algorithms and the policing of future crimes', in McDaniel, J.L.M. and Pease, K.G. (eds.) *Predictive policing and artificial intelligence*. Abingdon, Oxford: Routledge.
- Brayne, S. (2021). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.
- Burrell, J. (2016). 'How the machine 'thinks': Understanding opacity in machine learning algorithms', *Big Data & Society*, 3(1), <https://doi.org/10.1177/20539517156225>.
- Calder, M., McClory-Tiarks, E., Oswald, M., Sevegnani, M. and Taka, E. (forthcoming), 'Responsible AI in Criminal Justice: LLMs in policing and risks to case progression'.
- Canton, R. (2013). 'The point of probation: On effectiveness, human rights and the virtues of obliquity', *Criminology & Criminal Justice*, 13(5), pp. 577-593.
- Cobbe, J. 2019. 'Administrative law and the machines of government: Judicial review of automated public-sector decision-making', *Legal Studies*, 39(4), pp. 636-55.
- Corston, J. (2007). *The Corston Report: A Review of Women with Particular Vulnerabilities in the Criminal Justice System*. London: Home Office.
- Council of Europe (2024a). *Recommendation CM/Rec(2024)5 of the Committee of Ministers to member States regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services*. Available at: [https://cep-probation.org/wp-content/uploads/2018/10/CM\\_Rec20245.pdf](https://cep-probation.org/wp-content/uploads/2018/10/CM_Rec20245.pdf) (Accessed: 11/06/2026).
- Council of Europe (2024b). *Recommendation CM/Rec(2024)5 of the Committee of Ministers to member States regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services – Explanatory Memorandum*. Available at: [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%220900001680b1e593%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%220900001680b1e593%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}) (Accessed: 11/06/2026).
- Craik, A., Han, L., Sullivan, L., Landsiedel, J., Travers, T., Spaul, C. and Howard, P. (2024). *Revalidation: Risk of Recidivism tools: An evaluation of the actuarial instruments developed to assess recidivism risk in England and Wales*. London: Ministry of Justice. Available at: <https://www.gov.uk/government/publications/revalidation-risk-of-recidivism-tools-an-evaluation-of-the-actuarial-instruments-developed-to-assess-recidivism-risk-in-england-and-wales> (Accessed: 07/04/2024).

Crewe, B. (2011). 'Depth, weight, tightness: Revisiting the pains of imprisonment', *Punishment & Society*, 13(5), pp. 509-529.

Edwards, L. and Veale, M. (2017). 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for', *Duke Law & Technology Review*, 18, pp. 18–84.

Emeagi, C., Sullivan, L., Landsiedel, J., and Craik, A., and Howard, P. (2024). *The Actuarial Prediction of Sexual Reoffending: Responding to Changing Offending Patterns*. London: Ministry of Justice. Available at: <https://www.gov.uk/government/publications/actuarial-prediction-of-sexual-reoffending-responding-to-changing-offending-patterns> (Accessed: 07/04/2026).

Eubanks, V. (2018). 'Automating bias', *Scientific American*, 319(5), pp.68-71.

Gelsthorpe, L. (2013). 'Probation values and human rights', in Gelsthorpe, L. and Morgan, R. (eds.) *Handbook of Probation*. London: Willan, pp. 485-517.

Grimmelikhuijsen, S. and Meijer, A. (2022). 'Legitimacy of algorithmic decision-making: Six threats and the need for a calibrated institutional response', *Perspectives on Public Management and Governance*, 5(3), pp. 232-242.

HM Inspectorate of Probation (2025). *National Inspection – April 2025*. Manchester: HM Inspectorate of Probation. Available at: <https://hmiprobation.justiceinspectores.gov.uk/document/national-inspection-april-2025/> (Accessed: 07/04/2025).

Howard, A. and Borenstein, J. (2018). 'The ugly truth about ourselves and our robot creations: the problem of bias and social inequity', *Science and Engineering Ethics*, 24(5), pp. 1521-1536.

Joseph, J. (2025). 'Predicting crime or perpetuating bias? The AI dilemma', *AI & SOCIETY*, 40(4), pp. 2319-2321.

Justice AI Unit (2025). *Speech AI*. Available at: <https://ai.gov.uk/knowledge-hub/tools/speech-ai/> (Accessed: 07/04/2026).

Kelani, Z.A. (2024). 'Invisible Women: Data Bias in a World Designed for Men', *Journal of International Women's Studies*, 26(5), pp.1-6.

Knight, V. (2025). *Carefully and humanely progressing responsible and ethical digitisation in probation*, HM Inspectorate of Probation Academic Insights 2025/03. Available at: <https://hmiprobation.justiceinspectores.gov.uk/document/carefully-and-humanely-progressing-responsible-and-ethical-digitisation-in-probation/> (Accessed: 08/06/2026).

Lammy, D. (2017). *The Lammy Review: An independent review into the treatment of, and outcomes for, Black, Asian and Minority Ethnic individuals in the Criminal Justice System*. London: Lammy Review.

McNeill, F. (2012). 'Four forms of 'offender' rehabilitation: Towards an interdisciplinary perspective', *Legal and Criminological Psychology*, 17(1), pp.18-36.

Ministry of Justice (2023). *Justice and Home Affairs Committee Evidence Session*. Available at: <https://committees.parliament.uk/publications/41868/documents/207632/default/> (Accessed: 07/04/2026).

Ministry of Justice (2025a). *AI Action plan for justice*. Available at: <https://www.gov.uk/government/publications/ai-action-plan-for-justice> (Accessed: 07/04/2026).

Ministry of Justice (2025b). *MoJ: nDelius Contact Log Semantic Search*. Available at: <https://www.gov.uk/algorithmic-transparency-records/moj-ndelius-contact-log-semantic-search> (Accessed: 07/04/2026).

Ministry of Justice (2025c). *Independent Sentencing Review: Final report*. Available at: <https://www.gov.uk/government/publications/independent-sentencing-review-final-report> (Accessed: 16/05/2026).

Ministry of Justice (2025d). *Independent Review of the Criminal Courts: Part 1*. Available at: <https://www.gov.uk/government/publications/independent-review-of-the-criminal-courts-part-1> (Accessed: 16/05/2026).

Ministry of Justice (2026a). *Justice Transcribe Data*. Available at: <https://www.gov.uk/government/publications/justice-transcribe> (Accessed: 08/06/2026).

Ministry of Justice (2026b). *Independent Review of the Criminal Courts: Part 2*. Available at: <https://www.gov.uk/government/publications/independent-review-of-the-criminal-courts-part-2> (Accessed: 16/05/2026).

Ministry of Justice (2026c). *Criminal Justice Statistics Quarterly: October to December 2025*. Available at: <https://www.gov.uk/government/statistics/criminal-court-statistics-quarterly-october-to-december-2025/criminal-court-statistics-quarterly-october-to-december-2025> (Accessed: 16/05/2026).

Moore, R. (2015). *A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013*. London: Ministry of Justice.

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

Popa, D.M. and Paterson-Young, C. (2025). *Testing comprehensive frameworks for AI in heightened risk contexts*. Available at SSRN 5312787.

Prison and Probation Services (2002). *Offender Assessment System OASys: User Manual*. Available at: <https://prisons.org.uk/OASys-Manual.pdf> (Accessed: 07/04/2026).

PROBABLE Futures (2025). *Interactive map of AI Tools in England and Wales*. Available at: <https://probablefutures.github.io/ai-mapping/tools/?cjstages=5> (Accessed: 16/05/2026).

Robinson, G., McNeill, F. and Maruna, S. (2013). 'Punishment in society: The improbable persistence of probation and other community sanctions and measures' in Simon, J. and Sparks, R. (eds.) *The SAGE handbook of punishment and society*. SAGE Publications Ltd, pp.321-340.

Rudin, C., Wang, C. and Coker, B. (2020). 'The age of secrecy and unfairness in recidivism prediction', *Harvard Data Science Review*, 2(1), p.1.

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

Selten, F., Roberer, M. and Grimmelikhuijsen, S. (2023). 'Just like I thought': Street-level bureaucrats trust AI recommendations if they confirm their professional judgment', *Public Administration Review*, 83(2), pp.263-278.

UK Parliament (2025a). *HM Prison and Probation Service: Artificial Intelligence* [Online] Available at: <https://questions-statements.parliament.uk/written-questions/detail/2025-01-29/27101> (Accessed: 07/04/2026).

UK Parliament (2025b). *Ministry of Justice: Artificial Intelligence* [Online]. Available at: <https://questions-statements.parliament.uk/written-questions/detail/2025-09-01/73679> (Accessed: 07/04/2026).

University of Cincinnati (2023). *UC-developed chatbot to help supervision officers, parolees*. Available at: <https://www.uc.edu/news/articles/2023/06/uc-developed-chatbot-to-help-supervision-officers-parolees.html> (Accessed: 28/05/26).

Walley, P. and Glasspoole-Bird, H. (2025). *An Evaluation of the pilot application of Artificial Intelligence to Witness Statement and Report Generation at Hertfordshire Constabulary*. Centre for Policing Research and Learning, Open University.

Weaver, B., Gillon, F., Belton, I. and Heron, G. (2026). 'What do I not know? How does AI affect critical thinking?', *Criminology & Criminal Justice*, <https://doi.org/10.1177/17488958261453>

Whitmore, L. (2025). 'Unlocking Potential: How can AI Support Probation', *Probation Quarterly*, 38, pp. 39-43.

Young, M.M., Himmelreich, J., Bullock, J.B. and Kim, K-C. (2021). 'Artificial intelligence and administrative evil', *Perspectives on Public Management and Governance*, 4(3), pp. 244-58.

Youth Justice Board (2026). *Artificial intelligence (AI) use in youth justice services: Key principles*. Available at: <https://yjresourcehub.uk/artificial-intelligence-ai-use-in-youth-justice-services-key-principles/> (Accessed: 11/06/26).

Zerilli, J, Knott, A., Maclaurin, J. and Gavaghan, C. (2019). 'Algorithmic decision-making and the control problem', *Minds & Machines*, 29, pp. 555-78.

**© Crown copyright 2026**

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit [www.nationalarchives.gov.uk/doc/open-government-licence](http://www.nationalarchives.gov.uk/doc/open-government-licence) or email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third-party copyright information, you will need to obtain permission from the copyright holders concerned.

This publication is available for download at:  
[www.justiceinspectors.gov.uk/hmiprobation](http://www.justiceinspectors.gov.uk/hmiprobation)

**Published by:**

HM Inspectorate of Probation  
1st Floor Civil Justice Centre  
1 Bridge Street West  
Manchester  
M3 3FX

The HM Inspectorate of Probation Data Analysis & Research Team can be contacted via [HMIProbationResearch@hmiprobation.gov.uk](mailto:HMIProbationResearch@hmiprobation.gov.uk)

ISBN: 978-1-917531-11-5