

Artificial Intelligence (AI) Governance Policy for NHS Provider Organisations

Policy Title	Artificial Intelligence (AI) Governance Policy	
Version:	Ver 2.0	
Summary:	The purpose of this policy is to establish clear and comprehensive guidelines for the safe, lawful, and ethical development, implementation, and monitoring of Artificial Intelligence (AI) systems within NHS settings. It aims to protect personal data, uphold clinical and ethical standards, ensure transparency, and mitigate associated risks throughout the AI lifecycle.	
Ratified by	<i>Provider Policy Ratification Group</i>	Date: [Insert Date]
Has an Equality Impact Assessment been carried out?	Yes	Date: [Insert Date]
Name of originator/authors:	Dr Krishna Nair, Director of Quality and Innovation, Chief Clinical Information Officer and Clinical Safety Officer. Innovate Health Consulting Limited. www.InnovateHealthConsulting.com	
Executive Director	Executive Director, Provider Organisation	
Name of responsible committee:	<i>Provider Clinical Quality Committee.</i>	
Target audience:	All staff, students, bank colleagues, contractors, temporary workers, and third-party providers who are involved in the consideration, development, implementation, or use of AI technologies within the organisation.	
Date issued:	9 th April 2025 (Version 2.0)	
Next Review date:	9th April 2028 (or sooner if significant changes in regulation or practice occur)	

Version Control

Version No	Type of Change	Date	Description of Change
1.0	Initial draft	12/10/2024	New Policy drafted by Krishna Nair
2.0	Revised draft	09/04/2025	Generic version for the AI Governance policy developed by Krishna Nair, as requested by NHS England North

Contents

1. Introduction	3
2. Purpose.....	3
3. Scope of the Policy	4
4. Explanation of Terms.....	5
5. Duties	10
5.1. Individual colleagues	10
5.2. Senior Information Risk Owner (SIRO).....	10
5.3. Chief Clinical Information Officer	10
5.4. Clinical Safety Officer	10
5.5. Managers.....	11
5.6. Responsible committee	11
5.7. IT Directorate	11
5.8. Chief Executive.....	11
5.9. Caldecott Guardian.....	12
5.10. Information Governance (IG) Lead	12
5.11. Research and Development (R&D) Lead	12
6. Policy Document Requirements	13
6.1. Using AI produced content	13
6.2. Freely Available Artificial Intelligence Apps and Services.....	13
6.3. Procuring and implementing artificial intelligence products or systems that include AI features:	14
6.4. Defining the purpose and identifying a legal basis for the use of AI:.....	16
6.5. Developing Artificial Intelligence Products for Healthcare.....	17
6.6. Using AI for Research.....	17
7. Equality Impact Assessment	18
8. Consultation Process	18
9. Dissemination and Implementation.....	19
9.1. Dissemination	19
9.2. Competence/Training	19
10. Monitoring Compliance	19
11. References.....	20
12. Associated Policy Documentation	21
13. Appendices.....	22

1. Introduction

Artificial Intelligence (AI) has emerged as a transformative technology with significant potential to enhance patient care, improve clinical outcomes, optimise NHS resource management, and deliver operational efficiencies. As NHS (Provider) increasingly adopts innovative digital solutions, including AI and machine learning technologies, it becomes imperative to ensure these are deployed in a manner that safeguards patient safety, respects data privacy, and upholds public trust.

(NHS Provider) acknowledges the considerable benefits and the concurrent ethical, clinical, and data protection responsibilities associated with AI adoption. This policy provides a structured governance framework to ensure the appropriate, responsible, and transparent use of AI systems across NHS services and activities, including commissioned services delivered by external partners.

The integration of robust clinical safety and risk management processes is fundamental within this policy, recognising the unique challenges and hazards that AI technologies may introduce. This policy incorporates NHS Clinical Risk Management standards (DCB0129 for manufacturers/developers and DCB0160 for deploying organisations) as foundational elements to guarantee AI implementations meet rigorous safety criteria, thus protecting patients and clinicians alike.

Through this AI Governance Policy, (NHS Provider) commits to promoting best practices, enabling standardisation across Integrated Care Systems (ICSs) and key stakeholders, and aligning local initiatives with national guidance issued by bodies such as the NHS AI Lab, the Information Commissioner's Office (ICO), Medicines and Healthcare products Regulatory Agency (MHRA), and other relevant regulatory and professional organisations.

(NHS Provider) recognises that governance of AI systems is an evolving field, thus this policy will undergo regular review to ensure continued alignment with emerging technologies, evolving regulatory landscapes, and developing ethical considerations.

By clearly defining roles, responsibilities, standards, and procedures, (NHS Provider) aims to build confidence in AI systems, ensuring they deliver measurable benefits to patients, clinicians, and the wider healthcare ecosystem while mitigating associated clinical and information governance risks.

2. Purpose

The purpose of this policy is to establish clear guidelines for the responsible and effective development, implementation, use, and ongoing monitoring of Artificial Intelligence (AI) systems within (NHS provider organisation). This policy ensures that AI technologies are utilised in a manner that safeguards patient safety, protects personal data, upholds ethical standards, and mitigates potential risks.

(NHS provider organisation) acknowledges that AI technologies—including machine learning, deep learning algorithms, natural language processing (NLP), and generative AI—hold significant promise to enhance patient care, optimise clinical outcomes, streamline operational workflows, and reduce administrative burdens. AI can support clinical decision-making, personalise patient care, improve

diagnostic accuracy, and facilitate resource management, ultimately leading to safer, more efficient, and higher quality healthcare services.

However, to fully realise these benefits, it is essential to ensure strict adherence to legal and regulatory frameworks, such as the UK General Data Protection Regulation (UK GDPR), Data Protection Act 2018, NHS Digital Clinical Safety Standards (DCB0129 and DCB0160), and relevant professional guidance. Additionally, ethical principles including fairness, transparency, accountability, and patient autonomy must be prioritised to maintain public trust and confidence in the NHS.

This policy sets out fundamental principles and procedural requirements covering key aspects such as data protection, clinical risk management, algorithmic transparency, bias mitigation, and accountability. By clearly articulating roles, responsibilities, and standards for AI use, this policy aims to promote consistency, safety, and transparency across all NHS services.

Recognising the evolving landscape of AI technology and governance, (NHS provider organisation) commit to periodically reviewing and updating this policy to remain aligned with advances in technology, emerging regulatory requirements, ethical standards, and best practices. This continuous improvement approach ensures NHS providers stay at the forefront of responsible AI implementation, thereby safeguarding patient welfare, enhancing public confidence, and promoting ethical innovation within healthcare.

3. Scope of the Policy

This AI Governance Policy applies to (NHS provider organisation) and is relevant to all individuals working within or on behalf of these organisations. This includes, but is not limited to, substantive employees, bank staff, temporary workers, students, contractors, suppliers, volunteers, third-party service providers, and other associated personnel who interact with or use AI technologies in the delivery of NHS services.

The policy extends to all departments, clinical services, administrative functions, and operational areas within the provider organisations, regardless of the size, scale, or complexity of the AI systems in use.

The scope specifically encompasses:

- AI technologies developed internally by the provider organisation, including any in-house development, piloting, or customisation of AI systems or models.
- AI technologies procured from external vendors or suppliers, covering all forms of contractual agreements, partnerships, joint ventures, licensing, and subscription-based AI services.
- AI solutions employed for direct patient care, clinical decision support, administrative workflows, operational efficiency, research, or educational purposes.
- Any freely available or commercial AI applications and services used on organisational devices or platforms, subject to compliance with established organisational governance and IT security policies.

All relevant stakeholders must comply with the requirements outlined within this policy, ensuring that AI deployment across NHS providers adheres strictly to clinical safety standards, information governance regulations, data protection legislation, and ethical guidelines.

4. Explanation of Terms

Artificial Intelligence (AI)	AI refers to computer systems capable of performing tasks that historically required human intelligence, such as speech recognition, decision making and pattern identification. AI is an umbrella term that encompasses various technologies, such as machine learning, deep learning, and natural language processing (NLP). When people use the term “AI” today, they often refer to machine learning-powered technologies like chatbots, computer vision, and other systems that enable machines to perform tasks previously exclusive to humans.
Generative Artificial Intelligence (GenAI)	GenAI, also known as generative AI or GAI, is a subset of AI that can generate new content (such as text, images, videos, or other data) in response to prompts. Generative models learn patterns and structures from training data and then create new data with similar characteristics. Examples of GenAI include chatbots like ChatGPT, text-to-image systems like DALL-E, and text-to-video generators. It has applications across various industries, including healthcare, entertainment, and software development.
Data Protection Impact Assessments (DPIA)	A Data Protection Impact Assessment (DPIA) is a process to help identify and minimise the data protection risks of a project. At (NHS Provider), DPIAs must be considered and, where necessary, completed in full for any new data processing activities, new systems, services, and commissioning activities. The Information Governance (IG) Team reviews and approves DPIAs and advises on requirements and recommended actions.
Digital Technology Assessment Criteria (DTAC)	The Digital Technology Assessment Criteria (DTAC) is a framework designed to evaluate digital health technologies in health and social care settings. The DTAC provides a standardised approach assessing the safety, effectiveness, and usability of digital health products.
Machine Learning (ML)	Machine Learning is a field of study within artificial intelligence (AI) that focuses on developing and analysing statistical algorithms. These algorithms learn from data and generalize their knowledge to perform tasks without explicit

	instructions. In other words, ML enables machines to improve their accuracy over time by imitating how humans learn. It finds applications in various domains, including natural language processing, computer vision, speech recognition, email filtering, agriculture, and medicine
Natural Language Processing	Natural language processing is a branch of computer science / AI that gives computers the ability to understand interpret, and generate human language in both written and spoken forms.
Processing	In relation to information or data means; obtaining, recording or holding the information or data or carrying out any operation or set of operations on the information or data, which may include adaptation or alteration of the information; retrieval, or use of the information or data; disclosure of the information or data by transmission, dissemination or otherwise making available, or alignment, combination, blocking, erasure or destruction of the information or data. In summary anything you do with data is “processing”.
Robotic Process Automation (RPA)	Robotic Process Automation (RPA) is a form of business process automation technology that uses software robots to automate tasks typically performed by humans. These software robots, also known as bots, follow predefined workflows to perform repetitive office tasks such as extracting data, filling in forms, and moving files. RPA combines APIs and user interface (UI) interactions to integrate and execute tasks across different enterprise and productivity applications. By deploying scripts that emulate human processes, RPA tools enable autonomous execution of activities and transactions, freeing up human resources to focus on more complex tasks. RPA accelerates digital transformation efforts and enhances return on investment (ROI) by streamlining routine operations. Remember that RPA should not be confused with artificial intelligence (AI). While RPA is process-driven, AI is data-driven. RPA bots follow predefined rules, whereas AI bots use machine learning to recognize patterns in data and learn over time. Both RPA and AI complement each other, with AI enhancing RPA’s capabilities and handling more complex use cases.
Use Case	A use case is a methodology used in system analysis to identify, clarify and organise system requirements. The use case is made up of a set of possible sequences of interactions between systems and users in a particular

	environment and related to a particular goal. The method creates a document that describes all the steps taken by a user to complete an activity.
AI Hallucinations	An AI hallucination is a phenomenon where an artificial intelligence model generates or perceives patterns, objects, or meanings in data that do not actually exist. This has been observed across various AI models and continues to be prevalent today.
Shadow AI	Shadow AI refers to the use of artificial intelligence (AI) systems and tools within an organisation without explicit approval or oversight from the IT or data governance teams. This often involves colleagues using generative AI tools like ChatGPT or other AI applications to perform tasks such as drafting documents, creating images, or creating code, without the knowledge or consent of (NHS Provider)'s IT department.
AI Bias	In the context of artificial intelligence (AI), bias refers to the tendency of an AI system to produce results that are systematically prejudiced due to erroneous assumptions in the machine learning process. This can occur due to various factors, including: Biased Data: If the training data used to develop the AI model contains biases, the AI system is likely to replicate those biases in its outputs. Algorithmic Limitations: The design of the AI algorithms themselves can introduce biases, especially if they are not properly calibrated to handle diverse data. Human Influence: The biases of the developers and data scientists who create and train the AI systems can inadvertently become embedded in the models. These biases may result in unfair or discriminatory outcomes, affecting various applications such as hiring, lending, law enforcement, and healthcare.
Algorithmic Transparency	Algorithmic transparency refers to the ability to understand and explain how an AI system makes decisions. This includes being able to audit the processes and decisions of

	AI models to ensure they align with ethical standards and regulatory requirements. In healthcare, transparency is crucial to maintain trust and ensure that AI-driven decisions can be justified
AI Ethics	AI ethics encompasses the moral implications of AI technologies, including fairness, transparency, accountability, and non-discrimination. Ethical AI ensures that AI technologies are designed and implemented in ways that protect human rights, promote fairness, and avoid harm, especially in sensitive sectors like healthcare.
AI Lifecycle	The AI lifecycle refers to the stages an AI system goes through from conception to deployment and ongoing monitoring. It includes data collection, model development, training, deployment, performance evaluation, and continuous improvement.
Bias Mitigation	Bias mitigation refers to the processes and techniques used to reduce bias in AI models, including strategies such as data pre-processing, algorithmic adjustments, and post-processing methods that aim to minimise unfairness in AI outcomes.
Automated Decision Making	Automated decision making refers to decisions made by AI systems without human intervention. These decisions are typically based on data inputs and pre-set algorithms. In healthcare, automated decision making can assist in diagnostics, treatment recommendations, or administrative tasks, but must be overseen to ensure compliance with legal and ethical standards.
Data Governance	Data governance refers to the policies and processes in place to manage the availability, usability, integrity, and security of data within an organisation. Effective data governance ensures that data used for AI systems is accurate, secure, and used responsibly in compliance with data protection regulations such as GDPR.
Clinical Risk Management	Clinical risk management involves identifying, assessing, and mitigating potential risks to patient safety that could arise from the use of AI systems. In the context of AI, this means ensuring that the deployment of AI tools does not introduce errors or safety hazards into clinical workflows or decision-making.

Clinical Safety Case Report	A clinical safety case report is a comprehensive document prepared by the Clinical Safety Officer that outlines the potential risks associated with an AI system and the steps taken to mitigate those risks. It includes evidence that the AI system has been designed and tested with patient safety in mind, and it confirms that the AI solution is safe for use in a clinical setting.
Hazard Log	A hazard log is a dynamic document used to record identified hazards associated with AI systems in healthcare. It documents the risks, their potential impact on patient safety, and the controls implemented to mitigate these risks. The hazard log is regularly updated as new risks emerge and as part of ongoing monitoring.
Clinical Safety Standards	Clinical safety standards refer to the regulatory and industry-specific guidelines that AI systems must comply with to ensure they are safe for use in healthcare. In the UK, this would include adherence to the <i>DSC NHS DCB0129</i> and <i>DCB0160</i> standards, which outline safety guidelines for the manufacture and deployment of clinical IT systems, including AI.
Safety Critical Systems	Safety critical systems are systems whose failure or malfunction could result in significant harm to patients. In the context of AI, this refers to any AI system that plays a role in diagnosing, treating, or monitoring patients. These systems require rigorous testing, validation, and ongoing monitoring to ensure they meet the highest safety standards.
Algorithm Validation	Algorithm validation is the process of testing and evaluating an AI model to ensure it performs as intended and that its outputs are accurate, reliable, and safe for use in clinical environments. This includes testing the AI model against real-world clinical data to ensure it delivers accurate results without introducing risks to patients.
Human-AI Collaboration	Human-AI collaboration refers to the interaction between healthcare professionals and AI systems. This includes how AI tools assist clinicians in decision-making while maintaining human oversight and control to ensure patient safety. In clinical settings, AI should support, not replace, healthcare professionals, ensuring that final clinical decisions are made by qualified individuals.

5. Duties

5.1. Individual colleagues

All colleagues have a responsibility to use approved Artificial Intelligence (AI) systems strictly in accordance with established guidelines and organisational protocols. Colleagues are expected to actively provide feedback and share insights regarding the effectiveness, usability, and clinical impact of AI technologies to support ongoing evaluation and improvement. Clinicians, specifically, are accountable for ensuring that their use of AI aligns fully with professional standards and guidance issued by their respective professional bodies. While AI-generated content can support clinical practice, ultimate clinical responsibility remains with clinicians; consequently, all outputs produced by AI systems must undergo rigorous clinician review and approval before being adopted clinically. Furthermore, colleagues have an obligation to promptly report any incidents, risks, concerns, or potential issues related to AI technologies directly to the Information Governance (IG) team, Clinical Safety Officer (CSO), and Senior Information Risk Owner (SIRO).

5.2. Senior Information Risk Owner (SIRO)

The Senior Information Risk Owner (SIRO) holds overarching responsibility for governance and the proactive management of information-related risks associated with AI technologies. It is the SIRO's duty to ensure that robust frameworks, comprehensive risk management processes, policies, and controls are effectively implemented and maintained throughout the organisation. Additionally, the SIRO is expected to work collaboratively with key stakeholders, including the Chief Clinical Information Officer (CCIO), CSO, IG leads, and IT Directorate, to identify, assess, manage, and mitigate risks arising from the use and deployment of AI systems, ensuring that all information governance requirements are rigorously upheld.

5.3. Chief Clinical Information Officer

The Chief Clinical Information Officer (CCIO) provides strategic and clinical leadership for the safe and effective integration of AI technologies across NHS clinical services. It is the CCIO's responsibility to ensure that AI systems align seamlessly with existing clinical workflows, enhance patient outcomes, and are securely embedded into the digital infrastructure. Additionally, the CCIO oversees the adherence to interoperability, clinical safety standards, data governance, and ethical guidelines, and supports initiatives aimed at building clinician competence and confidence in the use of AI. The CCIO also works collaboratively with the CSO, SIRO, and relevant governance committees to ensure AI deployments are appropriately reviewed, approved, and aligned with organisational strategic objectives.

5.4. Clinical Safety Officer

The Clinical Safety Officer (CSO) is responsible for ensuring comprehensive compliance of all AI systems with relevant clinical safety standards and regulations, particularly those outlined in NHS Digital's Clinical Risk Management standards (DCB0129 and DCB0160), as well as obligations under the Health and Social Care Act (2012). The CSO leads the clinical risk management activities associated with AI technologies, which includes identifying, assessing, documenting, and mitigating

clinical risks through detailed hazard logs and Clinical Safety Case Reports. It is also the CSO's duty to ensure adherence to the organisational 'Clinical Safety Policy and Risk Management Procedure – IT Health Systems' throughout the lifecycle of AI technology deployment and operation, actively engaging with clinical, technical, and IG colleagues to maintain continuous oversight of clinical safety.

5.5. Managers

Managers across (NHS provider organisation) hold a pivotal role in ensuring that they, and all colleagues within their line management remit, fully understand and adhere strictly to this AI Governance Policy. Managers must ensure that colleagues are provided with appropriate training, resources, and support necessary to enable the safe, ethical, and effective implementation of AI solutions within their respective areas. Additionally, managers are expected to encourage colleagues to escalate any AI-related incidents, risks, concerns, or opportunities for improvement promptly and effectively to the relevant governance leads or committees.

5.6. Responsible committee

A dedicated governance committee is established to ensure robust oversight, effective implementation, and continuous evaluation and improvement of AI-related policies and practices, including this AI Governance Policy. This governance committee is tasked with providing continuous oversight, conducting regular reviews to ensure policies reflect technological advances, evolving clinical practices, regulatory updates, and emerging ethical considerations. The committee is responsible for undertaking holistic evaluations of AI systems, incorporating diverse perspectives from clinicians, technical experts, Information Governance leads, SIRO, CSO, and CCIO, thereby ensuring comprehensive governance across clinical, technical, ethical, and information governance domains. Furthermore, the committee holds clear accountability for the timely identification, management, and resolution of AI-related issues, ensuring alignment with organisational standards and statutory requirements.

5.7. IT Directorate

The Information Technology (IT) Directorate, in close collaboration with the Clinical Safety Officer (CSO) and Chief Clinical Information Officer (CCIO), is responsible for the technical assessment, assurance, and approval of AI technologies proposed for piloting or deployment within NHS services. Following approval, the IT Directorate supports the technical implementation, secure integration, and ongoing maintenance of AI systems, ensuring appropriate configuration, cybersecurity measures, data protection compliance, and compatibility with existing NHS IT infrastructures. Moreover, the IT teams are responsible for actively collaborating with AI vendors, suppliers, and other relevant stakeholders to address and resolve technical challenges, providing ongoing technical support throughout the AI system lifecycle.

5.8. Chief Executive

The Chief Executive holds ultimate strategic and operational responsibility for ensuring that all procedural documents relating to AI and digital technologies comply fully with relevant legal, statutory, regulatory, and best practice requirements. The Chief Executive is responsible for fostering an organisational culture that supports safe, ethical, and effective AI adoption, ensuring adequate

resources and organisational commitment to achieving excellence in AI governance. The Chief Executive also advocates transparency, accountability, and continuous learning related to AI implementation and governance throughout the organisation.

5.9. Caldecott Guardian

The Caldicott Guardian is responsible for ensuring that the use of patient-identifiable information in AI technologies complies with the principles of confidentiality, particularly where personal data is used for clinical, secondary, or research purposes. The Guardian plays a key role in reviewing proposals involving identifiable data, advising on the lawful sharing of information, and ensuring that privacy and confidentiality risks are appropriately mitigated. In collaboration with the IG Team and CSO, the Caldicott Guardian supports decision-making around ethical data usage in AI systems and advises on safeguards where anonymisation or pseudonymisation may be insufficient.

5.10. Information Governance (IG) Lead

The Information Governance Lead is responsible for ensuring that all AI systems meet data protection, privacy, and information security requirements, in accordance with the UK GDPR, Data Protection Act 2018, and national NHS IG frameworks. The IG Lead works closely with project teams to support the completion of Data Protection Impact Assessments (DPIAs), assess legal bases for data processing, and advise on data minimisation, transparency, and accountability principles. The IG Lead also monitors compliance with information governance obligations throughout the lifecycle of AI technologies.

5.11. Research and Development (R&D) Lead

Where AI technologies are proposed for research or evaluation purposes, the R&D Lead plays a pivotal role in ensuring that such initiatives are reviewed and approved in line with the Health Research Authority (HRA) and NHS Research Ethics Committee (REC) requirements. The R&D Lead provides guidance on research governance, including ethical approvals, consent, and data use in research contexts. They also liaise with external partners, academic institutions, and internal governance teams to ensure that AI research is conducted ethically, legally, and in accordance with organisational policy.

(Note: *The responsibilities outlined in this section may evolve as artificial intelligence technologies and governance practices advance. Effective communication, collaboration, and interdisciplinary working among the roles listed above—and any future governance roles established—are essential for ensuring the safe, lawful, and ethically sound implementation of AI systems within NHS provider organisations.*)

6. Policy Document Requirements

6.1. Using AI produced content

All colleagues need to maintain awareness that content produced by any AI service, tool, or product, may not always be accurate, complete or correct. AI generated content should never replace the need for human review to ensure its accuracy and appropriateness for the specific use case. This applies to all AI content irrespective of the source.

When using AI, all colleagues:

1. Must not rely on AI content as a sole source of truth or factual information, nor as a substitute for professional advice or opinion.
2. Must always evaluate AI content for accuracy and relevance to their specific use case, including using human review, before using, sharing or recording AI generated content.
3. Must not use any AI generated content relating to a patient or individual for purposes that could have a legal or material consequences, such as providing medical advice or significant decisions about that person.

6.2. Freely Available Artificial Intelligence Apps and Services

Artificial Intelligence (AI) applications and services are becoming increasingly prevalent, including freely accessible apps or bots available via app stores and those integrated by default within modern internet browsers. It is essential that all colleagues utilise these publicly available AI tools responsibly to ensure that their use does not compromise personal data, organisational confidentiality, or patient safety, nor breach established organisational policies.

(NHS Provider) advises colleagues to exercise caution when using freely available AI software, such as ChatGPT. While such tools can provide useful supplementary insights or background information, they should not be relied upon as the primary or sole source of information, as these tools may generate inaccurate, biased, or misleading outputs—known as "AI hallucinations".

The UK's National Cyber Security Centre (NCSC) explicitly advises against entering sensitive personal details or organisational intellectual property into publicly accessible AI chatbots. Additionally, colleagues should strictly avoid submitting queries or information that could cause complications or embarrassment if publicly disclosed

When using publicly available AI apps and services, colleagues must adhere to the following requirements:

- No personal data should be entered into these apps or services.
- No business-sensitive or confidential information should be entered into these apps or services.
- These applications must be used strictly for non-clinical purposes only.
- Colleagues must inform the Information Governance (IG) team if they intend to use these services routinely for professional purposes.

- Colleagues must remain mindful of copyright and intellectual property considerations when using generative AI services.
- Users must be aware of potential ethical implications associated with AI outputs, including risks of propagating biased, discriminatory, or harmful content.
- Outputs generated by these AI tools must be carefully verified for accuracy before use or dissemination.

While it is generally recommended that AI software used for work purposes is accessed via corporate devices to maintain robust security and data governance standards, (NHS Provider) recognises that there may be situations where safe and approved AI use need not be strictly constrained to corporate devices. With appropriate safeguards, certain AI apps (such as ChatGPT) can be safely used on personal devices.

To maintain appropriate control and mitigate associated risks, (NHS Provider) will explore enabling approved AI applications on personal devices, subject to the following conditions being met:

- **Approval and Security Measures:** All AI applications intended for use on personal devices must be pre-approved by the IT Department, ensuring compliance with organisational security standards, configured appropriately with defined permissions and restrictions to minimise the risk of accidental data leaks.
- **App Management:** Approved AI applications will be made available through corporate app stores or equivalent secure platforms, providing colleagues with the flexibility to securely download and utilise these applications on personal devices.
- **Compliance:** Usage of personal devices for approved AI applications must remain fully compliant with organisational data governance, security, and privacy policies, ensuring that no personal or business-sensitive data is compromised or inadvertently disclosed.

By enabling controlled and secure use of approved AI apps on personal devices, (NHS Provider) aims to provide greater flexibility to colleagues, balancing convenience and innovation with robust data protection and security requirements

6.3. When procuring and implementing artificial intelligence products or systems that include AI features, the following must be adhered to:

- The use of AI must be transparent to all colleagues and patients, ensuring they clearly understand how, where, and why AI is being used, including the potential impacts on employment, clinical practice, or patient care. The underlying logic and decision-making processes of the AI system must be clearly explainable.
- All AI-generated outcomes or outputs must be reviewed and verified by an appropriately qualified human. AI systems must not be relied upon as sole decision-makers; substantial human involvement is required in all decision-making processes involving AI.

- All procurements must adhere strictly to the procurement processes detailed within the organisation's Standing Financial Regulations and Scheme of Delegation.
- It is strongly recommended by the Information Commissioner's Office (ICO) that a Data Protection Impact Assessment (DPIA) is completed for all AI systems involving personal data. Both the service area and AI solution supplier must engage fully with this DPIA process.
- Colleagues must engage proactively with the IT Directorate to ensure completion of an IT Technical Assurance Assessment before implementation of any AI solution.
- Risks associated with AI systems and appropriate mitigation measures must be carefully considered and documented using the ICO's 'AI and Data Protection Risk Toolkit'.
- If the AI system is associated with healthcare provision, such as diagnostic image interpretation, clinical decision-making support, or other direct patient-care tools, a Digital Technology Assessment Criteria (DTAC) assessment must be completed. This assessment is managed by the Clinical Safety Officer (CSO) in accordance with the organisation's 'Clinical Safety Policy and Risk Management Procedure – IT Health Systems'.
- Any biases or ethical concerns identified during the DPIA and DTAC processes must be explicitly documented, addressed, and mitigated. A thorough assessment of societal impacts and ethical implications of AI deployments is required.
- If the AI system is associated with research, formal approval must be obtained from the Health Research Authority (HRA).
- The Clinical Safety Officer (CSO), and, if applicable, the Medical Device Safety Officer (MDSO)—in cases involving the development of medical devices—must be actively involved and consulted throughout the entire lifecycle of AI products, from initial procurement and implementation through to decommissioning.
- Compliance with Article 22 of the UK General Data Protection Regulation (UK GDPR) concerning automated individual decision-making, including profiling, is mandatory. Individuals have the right not to be subject solely to automated decisions without appropriate human oversight.
- An agreed and clearly documented process must be in place to flag, report, and address concerns regarding the outputs or performance of AI products or systems.
- Data used within AI systems must be collected and processed lawfully, ethically, and transparently, with appropriate consent mechanisms or anonymisation measures in place.
- Access to and sharing of data used by AI systems must be strictly controlled. Data must be securely stored throughout its lifecycle, with robust cybersecurity and data protection measures enforced at all times.

- Patient and public engagement activities must be undertaken proactively to ensure transparency and to assess whether individuals support the intended use of their data, addressing any concerns related to privacy and data usage effectively.
- Where AI implementation involves significant service change, formal consultation must be conducted with employees and their trade union representatives in alignment with the organisation's established Organisational Change Policy prior to implementation.
- Assurance must be obtained from AI suppliers that their products actively mitigate bias and discrimination risks.
- AI systems must be continuously monitored for suspicious activities, anomalies, and potential security breaches. All detected issues must be promptly reported to the appropriate governance or IT security teams for investigation and resolution.

6.4. Defining the purpose and identifying a legal basis for the use of AI:

Prior to piloting or implementing any AI system, it is essential to clearly define and agree upon its purpose and intended use. This should include the rationale for deploying AI and the specific value it is expected to bring to the organisation. Establishing this clarity ensures that AI implementation aligns with organisational priorities, ethical standards, and legal obligations.

A fundamental aspect of this process is determining whether a legal basis for the use of data is required before any processing occurs. Wherever feasible, data should be anonymised, thereby potentially removing the need for a legal basis. However, careful assessment is necessary to determine whether individuals can still be identified, even if common identifiers such as name, address, or phone number are removed.

For example, a combination of data points—such as a local geographical area, a rare medical condition, and a very young age—may render an individual identifiable. In such cases, the data must be treated as personal data, and a valid legal basis for processing must be established. Additionally, organisations must ensure compliance with the common law duty of confidentiality.

These requirements apply equally to data used during the testing, training, or development of AI systems. Even if the data does not inform outcomes or decisions affecting identifiable individuals, it is still considered to be processed and must therefore be managed in accordance with data protection and confidentiality obligations.

AI may be deployed in health and care settings in three principal ways:

- AI specifically designed for use in clinical or healthcare environments;
- AI used for population health management or health-related research;
- Freely available or commercially developed 'generic' AI tools.

Guidance for the responsible use of each of these types of AI in health and care settings is provided in the subsequent sections.

6.5. Developing Artificial Intelligence Products for Healthcare

The NHS's **AI and Digital Regulations Service** provides comprehensive guidance and regulatory resources for individuals and organisations that are developing or planning to adopt artificial intelligence (AI) or digital technologies within health and social care settings. The service consolidates relevant regulations, policies, and frameworks to support safe, effective, and compliant implementation of digital healthcare innovations.

This regulatory service is delivered through a collaboration of four national partners:

- National Institute for Health and Care Excellence (NICE)
- Medicines and Healthcare products Regulatory Agency (MHRA)
- Health Research Authority (HRA)
- Care Quality Commission (CQC)

Although NHS providers may not typically develop AI technologies directly, it is essential for organisations to be aware of this service, particularly given the increasing likelihood of adopting AI tools developed by third-party vendors. Furthermore, providers may in future partner with developers on bespoke innovations, such as the digital solutions designed to improve care delivery

The AI and Digital Regulations Service is therefore a key resource for ensuring alignment with regulatory standards and best practice in the deployment of AI-driven healthcare technologies. Further information is available at: [About the AI and Digital Regulations Service - AI regulation service - NHS \(innovation.nhs.uk\)](https://innovation.nhs.uk/about-the-ai-and-digital-regulations-service-ai-regulation-service-nhs)

6.6. Using AI for Research

Research involving the use of artificial intelligence (AI) within the NHS must be conducted in accordance with established ethical and regulatory frameworks. In England, approval from the Health Research Authority (HRA) is mandatory for all research studies conducted within NHS settings.

The HRA AI and Digital Regulations Service provides tailored guidance for NHS organisations intending to adopt AI technologies, as well as for digital health innovators seeking to develop or evaluate AI tools in research contexts. This includes regulatory advice, ethical considerations, and signposting to relevant standards and resources.

In addition to HRA approval, such research must also be reviewed by an NHS Research Ethics Committee (REC), which ensures that ethical principles are upheld in the design and execution of studies involving patients, data, or NHS staff. The HRA team will further assess the study's compliance with applicable regulations, including those relating to data protection, patient confidentiality, and the use of novel technologies.

Organisations planning to undertake AI research are strongly encouraged to consult their local Research and Development (R&D) Support Services for guidance. Within our organisation, this support is provided through the Innovation, Research and Improvement System (IRIS). The IRIS team offers expert advice on applying for research ethics approval and navigating the HRA's application process, ensuring that research activity meets all necessary standards before commencement

7. Equality Impact Assessment

(As part of its statutory and ethical commitment to equality, diversity, and inclusion, **(NHS Provider)** is dedicated to designing, commissioning, and implementing services, policies, and procedures that recognise and respond to the diverse needs of its service users, local population, and workforce. The organisation seeks to ensure that no individual or group is unfairly disadvantaged or excluded, either directly or indirectly, through its activities or decision-making processes.

To uphold these commitments, an Equality Impact Assessment (EIA) is undertaken as an integral component of the policy ratification and review process. The EIA tool enables the organisation to evaluate whether a proposed policy, function, or service aligns with the three aims of the Public Sector Equality Duty (PSED), as outlined in Section 149(1) of the Equality Act 2010. These are:

1. To eliminate unlawful discrimination, harassment, and victimisation and other conduct prohibited by the Act;
2. To advance equality of opportunity between people who share a protected characteristic and those who do not;
3. To foster good relations between people who share a protected characteristic and those who do not.

Through the use of the EIA, potential impacts on individuals or communities with protected characteristics—such as age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation—are identified, considered, and appropriately addressed. This process supports inclusive policy development and ensures that adjustments or mitigations can be applied where any adverse impacts are identified.

The outcomes and learning from the Equality Impact Assessment have been fully considered and incorporated into the development of this AI Governance Policy. This ensures the policy not only complies with legal obligations but also reflects best practice in promoting equity and inclusion across all areas of the organisation's digital and data-related activities.

8. Consultation Process

A structured consultation process was undertaken during the development of this policy to ensure that it reflects the expertise, perspectives, and governance requirements of the organisation. This collaborative approach supports the development of a robust and inclusive policy framework, aligned with both strategic objectives and operational realities.

Engagement was carried out with a range of key stakeholders, including senior managers and leads from relevant specialist areas such as digital transformation, clinical informatics, quality improvement, data protection, and operational service delivery. Their insights were instrumental in shaping the strategic direction, scope, and governance mechanisms outlined in this policy.

In addition, a dedicated working group was convened to guide and support the policy's development. This group was comprised of representatives from Information Technology (IT), Information Governance, and Clinical Services,

ensuring a multidisciplinary contribution that addressed both clinical and non-clinical aspects of AI governance.

The working group also provided input on practical implementation considerations, risk identification, and alignment with existing organisational policies. Looking ahead, this group may continue to provide oversight or advisory input in the evaluation and governance of any AI tools proposed for adoption, ensuring that appropriate review and assurance processes are applied in future deployments.

This inclusive consultation process has contributed to a policy that is both informed and adaptable, supporting the safe, ethical, and effective integration of AI technologies within the organisation.

9. Dissemination and Implementation

9.1. Dissemination

This policy will be disseminated through a structured internal communication plan, ensuring that all relevant colleagues are informed of its purpose, scope, and application. Communication will include targeted briefings to key teams and departments, alongside organisation-wide updates via established internal channels.

The policy will be published in the designated Policies section of the organisational SharePoint site, ensuring ease of access for all staff. Where future reviews or amendments are made to the policy, these changes will be communicated clearly and promptly through appropriate channels, including staff bulletins, team updates, or direct communication to affected services, as necessary.

9.2. Competence/Training

There is currently no additional specific training requirements associated with this policy. However, relevant teams, such as those involved in digital development, information governance, or clinical safety, should ensure familiarity with the policy's content and apply it within the scope of their roles.

Where AI systems are proposed for implementation, appropriate training or awareness sessions may be arranged as part of wider project governance or change management processes. This will be determined on a case-by-case basis in alignment with the complexity of the AI solution and its intended use.

10. Monitoring Compliance

Compliance with this policy will be monitored on an ongoing basis to ensure that its principles and requirements are being effectively implemented across the organisation. This is essential for maintaining robust governance of artificial intelligence (AI) and digital technologies, as well as for meeting legal, regulatory, and ethical obligations.

Monitoring activities will be led by (NHS Provider)'s Information Governance Team and will include the following mechanisms:

- **Staff Awareness:** Evaluation of staff understanding and application of the policy through feedback mechanisms, engagement forums, and internal communications.

- Completion of Data Protection Impact Assessments (DPIAs): Review of DPIAs submitted for projects involving AI or digital technologies, ensuring that data protection, privacy, and risk mitigation measures are appropriately considered and documented.
- Spot-Checks: Periodic, unannounced reviews of selected AI-related activities or projects to confirm adherence to policy requirements.
- Internal Audits: Structured audits of AI projects or data use practices to identify any gaps in compliance, assess implementation fidelity, and recommend areas for improvement.

Findings from these monitoring activities will be reported to the appropriate governance forums, such as the Data Security and Protection Group or Digital Governance Committee, and will inform continuous improvement of the policy and associated practices.

Where non-compliance is identified, appropriate remedial actions will be taken in accordance with the organisation's policies on information governance, data protection, and risk management.

11. References

The development of this policy has been informed by a range of national guidance documents, regulatory frameworks, and best practice standards relating to artificial intelligence, data protection, clinical safety, and digital health. The following key resources are referenced:

- **A Clinical Safety Officer's Guide to Artificial Intelligence – NHS England** Provides an overview of clinical safety considerations and responsibilities for Clinical Safety Officers (CSOs) involved in AI deployment within healthcare settings.
- **Clinical Safety (CSO) Policy and Risk Management Procedure for Health IT Initiatives.** A mandated policy framework for provider organisations implementing health IT solutions. (*For organisations seeking to develop a bespoke version of this policy, please contact: krishna@innovatehealthconsulting.com*)
- **UK General Data Protection Regulation (UK GDPR)** The legal framework governing the processing of personal data in the UK, incorporating principles of lawfulness, fairness, transparency, and accountability.
- **Data Protection Impact Assessment (DPIA) Guidance – Information Commissioner's Office (ICO)** Official guidance on when and how to complete DPIAs to assess and mitigate data protection risks associated with new technologies, including AI.
- **National Cyber Security Centre (NCSC) – Guidance on AI and Data Security** Provides cybersecurity considerations specific to the use of artificial intelligence, with a focus on safeguarding data and system integrity.
- **NHS Code of Conduct for Data-Driven Health and Care Technology** Sets out the principles that health and care organisations should follow when

designing, developing, or deploying data-driven technologies, ensuring ethical, transparent, and safe use.

12. Associated Policy Documentation

This policy should be read in conjunction with the following organisational documents, which collectively form part of the governance and assurance framework for the safe, lawful, and effective use of artificial intelligence and digital technologies within (NHS Provider). These documents provide supporting guidance on data protection, cyber security, professional accountability, and organisational conduct.

- **Information Security Policy – (NHS Provider)** Sets out the organisation's approach to protecting the confidentiality, integrity, and availability of information systems and data, including measures relevant to the deployment of AI tools.
- **Disciplinary Procedure – (NHS Provider)** Outlines the formal procedures for addressing breaches of organisational policies, including provisions for disciplinary action in the event of non-compliance with this AI Governance Policy.
- **Cyber Incident Response Plan (CIRP)** Describes the protocols for responding to cyber security incidents, including detection, containment, investigation, and recovery. This plan is particularly relevant where AI tools form part of the digital infrastructure affected by a security breach.
- **Data Protection and Confidentiality Policy** Defines responsibilities for the lawful processing of personal data and upholding the common law duty of confidentiality, aligned with the UK GDPR and Caldicott principles.
- **Digital and Information Governance Policy** Provides overarching governance principles for the acquisition, development, and use of digital technologies, ensuring alignment with national standards and local procedures.
- **Records Management Policy** Details how records, including those generated or used by AI systems, should be created, stored, retained, and disposed of in accordance with legal and regulatory requirements.
- **Clinical Safety (CSO) Policy** Outlines the organisation's responsibilities for clinical risk management in the deployment of digital health technologies, ensuring patient safety in line with DCB0129 and DCB0160 standards.
- **Clinical Governance Framework / Policy** Ensures that any AI implementation aligns with the principles of high-quality care, clinical effectiveness, patient safety, and continuous improvement. This is particularly relevant where AI informs or supports clinical decision-making.
- **Corporate Governance Framework / Policy** Provides a strategic and organisational assurance context, supporting oversight of AI-related decisions at board and committee levels, especially where AI is linked to transformation, innovation, or strategic risk.

- **IT and Digital Strategy** Sets the direction for the organisation's use of digital technology, including infrastructure, architecture, innovation, and resilience—key considerations for any AI deployment.
- **Digital Transformation / Innovation Policy** Establishes the governance and process for trialling, approving, and scaling new digital technologies. Including this supports alignment between innovation processes and AI governance.
- **Clinical Risk Management Policy** Complements the Clinical Safety Policy by addressing broader clinical risks, which may include risks introduced through automation or AI-driven changes in workflow.
- **Workforce Development and Training Policy** Covers training expectations, digital literacy development, and potential AI-related competencies for staff—essential for safe adoption and future readiness.
- **Change Management / Service Redesign Policy** Guides how services are adapted or reconfigured, ensuring any AI-induced changes go through appropriate governance and impact assessment processes.
- **Procurement and Supplier Management Policy** AI solutions often involve third-party vendors. This policy helps ensure that procurement processes include due diligence for data protection, security, and ethical considerations.
- **Audit and Assurance Policy** Supports systematic internal auditing of AI deployments to verify compliance, effectiveness, and continuous improvement.
- **Risk Management Policy** Provides a structured approach to identifying, evaluating, and mitigating organisational risks—including those introduced by new technologies.
- **Business Continuity and Resilience Policy** Ensures that the impact of AI system failures is considered in service continuity planning.

(If you would like to discuss or seek support with the development of any of the above policies or related governance frameworks, please contact krishna@innovatehealthconsulting.com)

13. Appendices

Appendix A – Glossary of Terms (Section 4)

A list of key terms used in the policy (e.g., AI, algorithm, model training, DPIA, anonymisation, data minimisation, clinical safety case, etc.), with clear definitions to aid reader understanding.

Appendix B – Roles and Responsibilities (Section 5)

A summary table or matrix detailing key internal roles (e.g., CCIO, CSO, Caldicott Guardian, IG Lead, Clinical Leads, Digital Programme Manager), with corresponding responsibilities in relation to AI governance.

Appendix C – AI Project Checklist (Pre-implementation Requirements)

A practical checklist for teams proposing the use of AI, covering points such as legal basis, data protection, DPIA, clinical safety case (if applicable), ethical review, regulatory compliance, and risk assessment.

Appendix D – Data Protection Impact Assessment (DPIA) Template or Link

A standard DPIA template or a link to your organisation's approved version, tailored to include AI-specific risks such as model bias, explainability, and automated decision-making.

Appendix E – AI Use Case Assessment Form

A structured form to support the initial triage of proposed AI tools, assessing risk level, purpose, data usage, type of AI (predictive, generative, classification, etc.), and required governance pathways.

Appendix F – Policy Equality Impact Assessment Summary

A completed summary or excerpt of the Equality Impact Assessment (EIA) undertaken for this policy, showing key findings and mitigations.

Appendix G – AI Lifecycle Governance Flowchart

A visual representation of the stages from AI proposal to deployment and monitoring, aligned with governance steps such as clinical safety, procurement, DPIA, ethical review, and audit.

Appendix H – Useful Resources and Links

Hyperlinks or references to national guidance (e.g., NHS England, ICO, NCSC, NICE, HRA), best practice toolkits, and further reading relevant to AI in health and care.

Appendix I – Escalation and Review Pathways

An outline of internal escalation points for AI-related concerns (e.g., data breaches, clinical safety issues, ethical concerns), including reporting structures and governance committees.