Risk Control

Artificial Intelligence in the Public Sector

In partnership with

GALLAGHER BASSETT
GUIDE. GUARD. GO BEYOND.

Risk Management Partners
rmpartners.co.uk

Artificial Intelligence in the Public Sector

Introduction

Artificial Intelligence (AI) is rapidly transforming the operational landscape of public sector services across the United Kingdom. Its integration into the workflows of local councils, fire and rescue services, and policing, offers the potential to streamline operations, increase efficiency, and support data-driven decision-making. While many AI applications promise to optimise resource allocation and enhance user experiences, the introduction of such technologies also presents complex legal, ethical, and practical challenges. For organisations entrusted with public safety and welfare, the stakes are high, necessitating measured and inclusive implementation[1, 2].

This guidance provides a comprehensive overview of how AI is being used in UK public sector services. It outlines practical examples, identifies associated risks, and proposes responsible approaches to governance. The intended audience includes professionals working in or with public sector service bodies, particularly those in frontline roles, policy development, or digital transformation. It aligns with regulatory expectations and draws from relevant UK Government frameworks to ensure lawful and ethical deployment[3, 4].

Overview of AI in Public Sector Practice

Artificial Intelligence encompasses a range of technologies that allow machines to simulate human intelligence — processing information, making decisions, and learning from data. Common examples include machine learning algorithms, natural language processing, and computer vision systems[1]. In UK public sector services, AI is increasingly being used to enhance operational planning, improve public engagement, and support predictive modelling in resource deployment[4, 5].

Local authorities, for instance, have begun using AI-powered chatbots to manage routine service requests, automate housing assessments, and forecast demand for public sector services[1]. In fire and rescue services, AI tools are used to model incident likelihood based on environmental data, enabling proactive fire prevention strategies[6]. Police forces leverage AI in digital forensics, video surveillance analysis, and strategic deployment of officers based on crime trends[5, 7]. Similarly, social services apply AI to prioritise cases, identify early intervention opportunities, and analyse outcomes across demographics[4].

These examples demonstrate a fundamental shift in how services are delivered and accessed. However, because many decisions made using AI can directly impact people's lives, public bodies must balance innovation with caution[2].

Strategic Benefits of AI Adoption

The adoption of AI in public sector services is driven by several compelling advantages. Chief among them is operational efficiency. AI systems can automate time-consuming administrative tasks, such as form processing, appointment scheduling, or routine triage. This enables human professionals to redirect their expertise toward higher-value, complex, or emotionally sensitive cases that require judgment and empathy[1].

Another key benefit is enhanced responsiveness. AI tools such as virtual assistants and automated messaging platforms allow public sector services to offer support outside of traditional working hours. Members of the public can report issues, receive guidance, or access information at any time, improving satisfaction and service coverage, particularly for individuals with caregiving responsibilities, unpredictable work schedules, or accessibility needs[4].

AI also supports data-driven policymaking. Through predictive analytics and real-time modelling, public bodies can anticipate service demands, identify emerging risks, and allocate resources more effectively. Fire services can forecast peak risk periods based on historical patterns and weather data[6]. Police forces can map hotspots using crime statistics, enabling targeted and timely intervention[5]. Local councils can estimate public health needs using regional demographics, environmental indicators, and behavioural trends[1].

Moreover, AI has the potential to improve accessibility and inclusion. Technologies such as real-time translation, automatic transcription, and content simplification support diverse users, including those who are neurodivergent or speak English as an additional language[8]. When designed with care, these features can significantly reduce access barriers, improving equity in service provision.

Finally, AI enables proactive approaches to safety and risk management. For emergency services, the ability to simulate incidents and plan resource deployment can save lives and reduce disruption[6]. In social care, AI systems that flag indicators of neglect or financial hardship allow for earlier intervention and more holistic support planning[4].

Human-Centred Design and Ethical Implementation

Despite the efficiency gains, it is essential that AI in public sector services is implemented with a strong commitment to human-centred design. These sectors often involve nuanced, sensitive interactions where individual context and professional judgment are critical. Decisions in policing, social services, or emergency response must uphold dignity,

privacy, and equality — values not easily replicated by automated systems[5, 8].

To achieve ethical implementation, public bodies must ensure transparency about AI use. This means informing individuals when their data is being processed by automated systems and providing clear explanations of how decisions are made[9]. Transparency is particularly vital in situations involving eligibility for services, legal enforcement, or resource allocation.

Equally important is accountability. AI systems should never operate in isolation when the consequences of decisions are high. A human-in-the-loop approach, where professionals retain final oversight, ensures that errors, exceptions, or ethical considerations can be addressed promptly[1, 3]. Staff must be equipped to understand AI outputs, question anomalies, and intervene when necessary.

Public involvement plays a critical role in trust-building. Engaging service users and stakeholders in the design, review, and evaluation of AI systems helps ensure these technologies align with community expectations and needs[4]. This is particularly important in marginalised or underserved populations, where historical inequities may influence outcomes and system design.

To support frontline professionals, training should be provided that equips staff with foundational digital literacy, ethical reasoning, and interpretive skills[3]. This enables employees to understand how AI functions, assess its limitations, and apply its insights effectively. Structured training plans should be inclusive and tailored to different roles and responsibilities.

Frameworks such as the UK Government AI Playbook[3] and the Algorithmic Transparency Recording Standard (ATRS)[1, 9] offer practical support in achieving these goals. They outline design principles, documentation standards, and public communication strategies that can strengthen ethical oversight and legal compliance.

Risks and Ethical Dilemmas in AI Deployment

Alongside its benefits, AI introduces a range of risks that must be considered and mitigated. First among these is the potential for privacy violations. AI systems frequently rely on large volumes of personal data, including health records, behavioural information, and demographic indicators. Improper handling of this data can breach UK General Data Protection Regulation (GDPR), and the Data Protection Act 2018[10], particularly if systems are not clearly documented or subject to regular review[2].

Another pressing concern is algorithmic bias. If the data used to train AI systems reflects historical inequalities, these patterns may be perpetuated or even intensified[4, 5]. For instance, facial recognition tools have faced criticism for lower accuracy rates among certain ethnic groups. Predictive policing models based on historical arrest data may unfairly target specific communities[7]. These risks are particularly salient in services where public trust is paramount, and discriminatory outcomes carry serious consequences[8].

Opacity, or the "black box" problem, is also a concern. Many AI systems are complex and difficult to interpret, making it hard for users, managers, or auditors to understand how decisions were reached. This lack of transparency can hinder accountability and limit recourse for affected individuals[9].

Over-reliance on automation poses further risk. Professionals may defer to AI outputs without critical evaluation, leading to de-skilling or unjustified decision-making. Conversely, in crisis situations where human empathy and discretion are paramount, rigid adherence to AI guidance may be inappropriate or harmful[3].

Job displacement is another consideration, particularly in administrative roles. While automation can reduce repetitive tasks, it may also lead to workforce restructuring. Public bodies must plan carefully to upskill employees, redeploy talent, and ensure that transitions are managed equitably and with respect for labour rights[2].

Taken together, these risks underscore the need for robust governance structures and ethical safeguards, not as an afterthought, but as a core component of strategic planning[3].

Bias Mitigation and Inclusive Practice

Public sector service organisations have a duty to promote equality, and this extends to their use of AI. Ensuring that AI systems are fair, inclusive, and representative requires deliberate action[8]. One foundational step is to source diverse datasets when developing machine learning models. Data should reflect the full spectrum of community experiences, avoiding skewed results based on geography, ethnicity, gender, or socio-economic status.

Technical interventions also support bias mitigation. Fairness-aware algorithms can adjust outputs based on detected disparities, while regular audits can identify unintended consequences or performance gaps[4]. In practice, this means tracking outcomes by protected characteristic and responding to disparities with appropriate policy or technical adjustments.

Explainability is another key factor. AI systems should be designed in ways that make their logic understandable to users and stakeholders. This includes providing documentation, example cases, and context-sensitive guidance. The ATRS is a useful tool for this purpose, offering a structured format to record purpose, design logic, input / output mechanisms, and review procedures[9].

Ongoing user engagement is critical, particularly with communities at risk of exclusion or harm. Service users should be invited to provide feedback, participate in pilot schemes, and suggest improvements. Engagement methods should be inclusive, ranging from online forums and surveys to in-person workshops or community briefings[4].

By embedding these practices, public bodies can ensure that AI serves as a tool for inclusion, not exclusion, and supports more equitable service provision[8].

Governance and Legal Compliance

Effective governance is essential to the responsible use of AI in public sector services. At a minimum, organisations must ensure compliance with relevant legislation. The UK GDPR governs personal data handling[1], while the Equality Act 2010 requires public bodies to actively promote equality and prevent discrimination[8]. The Freedom of Information Act 2000 supports transparency in decision-making and public accountability[9].

Documentation of AI systems should be consistent and comprehensive. This includes details of system purpose, decision logic, technical specifications, and risk mitigation strategies. The ATRS offers a useful framework for this documentation, making it easier to share information internally and with regulators[1].

Organisations should also appoint clear oversight roles. This may involve designating a Senior Responsible Owner (SRO) for AI governance, establishing an ethics committee, or integrating AI considerations into existing risk management structures[3]. Such oversight enables consistent monitoring, encourages cross-departmental collaboration, and ensures that AI deployment aligns with organisational deployment goals. Governance should also extend to procurement practices. When sourcing AI systems from external vendors, public bodies must specify requirements for transparency, interpretability, and ethical safeguards. Contracts should include provisions for documentation, training support, and ongoing access to performance metrics. Where appropriate, public buyers should seek open standards or interoperable designs to reduce dependency and promote flexibility[2].

AI risk registers should be developed and maintained, identifying specific use cases, associated risks, and mitigation strategies. This documentation enables consistent oversight across departments and supports compliance during audits or regulatory reviews. Local authorities may consider adopting cross-sector governance boards to harmonise approaches and share lessons learned[4].

Ultimately, legal compliance should be viewed not as a constraint but as a framework for building trust. By aligning with established laws and ethical standards, public bodies reinforce the legitimacy of AI use and safeguard the rights of service users[3, 8].

Risk Mitigation and Responsible Innovation

Effective risk mitigation begins with early-stage planning and continues throughout the system lifecycle. Prior to deployment, organisations should conduct impact assessments that examine legal, social, and operational implications. These assessments should be structured, evidence-based, and informed by stakeholder input. In high-risk scenarios, such as AI used in policing, safeguarding, or health triage, independent review is advisable[2, 6].

During deployment, human oversight must be embedded into workflows. This may involve decision escalation pathways, interpretability dashboards, or case review protocols that allow staff to contest or override AI outputs. Systems should be designed with guardrails that flag anomalies, detect unintended consequences, and prevent algorithmic drift[3].

Public communication plays a critical role in risk management. Service users should have access to plain-language summaries that describe AI systems in use, their purpose, and how they may affect outcomes. These communications should be accessible, translated where necessary, and distributed via multiple channels, including service portals, correspondence, and community engagement[1, 9].

Feedback loops are essential to continuous improvement. AI systems must be monitored over time, with mechanisms to receive and respond to concerns raised by users or staff. This includes structured incident reporting, satisfaction surveys, and regular user forums. Insights from these sources should feed into model updates, training refreshers, and governance adjustments[4].

Innovation should be guided by prudence. While pilot schemes and proof-of-concept initiatives offer valuable learning, they must be accompanied by appropriate safeguards, evaluation frameworks, and exit plans. By treating risk mitigation as an enabler of innovation, rather

than a barrier, public bodies can foster experimentation that is safe, inclusive, and responsive[4, 6].

Conclusion

Artificial Intelligence is poised to become a cornerstone of modern public sector service delivery in the UK. Its capacity to enhance operational efficiency, improve predictive insights, and support accessible services offers immense promise across local councils, fire and rescue, and policing. However, this promise comes with significant responsibility.

Responsible AI deployment requires more than technical competence. It demands human-centred values, legal awareness, and inclusive engagement strategies. Public bodies must balance speed and ambition with caution and accountability. They must empower staff and service users to understand and shape AI systems, ensuring that technology enhances, not replaces, professional judgment. To support this, organisations should conduct impact assessments including DPIA and EIA prior to deployment [2, 3]; embed human oversight mechanisms into decision workflows[1, 4]; develop and maintain AI risk registers with documented use cases and controls[3]; specify interpretability and ethical safeguards in procurement contracts[2]; provide plain-language summaries of AI systems to service users[1, 9]; create feedback loops through user forums, reporting channels, and monitoring reviews[4]; deliver staff training on AI literacy, ethics, and escalation protocols[3]; engage diverse stakeholders in design and testing, prioritising inclusive accessibility[5, 8]; and use pilot schemes with safeguards and evaluation plans to support responsible innovation[2, 6].

Through principled design, robust oversight, and evaluation shaped by those most affected, UK public sector service organisations have an opportunity to set the standard for socially accountable AI; demonstrating that digital systems can reflect public values, operate transparently, and operate within legal frameworks.

References

1. UK Government – https://www.gov.uk/data-ethics-guidance/artificial-intelligence-in-public-services
2. Geldards LLP – https://www.geldards.com/insights/ai-in-the-uk-public-sector-use-cases-and-regulatory-overview/
3. Government Digital Service – https://gds.blog.gov.uk/2025/02/10/launching-the-artificial-intelligence-playbook-for-the-uk-government/
4. Local Government Association – https://www.local.gov.uk/our-support/cyber-digital-and-technology/artificial-intelligence-hub
5. National Police Chiefs' Council – https://www.npcc.police.uk/SysSiteAssets/media/downloads/publications/publications-log/science-and-innovation/2025/npcc-ai-strategy.pdf
6. AI Knowledge Hub – https://ai.gov.uk/knowledge-hub/
7. College of Policing – https://library.college.police.uk/docs/NPCC/Artificial-intelligence-playbook-policing-2025.pdf
8. Equality and Human Rights Commission – https://www.equalityhumanrights.com/guidance/public-sector-equality-duty-and-data-protection?return-url=https%3A%2F%2Fwww.equalityhumanrights.com%2Fsearch%3Fkeys%3Dpublic%2Bsector%2Bequality%2Bduty
9. Local Government Lawyer – https://www.localgovernmentlawyer.co.uk/projects-and-regeneration/317-projects-features/61259-ai-in-the-uk-public-sector
10. UK Government – Data Protection Act 2018 and UK GDPR - https://www.gov.uk/data-protection

Further information

For access to further RMP Resources you may find helpful in reducing your institution's cost of risk, please access the RMP Resources or RMP Articles pages on our website. To join the debate follow us on our LinkedIn page.

Get in touch

For more information, please contact your broker, RMP risk control consultant or account director.

contact@rmpartners.co.uk

Risk Management Partners

The Walbrook Building
25 Walbrook
London EC4N 8AW

020 7204 1800
rmpartners.co.uk

FP1152-2025