Artificial Intelligence Policy

Policy ref no:	89
Responsible Executive Director:	Deborah El-Sayed, Chief Transformation and Digital Officer
Author and Job Title:	**Chris Borman Head of Digital ICB** **Alison Gane, Information Consultant, SCW CSU**
Date Approved:	09 February 2026
Approved by:	Shane Devlin, Chief Executive
Date of next review:	September 2027 To be co-terminus with the IT AUP

-Policy Review Checklist

	Yes/No/NA	Supporting information
Has an Equality Impact Assessment Screening been completed?	Yes	Awaiting Signoff – by the end of March 2026
Has the review taken account of latest Guidance/Legislation?	Yes	EU AI Act
Has legal advice been sought?	No	
Has HR been consulted?	No	
Have training issues been addressed?	Yes	See Section 7
Are there other HR related issues that need to be considered?	No	
Has the policy been reviewed by Staff Partnership Forum?	No	
Are there financial issues and have they been addressed?	No	
What engagement has there been with patients/members of the public in preparing this policy?	None	
Are there linked policies and procedures?	Yes	The policy is one of a suite of IG/IT related document which support the ICB's responsibilities listed in the Data Security and Protection Toolkit (DSPT)
Has the lead Executive Director approved the policy?	Yes	
Which Committees have assured the policy?	Yes	IGG, CPRG
Has an implementation plan been provided?	Yes	
How will the policy be shared with staff?	Yes	Using internal communication channels
Will an audit trail demonstrating receipt of policy by staff be required; how will this be done?	No	
Has a DPIA been considered in regard to this policy?	Yes	A DPIA is not required for policy

	Yes/No/NA	Supporting information
Have Data Protection implications have been considered?	Yes	via input from SCW CSU IG Consultant and DPO

Version	Date	Consultation
0.1	22/11/2024	Draft Version for review/comment
0.2		Draft version including comments from review meeting
0.3	28/07/2025	Updated with additional AI specifics
0.4	15/09/2025	Removed AI Strategy components to ensure only a policy

Table of contents

Use of Artificial Intelligence Policy

1 Introduction

The integration of Artificial Intelligence (AI) within Bristol, North Somerset, and South Gloucestershire (BNSSG) Integrated Care System (ICS) represents a transformative step towards enhancing patient care, management of services, operational efficiency, strategic development and clinical outcomes.

BNSSG ICB recognises the importance that AI will have in the workspace whilst via this policy and associated processes the need to ensure all use of AI is ethical, legal, responsible, proportionate, and effective. The use of AI will reduce some risks but will introduce new risks that will need appropriate consideration and control.

AI is the ability of a computer system to perform tasks that normally would be conducted by a human. This could be in the form of Machine Learning, Large Language Models, and the types of technology we are using in our everyday lives, such as ChatGPT, Siri, Alexa, and Gemini.

As we start to adopt AI into our work practices it is important that we all understand how to use these products legally, safely and appropriately to reduce any risks. It is everyone's responsibility to make sure that before using any AI products they fully understand the implications and have accepted all relevant acceptable uses policies as mentioned in the preface.

One critical distinction that is important for people to be aware of is the difference between internally deployed and governed AI systems, such as BNSSG Copilot and external products such as ChatGPT that can be accessed online. Internally deployed and governed AI systems are better for enterprises than consumer-grade systems like ChatGPT due to enhanced data privacy and security, regulatory compliance, customisation for specific business needs, and seamless integration with existing internal tools. They offer greater control over the AI's behaviour, ensuring accountability and alignment with organisational ethics and goals, which builds trust and mitigates risks such as biased outputs or security breaches.

The Acceptable Use Policy (AUP), Secure Data Environment (SDE) standards, IT policies, Artificial Intelligence (AI) governance, Information Governance (IG), and Microsoft 365 (N365) controls are interconnected components of a comprehensive digital strategy. Each exists to address distinct but overlapping risks and responsibilities: AUP defines user behaviour expectations; SDE ensures technical safeguards for secure operations; IT policies set overarching rules for technology management; AI governance addresses ethical and compliant use of intelligent systems; IG ensures data protection and regulatory compliance; and N365 provides platform-specific security and collaboration controls. Together, they create a layered approach that mitigates risk, supports legal and regulatory obligations, and enables safe innovation. The reason for multiple policies is that no single

document can cover all dimensions—technology, people, processes, and platforms—without losing clarity or enforceability.

1.1 BNSSG ICB Values

This AI policy contributes to the BNSSG ICB values by establishing a framework inclusive of each of the values:

Acting with Integrity by establishing a policy for the ethical, transparent, and responsible use of AI technologies, the policy provides guidelines to prevent harm, mitigate bias, protect data privacy, and ensure accountability, all of which are essential components of integrity in the use of AI

We Support Each Other by focusing on collaboration, training, and open communication, the policy ensures that AI tools empower employees rather than alienate them, fostering trust and psychological safety throughout BNSSG ICB.

We Embrace Diversity by establishing clear principles and practices that actively mitigate bias and promote inclusivity it ensures that AI systems are designed, developed, and deployed to expand, rather than limit, opportunities for people from all backgrounds, including those who are typically underrepresented.

We Work Better Together ensuring AI is used to enhance, not undermine, human collaboration, fosters trust, defines roles, improves communication, and empowers teams by leveraging AI for shared goals.

We Strive for Excellence by providing the strategies and ethical guardrails necessary to pursue innovation responsibly and effectively. By standardising processes and leveraging AI to improve performance, an AI policy ensures that the pursuit of excellence is both ambitious and reliable, setting a foundation for continuous improvement.

We Do the Right Thing translates abstract ethical principles into concrete, enforceable rules for using AI. It provides the necessary framework to navigate AI's complex ethical landscape and ensure the technology is used responsibly, transparently, and safely. It codifies ethical principles into actionable rules and sets standards that guide employees' moral conduct and decision-making, ensuring AI applications align with the BNSSG ICBs ethical standards. I.e. the policy must:

- Mitigate biases and avoid discriminatory outcomes.
- Protect data privacy and respect user rights.
- Prioritise safety and avoid causing harm.

2 Purpose and scope

The purpose of this policy is to establish standards relating to the procurement, development, implementation, monitoring and use of AI systems to ensure that any AI applications are used responsibly, transparently, and in alignment with our commitment to

cyber security, patient safety, data privacy, data security, ensuring its ethical, safe use of IT, legally in line with regulations, cost effective, and continuous improvement.

AI systems, such as machine learning algorithms and natural language processing, can enhance productivity and resource management as an ICB. Nevertheless, it is crucial to ensure that the use of AI technologies complies with legal standards, respects and addresses the ethnical implication and preserves the trust and confidence of our population, staff, and stakeholders.

This policy sets forth essential principles and procedures for the controlled and responsible use of AI technologies within BNSSG ICB. It covers crucial areas including data privacy, algorithm transparency, accountability, and continuous monitoring of AI systems. By adhering to this policy, we strive to cultivate a culture of responsible AI usage, ensuring that the advantages of AI are realised while mitigating potential risks.

This policy applies to any individual authorised to undertake work on behalf of the ICB, and who have been provided with access to IT equipment and systems/apps.

This includes:

- permanent or temporary staff, contractors, and agency staff.
- any third-party accessing ICB IT systems including volunteers and students.
- all those engaged in duties for the ICB, under a letter of authority/honorary contract or work experience.

The term user is used throughout this policy to encompass the above.

It is relevant to all departments and services that use AI, regardless of their scale or scope. The policy covers both internally developed AI systems and those that are publicly available. procured from external vendors for use within the system.

The overall aim of the policy is to foster a culture of controlled and responsible AI use where benefits are maximised, and risks are minimised.

3 Duties – legal framework for this policy

There is currently no specific legal framework relating to the use of AI, Current legal acts still apply and should be considered when using or designing AI.

The legal framework on which this acceptable use policy and other related information security policies are based is as follows:

- UK General Data Protection Regulation 2016/679 (UK GDPR)
- Data Protection Act (DPA) 2018
- Law Enforcement Directive (Directive (EU) 2016/680) (LED) and any applicable national Laws, implementing them as amended from time to time.

In addition, consideration will also be given to all applicable Law concerning privacy, confidentiality and the processing and sharing of personal data including:

- Human Rights Act 1998
- Health and Social Care Act 2012 as amended by the Health and Social Care (Safety and Quality) Act 2015
- Common Law Duty of Confidentiality
- Privacy and Electronic Communications (EC Directive) Regulations
- The Security of Network & Information Systems Regulation (NIS Regulations) 2018
- Computer Misuse Act 1990 and as amended by the Police and Justice Act 2006 (Computer Misuse)
- Copyright, Designs and Patents Act 1998
- Regulation of Investigatory Powers Act 2000
- Electronic Communications Act 2000
- Freedom of Information Act 2000
- Other relevant Health and Social Care Acts
- Access to Health Records Act 1990
- Fraud Act 2006
- Criminal Justice and Immigration Act 2008
- Equality Act 2010
- Terrorism Act 2006
- Malicious Communications Act 1988
- Digital Economy Act 2010 and 2017
- Counterterrorism and Security Act 2015
- Bribery Act 2010
- Economic Crime & Corporate Transparency Act 2023

Data can lawfully be used to support AI developments. The ICB's IG lead, Data Protection Officer (DPO) and Caldicott Guardian should be involved in any decision to implement or share data to develop AI technology. This will be driven by the proper completion and authorisation of a DPIA. Additionally, you can contact the NHS IG Policy Team if you require further IG support for individual projects.

NHSE has an AI knowledge repository for assistance in using AI here: AI knowledge repository - NHS England Digital to enable the responsible adoption of AI in the NHS by providing a suite of resources for all those working within healthcare to access.

NHSE also has this guidance focusing on the IG implications of using AI in health and care settings, in summary the NHSE AI IG guidance has been developed to support lawful, safe use of data for AI innovations in health and care, and focuses on patient/service-user, healthcare organisations, and IG professionals.

The guidance takes into consideration the following:

- **What Patients/Service-Users Should Know**
    - AI exists in everyday devices (e.g., voice assistants, facial recognition).
    - In healthcare, it's used for diagnosing (e.g., X-rays, mammograms), monitoring (e.g., virtual wards), and urgent decision support.
    - Your consent is implied when data is used for your direct care, with clinicians always making final decisions.
- **What Health & Care Organisations Should Know**
    - Using personal data to provide care is legitimate with implied consent.
    - Re-using data beyond individual care (e.g., for training algorithms) requires explicit processing lawful bases, such as GDPR compliance and Common Law Duty of Confidentiality.
- **What IG Professionals Should Know**
    - Ensures AI projects are supported by GDPR, DPIs, data-sharing agreements, and transparency obligations.
    - Also addresses emerging requirements around data minimisation, purpose limitation, auditability, governance oversight, and mitigation of bias, fairness, and safety risks.
- **Key Use-Cases & Considerations**
    - Examples include X-ray & CT scan analysis, virtual wards, brain scan triage, and AI-assisted ambient scribing.
    - Re-use of data beyond consent scopes triggers data protection standards, DPIAs, and transparency rules.
- **Complementing Frameworks**
    - Works alongside other NHS AI governance tools (e.g., Assurance Framework, MIAA checklist, NICE/MHRA/CQC/HRA regulations, London AI framework, NHS policy templates).

The Information Commissioners office has guidance on AI and data protection which is regularly updated here: Guidance on AI and data protection | ICO.

There are some guiding principles from the ICO, and by following these guidelines, healthcare professionals can responsibly harness the potential of generative AI, ensuring

patient safety, data privacy, and the maintenance of high-quality, human-centered care. The four principles are:

- be transparent.
- be accountable.
- consider the context you are operating in; and,
- reflect on the impact of your AI system on the individuals affected, as well as wider society.

4 Responsibilities and Accountabilities

Implementation of AI solutions will put responsibilities on all staff as well as some specific responsibilities on key roles. These are set out in the following section.

Executive Management Team

It is the role of the ICB Executive Management Team to define the ICB policy in respect of AI, considering legislative and NHS requirements. The Executive Management Team is also responsible for ensuring that sufficient resources are provided to support the requirements of this policy.

BNSSG Information Governance Group (IGG)

The Information Governance Group (IGG) oversees and provides leadership within BNSSG ICB for Information Governance (IG), ensuring that it complies with statutory responsibilities and fulfils the requirements of data protection legislation. The IGG is responsible for the review of this Policy.

Chief Executive

The ICB Chief Executive has overall responsibility for Information Governance within the organisation. They are responsible for the management of the organisation and for ensuring appropriate mechanisms are in place to support service delivery and continuity. This includes assigning the role of SIRO and Caldicott Guardian roles. The Chief Executive is responsible for approval of this Policy.

Senior Information Risk Owner (SIRO)

The Senior Information Risk Owner for the ICB is a board member with allocated lead responsibility for the organisation's information risks and provides the focus for management of information risk at executive management level. The SIRO will assign the role of DPO to a postholder. The Chief Executive must receive assurances from the SIRO that information risk is being managed suitably and successfully throughout the ICB, and for any services contracted by the organisation. The Caldicott Guardian, the Data Protection Officer, the IG Manager (SCW), and the Information Asset Owners (IAOs) provide support to the SIRO. The SCW Information Governance Manager will support the SIRO in fulfilling this role. In the absence of the SIRO there is a Deputy SIRO assigned by the SIRO. For AI this means:

- Take responsibility for the overall governance and management of the information risks associated with AI solutions, ensuring that such risks are subject to effective assessment and mitigation actions.
- Provide oversight and strategic direction to ensure the responsible use of AI solutions and ensure appropriate stakeholder engagement.

Caldicott Guardian

The Caldicott Guardian is a member of the Executive Management Team and a senior health or social care professional with responsibility for promoting clinical governance or equivalent functions and advising on confidentiality issues. The Caldicott Guardian acting as the conscience of the organisation plays a key role in ensuring that the ICB satisfies the highest practical standards for handling patient/staff identifiable information. The Caldicott Guardian serves as part of a broader Caldicott function and is supported by the Data Protection Officer and SCW Information Governance Team. For AI this means:

- Ensure use of AI solutions are aligned with the Caldicott principles.
- Ensure any use of confidential patient information is ethical and appropriate.
- Provide advice and guidance on application of Caldicott principles to the use of AI solutions.

Digital Lead

The Digital Lead is responsible for developing, communicating, managing, and implementing IT Security policies/processes daily and for managing arrangements relating to access/use involving the third-party IT supplier. For AI this means:

- Ensure the proper configuration, security and compatibility of AI solutions.
- Ensure assessment of cyber security requirements of AI solutions.
- Support the implementation, integration and maintenance of AI solutions.
- Collaborate with vendors and other stakeholders to address technical issues and provide technical support for AI systems as required.

Data Protection Officer

When required the Data Protection Officer (DPO) will report directly to the ICB Board in matters relating to data protection assurance and compliance, without prior oversight by their line manager. For AI this means:

- Ensure any use of AI solutions is compliant with data protection legislation and any related guidance from the Information Commissioner's Office.
- Provide advice and guidance on data protection related to AI solutions

- Support the development of Data Protection Impact Assessments for AI solutions where personal data is used. DPIAs will continue to be subject to the approval process documented in NHS BNSSG ICB's IG policies.
- Serve as the point of contact for data subjects and supervisory authorities regarding any data protection concerns related to AI solutions.
- Investigate and address any incidents related to use of personal data in AI solutions (jointly with other key staff as required).

Directorate Information Governance Lead (DIG)
The Directorate Information Governance Lead role is a senior member of staff who has been identified by the responsible director to represent a ICB Directorate and support oversight of Information Governance processes. This includes providing support for the requirements of the Data Protection and Security Toolkit and awareness of this policy.

Information Asset Owners (IAO)
The SIRO and Directorate Information Governance Lead is supported by Information Asset Owners (IAOs). The role of the IAO is to understand what information is held, what is added and what is removed, who has access and why in their own area. As a result, they can understand and address risks to the information assets they 'own' and to provide assurance to the SIRO that information risks within their areas of responsibilities are identified, recorded and that controls are in place to mitigate those risks. They will also investigate and act on any potential breaches of this policy.

Information Asset Administrators (IAA's)
Information Asset Administrators are required to support the IAO's and SIRO who will work with the SCW Information Governance Team to ensure staff apply the data protection legislation and Caldicott Principles and Information Governance and IT Policies within daily working practices.

SCW Information Governance Team
The SCW Information Governance Team supports the ICB and is responsible for ensuring that the Information Governance programme is implemented throughout the organisation. The team is also responsible for the completion and annual submission of the Data Security and Protection Toolkit. The Information Governance Team will support the organisation in investigating Serious IG Incidents Requiring Investigation (SIRIs), offer advice and support for the organisation to comply with this policy and support with communication through established channels including the intranet and staff briefings.

Management
All managers are responsible for promoting good information governance within their team. This includes ensuring that staff complete induction training and annual Data Security and Awareness training and acceptance of this policy and also identifying training and development needs in relation to AI for their staff.

Line Managers will co-ordinate the leavers process to ensure the return of equipment and restriction of access to systems at the end of any employment or engagement with the ICB is initiated immediately.

Clinical Safety Officer:

- Assess any clinical safety risks associated with AI solutions, with reference to any relevant clinical safety standards in place at that time (e.g. DCB0129 and DCB 0160).
- Assess whether the proposed solution requires checking and compliance with Medical and Healthcare Products Regulatory Agency (MHRA) guidance, Medical Device Regulations and guidance from the National Institute for Clinical Excellence (NICE).
- Produce and maintain any required clinical safety reports for AI solutions that support, influence or impact clinical care.
- Investigate and address any incidents related to clinical safety in AI solutions (jointly with other key staff as required.)
- Establish any safety protocols and guidelines required for the safe utilisation of AI solutions supporting clinical care.

All staff
All staff have responsibility for reading, downloading and complying with this policy and with Data Protection Legislation, organisational policies and for completing annual Data Security and Awareness training. Staff are also responsible for taking the necessary steps to maintain the security or equipment and data and for reporting breaches. For AI this means:

- Utilise AI solutions in accordance with established BNSSG ICB guidelines and processes and the following policies:
    - Acceptable Use Policy
    - Attend training and maintain any necessary compliance with mandatory training requirements
- Highlight any AI solutions that fall within the scope of requiring consultation and approval before use. Currently the only supported solution is Copilot, with solutions we already have that are upgraded to include AI the solutions DPIA must be updated and resubmitted for approval. New solutions must have a DPIA completed and approved before use.
- Make the DPO/Senior IG Consultant aware of any AI solution tender/procurement requirements.
- Are responsible for the use of any output of AI solutions that they use, checking that the output is accurate, appropriate and usable.

- Provide feedback and insights on the effectiveness, usability and impact of AI solutions
- Be mindful of the use of new and emerging technology and have a heightened awareness of potential risks despite perceived benefits.
- Report any concerns or incident related to AI solution safety or performance via existing NHS BNSSG ICB incident/near miss procedures.
- Familiarise themselves with and adhere to the ICB's Information Governance & Security policies, protocols, and guidelines.
- Report any concerns or issues related to the AI systems.
- To not use any AI, including freely available AI, including on corporate mobile devices, for ICB business that has not been formally approved by the ICB.

Clinical Safety Officer:

- Assess any clinical safety risks associated with AI solutions, with reference to any relevant clinical safety standards in place at that time (e.g. DCB0129 and DCB 0160).
- Assess whether the proposed solution requires checking and compliance with Medical and Healthcare Products Regulatory Agency (MHRA) guidance, Medical Device Regulations and guidance from the National Institute for Clinical Excellence (NICE).
- Produce and maintain any required clinical safety reports for AI solutions that support, influence or impact clinical care.
- Investigate and address any incidents related to clinical safety in AI solutions (jointly with other key staff as required.)
- Establish any safety protocols and guidelines required for the safe utilisation of AI solutions supporting clinical care.
- Familiarise themselves with and adhere to the ICB's Information Governance & Security policies, protocols, and guidelines.
- Report any concerns or issues related to the AI systems.
- To not use any AI, including freely available AI, including on corporate mobile devices, for ICB business that has not been formally approved by the ICB.

It is important to note that these roles and responsibilities may vary, collaboration and clear communication among these roles are essential for the successful and responsible use of AI.

5 Definitions/explanations of terms used

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and learn like humans. These systems can perform tasks such as problem-solving, decision-making, and language understanding by processing large amounts of data and recognising patterns.

AI Systems refer to foundational technologies that enable AI. They include the algorithms, models, and infrastructure that process data, learn from it, and make decisions. They form the backbone of AI capabilities, providing the tools and methods to develop intelligence behaviours.

AI Applications refer to the practical application of AI systems in real-world scenarios. AI applications use capabilities of AI systems to perform specific tasks or solve problems. These are the 'end product' that users interact with, leveraging the power of AI systems.

Generative Artificial Intelligence is a subset of AI that focuses on creating new content, such as text, images, or music, based on the data it has been trained on. It uses models like Generative Adversarial Networks (GANs) or transformer-based models to generate outputs that are often indistinguishable from those created by humans.

Machine Learning (ML) is a subset of AI that focuses on developing algorithms and statistical models that enable computers to learn from and make predictions or decisions based on data. Unlike traditional programming, where explicit instructions are given for every task, ML systems improve their performance over time by identifying patterns and relationships within large datasets. This capability allows ML to be applied in various fields, such as image and speech recognition, recommendation systems, fraud detection, and predictive analytics, making it a powerful tool for solving complex problems and enhancing decision-making processes.

Natural Language Processing (NLP) is a subfield of computer science and artificial intelligence (AI) that focuses on enabling computers to understand, interpret, and generate human language. By combining computational linguistics with machine learning and deep learning techniques, NLP allows machines to process and analyse large amounts of natural language data. This technology is used in various applications, such as speech recognition, language translation, sentiment analysis, and chatbots, making it possible for computers to interact with humans in a more natural and intuitive way.

Robotic Process Automation (RPA) is a technology that uses software robots, or "bots," to automate repetitive, rule-based tasks traditionally performed by humans. These tasks can include data entry, form filling, and transaction processing across various applications and systems. By mimicking human interactions with digital systems, RPA enhances efficiency, accuracy, and productivity, allowing employees to focus on more complex and strategic activities. RPA is widely used in industries such as finance, healthcare, and customer service to streamline operations, reduce costs, and improve overall service quality.

AI Bots are software applications that use artificial intelligence (AI) and machine learning (ML) to automate tasks and interact with users, often in a conversational manner. They are designed to understand user input, process information, and provide responses or complete actions, sometimes mimicking human-like conversations.

6 AI principles

BNSSG are committed to the responsible and ethical use of artificial Intelligence (AI) systems.

Please follow the AI Policy for use of AI and be aware of the following specific standards.

- AI software (e.g., Co-pilot) must not be used to process personal, special category or business confidential information.
- Understand the privacy policies and data handling practices of any third-party AI tools you are permitted to use.
- Ensure caution when sharing AI-generated content. Clearly indicate when content has been generated by AI if there is a risk of misinterpretation.
- Be responsible for the consequences of disseminating AI-generated content. If you share or publish content created by AI (like text, images, or data), you are accountable for how that content is used, interpreted, and the impact it may have.
- Do not present AI-generated work as your own unless explicitly permitted for specific educational or research purposes with proper attribution.
- Do not intentionally attempt to overload, disrupt, or damage AI systems.
- Report any suspected vulnerabilities or malfunctions of AI systems to the designated IT Service desk.
- Generating Unreliable Information: Relying on AI-generated content without verifying its accuracy and completeness. AI is a tool to assist, not replace, human judgment and fact-checking.
- Circumventing Governance Policies: Using AI in ways that bypasses or undermines the ICB's information governance policies, including data security, privacy, and access control.
- Submitting personal Data in Prompts: Directly inputting personal identifiable data into AI prompts is strictly prohibited. Anonymised or aggregated data may be used where appropriate and compliant with information governance policies.
- Be mindful of the use of new and emerging technology and have a heightened awareness of potential risks despite perceived benefits.
- If in doubt, refer to IT Acceptable Use Policy

7 Details of the policy

Assessment of AI solutions

Staff with key responsibilities will collaborate to utilise and/or develop appropriate assessment tools for the procurement, development and use of AI solutions. Where possible these will be linked in with other existing assessments such as Data Protection Impact Assessments (DPIAs), digital clinical safety assessments and cyber security assessments (linked to Digital Technology Assessment Criteria – DTAC).

Any assessment of an AI solution will also look at ethics and fairness, considering potential for bias in an AI model and transparency and explainability so that the users can understand the functioning of the product. The assessment will also ensure there is a solid use case for the proposed AI solution.

Any AI solution will also adhere to the AI Safety Principles set out in Appendix D to this policy

Acceptable use of AI solutions without consultation

AI will have many uses ranging in scale from the simple 'write me an interview question' through to a system for clinical decision support. It is not possible or necessary for BNSSG ICB to assess every single use of AI. All staff will be informed of the sorts of uses that are acceptable without engaging formal consultation and approval. Staff will be permitted to use AI without consultation where:

- The AI solution does not require any confidential or sensitive data (including personal data or commercially sensitive data) to be input or uploaded or personal data of individuals other than the user.
- The AI solution generates content such as text for a document or images but does not provide any form of decision support.
- Any AI produced output must be thoroughly reviewed by the user before being relied on.

Consultation on potential use of an AI solution:

Any team, service or staff member considering the use of any AI solution (procured, developed or free) must undertake consultation with the following key roles:

- Data Protection Officer/IG Consultant to establish if any personal data is processed.
- Clinical Safety Officer to assess any clinical use aspects of the proposed product.
- Digital Team regarding implementation of any software.
- Cyber security lead regarding implementation of any software or the security of any externally provided solutions.
- Human Resources/People lead regarding any potential impact on workforce.

Whilst not all AI solutions will need detailed assessment from each of the above, it is important that each area is given due consideration and documented, even if the assessment is confirmed as no required input.

The above roles, when consulted will determine if any specific engagement is needed with the Senior Information Risk Owner and/or Caldicott Guardian.

In consultation with any of the above roles, consideration will be given to any wider consultation needs with the use of the AI solution in question this can include:

- consultation with staff and the BNSSG ICB Staff Partnership Forum where appropriate
- representatives of the public.
- partner organisations potentially impacted using AI solutions by BNSSG ICB.

Purpose and legal basis:

It is key that the purpose, intended benefits and basis of using any AI solution is clearly defined and agreed prior to any operational use.

AI proposals that potentially utilise personal data will be screened as part of BNSSG ICB's Data Protection Impact Assessment process. Where a full Data Protection Impact Assessment is undertaken, the DPIA will identify the appropriate legal basis for the use of any personal data and risks around the use of such data.

Procuring AI solutions:

Procurement of any AI solution should not commence until any consultation with key stakeholders (listed above) has concluded or at least reached a position where the stakeholder is happy for procurement to commence. This is particularly important where a stakeholder will have material input into the procurement activities related to the required solution, such as specifications etc.

Developing AI solutions:

Where BNSSG ICB is involved in work either solely or in collaboration with others to develop an AI solution, then staff leading the work must ensure that it takes into consideration the NHS AI and Digital Regulations Service: Home - AI and Digital Regulations Service for health and social care (innovation.nhs.uk). It will also take into consideration any other standards or regulations relevant at the time to either the development in general or the specific context of the solution (e.g. Digital Clinical Safey standards/Medical device regulations as appropriate).

Any AI product developed will be subject to the same requirements of this policy as any procured or freely available solution, e.g. DPIA and consultation.

This policy does not cover any other aspects related to the development of AI solutions such as intellectual property, collaboration agreements or contracts with collaboration partners.

Freely available AI solutions:

There are many freely available AI solutions and many also do not require any form of local implementation or installation i.e. AI Bots. Prior to use these must be assessed as set out above and if approved, these should be used with caution, particularly in any form of clinical context and the staff member using them is responsible for ensuring they review the outcomes/outputs of such products. Staff need to be aware of the following specific requirements:

- They must not be used to make decisions without appropriately skilled and knowledgeable staff reviewing the output. The staff remain responsible for the output created.
- Where they are used to support the authoring of any document, the use of the tool must be noted and clearly stated within the document. With the risk of tools producing inaccurate, biased or false information, any output must be reviewed by an appropriately skilled and knowledgeable member of staff.
- Personal data or organisationally sensitive data must not be used on such tools without consultation with the roles identified above.
- AI Bots depending on the meeting, it is the responsibility of the chair or meeting organiser to allow or disallow these AI Bots. Generally, these bots only record notes in a similar way to copilot, and most are safe, however, the decision should be made at the beginning of a meeting if these bots may be used. For a bot to be approved for use the attendee must also be in the meeting, if the attendee does not join the meeting the bot must be removed from the meeting.

Developing large language models (LLMs)

Developing large language models (LLMs) raises significant concerns spanning ethical, technical, and societal domains, primarily related to bias and fairness, data privacy, misinformation (hallucinations), security, and accountability.

Ethical and Societal Concerns

- **Bias and Fairness**: LLMs learn from vast datasets that may reflect prejudices in society or stereotypes. Consequently, the models may inherit and amplify these biases, resulting in discriminatory or unfair outputs, which can have serious implications in critical areas like healthcare or hiring.
- **Misinformation and Hallucinations**: LLMs can generate content that is convincing but factually incorrect or nonsensical (known as "hallucinations"). This poses a risk for the large-scale creation and spread of disinformation, which can damage reputations or influence public opinion on critical events like elections.
- **Accountability and Human Agency**: The hidden nature of complex neural networks makes it difficult to understand how they arrive at specific decisions, complicating accountability when errors occur.

- **Intellectual Property and Copyright**: The use of data scraped from the internet without clear consent or compensation may lead to lawsuits regarding intellectual property and copyright infringement.
- **Job Displacement**: There are concerns that the increasing automation capabilities of LLMs could lead to significant job displacement in various white-collar professions, with potential short-term catastrophic economic effects.

Each of the above points raise questions about human oversight and whether LLMs should be used in high-stakes decision-making processes, and when building LLMs there must be documented evidence of robust validation mechanisms covering the above attributes.

Technical and Security Concerns

- **Data Privacy and Security:** LLMs are trained on extensive datasets that may inadvertently include sensitive personal information. This creates risks of privacy breaches, where models might memorise and reproduce confidential data. Security vulnerabilities like prompt injection and data poisoning also pose threats.
- **Computational and Environmental Costs:** Training and running LLMs may require massive computing power and energy consumption, raising concerns about high operational costs and significant environmental impact.
- **Reliability and Limitations:** Despite impressive capabilities, LLMs struggle with complex reasoning, staying up-to-date with real-time information (without supplementary systems like RAG), and understanding nuance like sarcasm or idioms.
- **Malicious Use:** Malicious actors can exploit LLMs to generate malware, craft highly convincing phishing attacks, or create biological weapons information, necessitating strong safety measures and regulatory frameworks.

To address these points, when developing LLMs we must ensure we document and self-regulate our processes, promote transparency in the data used, implement robust data protection measures, and ensure human oversight in critical applications, there must be a Data Privacy Impact Assessment for each LLM created and must be updated for each new use of the LLM.

Vibe Coding

Vibe coding is the practice of using natural language prompts with generative AI to produce software with minimal human review, this provides several concerns regarding security flaws, technical debt, maintainability issues, and skill degradation among developers. Concerns include:

- **Security Vulnerabilities:** AI-generated code may contain security flaws, such as injection attacks, hardcoded API keys, and improper authentication, because models are trained on public, often insecure, code and lack a deep, adversarial understanding of security.

- **Accumulation of Technical Debt:** The focus on speed over structure often results in verbose, poorly organised, and undocumented code that is difficult to maintain, optimise, and scale over time.
- **Debugging Difficulties:** Developers who did not write the original code struggle to understand and debug AI-generated code, leading to time-consuming troubleshooting or complete project rewrites.
- **Skill Degradation and Over-reliance:** Over-reliance on AI tools can prevent junior developers from building foundational programming knowledge, debugging abilities, and architectural understanding, potentially making them less valuable in the job market.
- **Inconsistent Output and Lack of Control:** Asking the same LLM the same prompt can yield different results, and the AI may not always follow specific instructions, making consistent code quality difficult to achieve.
- **Supply Chain Risks:** AI can "hallucinate" non-existent software packages or libraries; malicious actors can then create these packages to compromise applications that automatically install them, introducing malware into the supply chain.
- **Compliance and Legal Challenges:** The opaque nature of AI code generation and potential intellectual property ambiguities regarding training data make it difficult to ensure compliance with regulations like GDPR or industry standards such as OWASP guidelines.

To manage concerns on Vibe Coding AI should be treated as a copilot and not an autopilot. This requires significant human oversight, rigorous testing, and code review for any production-level application. Any vibe coding must include the following areas which must be documented:

- Mandatory Code Reviews
- Comprehensive Testing
- Integrate DevSecOps (Development, Security, & Operations) solutions and security scanning tools to automatically scan for common vulnerabilities.
- Validate Inputs and Outputs: Ensure every input is validated and sanitised to prevent injection attacks and other security breaches, following guidelines from resources like the OWASP Cheat Sheet Series (https://cheatsheetseries.owasp.org/).
- Define Clear Requirements and Architecture
- Maintain Control and Accountability, ensuring version control and documentation
- Establish human-led security review gates as a mandatory part of the development lifecycle before any code moves to production.
- Promote and adhere to a secure by design principle.

Using AI for Research

- Health Research Authority (HRA) approval is required for research studies that take place in the NHS in England. The 'HRA AI and Digital Regulations Service' can provide guidance for NHS AI adopters, and digital health innovators. Review by an NHS Research Ethics Committee (REC) is required, as well as an assessment of regulatory compliance and related matters undertaken by dedicated HRA staff. If you are planning to develop an AI research programme within the NHS, the Research and Development Support Services team within the Innovation, Research and Improvement System (IRIS) will be able to provide advice and guidance on how to apply for research ethics and approvals via the Health Research Authority.

Routes to approval for use:

With the appropriate engagement of key roles and other stakeholders any proposal for the use of AI, whether that is a procured product, a locally developed product or use of freely available tool will go through the following governance approvals as indicated:

- Information Governance Group (all AI proposals to check impact/use of personal data)
- Where the product is to support the requirements of a customer, the above approvals will be noted for the customer, and the approval will also be subject to any approval process required by the customer(s). This will be through the Digital and IG teams.

Where there is an urgent need to approve an AI solution (generally a freely available tool), then a fast-track process will be set up, consisting of review and agreement by key leads in IT/Cyber Security, Data Protection Officer, Caldicott Guardian (if applicable) and the Senior Information Risk Owner.

A process flow will be documented and used along with an organisational register of AI products in use.

8 Monitoring compliance and effectiveness

Adherence to the requirements of this policy will be monitored in the following ways:

- Initial AI proposal assessment and screening of AI proposals for Data Protection Impact Assessments and completion of full DPIAs where required. In line with existing IG policies and conducted via existing procedures.
- Incident & near miss monitoring for any mention or potential AI issues or threats.

The Information Governance Group will be alerted to any significant issues identified from any monitoring activities and have responsibility for identifying and undertaking any remedial or improvement actions.

9 Review

The policy will be reviewed at least annually. AI is an area of developing legislation and regulation, so in addition to annual review, any significant legal, regulatory or sector impacts will need to be taken into consideration and policy review undertaken if required in advance of any annual review. Decision on any interim review will be taken by the Information Governance Group overseeing the policy

10 Training requirements

Before using AI products, all staff are required to be current with mandatory training relating to the use of data (Data Security Awareness, Data Security Awareness Handbook & Policy). In addition, users are required to complete any specific training available relating to the use of AI products or systems. This will vary depending on the system or service in use. If unsure of training requirements, users are to seek guidance from the system implementation partner.

This policy has the following training requirements to support complying with the policy statements:

- All staff: As a minimum all staff will be made aware of the policy and key responsibilities via all staff communication channels and the provision of guidance via these routes.
- Staff with additional knowledge needs in respect of AI. Staff in areas such as Digital, and Information Governance will be expected to discuss their knowledge and skill requirements regarding AI as part of their appraisal and development reviews. BNSSG ICB will endeavour to support staff with any development requirements related to AI.
- Roles identified as specific responsibilities in this policy will be specifically required to identify training and development needs to fulfil these responsibilities. Any appropriate support to meet these will be provided.

11 Equality Impact Assessment

EHIA included as Appendix C

12 Implementation and Monitoring Compliance and Effectiveness

This is included as appendix B

13 Countering Fraud, Bribery and Corruption

The ICB is committed to reducing and preventing fraud, bribery and corruption in the NHS and ensuring that funds stolen by these means are put back into patient care. During the development of this policy document, we have considered how fraud, bribery or corruption may

occur in this area. We have ensured that our processes will assist in preventing, detecting, and deterring fraud, bribery and corruption and considered what our responses to allegation of incidents of any such acts would be.

In the event that fraud, bribery or corruption is reasonably suspected, and in accordance with the Local Counter Fraud, Bribery and Corruption Policy, the relevant Team will refer the matter to the ICB's Local Counter Fraud Specialist for investigation and reserve the right to prosecute where fraud, bribery or corruption is suspected to have taken place. In cases involving any type of loss (financial or other), the ICB will take action to recover those losses by working with law enforcement agencies and investigators in both criminal and/or civil courts.

14 References, acknowledgements and associated documents

BNSSG ICB Data Security Awareness and Information Governance – Staff Handbook

BNSSG ICB Information Governance Management Framework and Strategy

Information Commissioner's Office: AI and Data Protection Risk Toolkit - AI and data protection risk toolkit | ICO

Humber and North Yorkshire Health Care ICB: Artificial Intelligence (AI) Governance Policy (October 2023)

Somerset NHS Foundation Trust: Approval & Use of Artificial Intelligence (AI) Policy (November 2024)

BNSSG ICB AUP Policy

15 Appendices

Appendix A - Equality Health Impact Assessment

See attached EHIA.

Equality and Health Inequalities Impact Assessment - AI

Appendix B - Implementation Plan

Target Group	Implementation or Training objective	Method	Lead	Target start date	Target End date	Resources Required
All	Implementation	Comms via HIGNFY, The Hub, and The Voice	CB	11/25	N/A	Comms Team
All	Training	Updated AUP to include AI	CB	04/26	06/26	IG
All	Training	As advised in the policy				

Appendix C - Equality Health Impact Assessment

Section 1: What is it about?
1 Describe the proposal or policy and the outcomes and benefits you are hoping to achieve. The policy has been developed to ensure sufficient oversight and governance of the use of AI solutions in the services provided to the wider Health & Care Sector by BNSSG ICB. This should include both improvements in organisational efficiency and effectiveness of clinical care. Specific benefits will depend on the AI solutions in the context that they are used. The power of the technology to make improvements is very significant, however any such development needs to consider and address any risk factors both to the data the solution uses and the outputs that will affect individuals.
2 Who is it for? The Policy is applicable to all BNSSG ICB staff (including interim/off payroll workers and bank staff).
3 How will the proposal or policy meet the equality duties? AI has the potential for massive benefits, but these can only be realised if the use of solutions is appropriately governed. Poor governance of the use of such solutions may have a detrimental impact on equality aspects. AI solutions have a risk of bias that can occur due to the data used in AI modelling and this could be linked to protected characteristics. A poorly designed and implemented solution could discriminate between individuals in an unfair manner. On that basis the policy has been designed to ensure as far as reasonably possible that the governance controls detailed in the policy will reduce and mitigate such bias and safety risks to a minimum that can then be assessed to meet legal and regulatory standards in the areas of data protection, clinical safety, cyber security and as required medical device regulation. The accessibility features that are available through the technology i.e. screen reader compatibility for visually impaired or for Mental Health-Friendly Interaction by providing calm, clear language, avoiding jargon, and offers step-by-step guidance to reduce

cognitive load. By using these features, it will help reduce barriers for those with specific requirements

4 What are the barriers to meeting this potential?

The barriers are that this will require significant raising of awareness across all BNSSG ICB staff and sufficient time and expertise for key staff to embed the necessary governance that AI requires into existing governance and programme frameworks and purposes. With numerous AI solutions being freely available there is a risk that staff start to use such solutions with all good intents, but without sufficient control.

Section 2: Who is using it?
Consider all equality groups

Describe the current or proposed beneficiaries and include an equality profile if possible.

The beneficiaries are:

- BNSSG ICB Staff who will be able to use AI solutions with confidence, staff who are less confident can request training, as per the roles and responsibilities this is completed by the line manager.
- BNSSG ICB customers who will benefit from the use of AI solutions
- Patients either directly or indirectly from the improvements/enhancements that properly governed and managed AI can bring

5 How have you/can you involve your patients or service users in developing the proposal or policy?

Patients/service users have not been involved in developing the policy. However, the policy notes that where required to implement an AI solution consideration will be given to consultation with public representatives. This may be a specific activity or part of a Data Protection Impact Assessment. It is also possible BNSSG ICB will utilise AI solutions developed elsewhere that have in their development included such consultation. Data

Protection legislation also requires a degree of transparency about the use of data and any AI use will also be subject to such transparency.
6 Who is missing? Do you need to fill any gaps in your data? No gaps noted.

Section 3: Impact **Consider how it affects different dimensions of equality and equality groups, using the information from steps 1 and 2 above**
a) Does (or could) the proposal or policy create an adverse impact for some groups or individuals? Is it clear what this is? The policy is required to ensure as far as possible that there are no adverse impacts on any groups or individuals. Not having the policy and sufficient governance processes may create such an impact.
b) What can be done to change this impact? If it can't be changed, how can this impact be mitigated or justified? As above, sufficient governance through the policy and related processes is the mitigation against such impacts.
c) Does (or could) the proposal or policy create a benefit for a particular group? Is it clear what this is? Can you maximise the benefits for other disadvantaged groups? The policy and AI solutions could create such a benefit if that is what they are designed to do, e.g. algorithms to support identification of health inequalities in ethnic communities. Such benefit could only be maximised by increased use of such AI solutions in a controlled and governed manner.

d) Is further consultation needed? How will the assumptions made in this analysis be tested? No plans for further consultation currently. Adherence to the governance processes that will flow from the policy will test the proposed AI solutions to ensure that no adverse impacts on equality arise. This will be overseen by groups such as the Information Governance Steering Group and the Corporate Governance Assurance Group.

Section 4: So what? (The outcome of this EIA) **Link to the business planning process**
What changes have you made during this EIA? No changes
7 What will you do now and what will be included in future planning? As AI solutions are developed the governance model will be checked and reviewed for any necessary adaptions.
8 When will this EIA be reviewed? Annually subject to revisions of the policy.
9 How will success be measured? Success will be measured by the effective implementation of AI solutions into BNSSG ICB services. This can be measured by the number of projects successfully completing

and realising benefits with an AI solution as part of the project. Another measure will be the absence of incidents (or acceptable level of) related to AI solutions.

Person(s) completing EIA
Chris Borman
Date EIA completed; 15/09/25
Review Date; Annually subject to revisions of the policy

Appendix D - Safety principles

The following principles demonstrate how not to use AI Safely.

1. Misuse AI for cognitive behavioural manipulation

AI should not be used to manipulate the behaviour of people, particularly vulnerable groups. This includes:

- Avoiding the deployment of automated chatbots to provide advice or information to vulnerable individuals without appropriate human oversight.

We will also ensure that AI applications are designed to support rather than exploit or manipulate users.

2. Misuse AI in a way that invades privacy

AI applications must respect privacy and confidentiality. This includes:

- It is not permitted to use Artificial Intelligence to track the physical location of any patient, employee, or member of the public without their informed consent.
- The use of software that trains on a person's likeness, such as identifying a person's identity from image or video recognition software is not permitted.

Any individuals who have concerns about surveillance can raise an objection under their data subject rights and this will be considered in line with requirements of current data protection legislation.

3. Misuse AI in a way that contradicts safety or technology recommendations

AI applications must comply with established safety and technology guidelines. This includes:

- Avoiding the use of AI in ways that conflict with established guidelines; including for example, terms of service, agreed contracts or recommendations, SCW policies and guidance.

We must also ensure that any deviations from guidelines are approached through a well-documented research methodology.

4. Misuse AI in a way that negatively impacts the welfare of individuals

AI should not negatively impact the welfare of individuals. This includes:

- Prohibiting the use of generative AI to create multimedia or content of individuals without their explicit permission.

We will ensure that AI applications do not cause harm, distress, or disadvantage to individuals.

5. Allow AI to become part of critical infrastructure without appropriate fail-safes

AI should not be integrated into critical infrastructure without appropriate fail-safes.

Fail-safes should:

- Be implemented to mitigate failures in critical systems where AI is deployed.
- Be regularly tested and validated to ensure appropriate availability, redundancy, reliability and safety.

We will ensure that Business Continuity Plans effectively consider AI processes, addressing how AI systems will be managed and restored in case of disruptions or failures.

6. Misuse AI to edit or delete legally binding documents or records without significant safeguards in place

AI systems must adhere to data protection laws and respect legally binding documents. This includes:

- Ensuring AI does not edit or delete individuals notes, diagnostics, official documents, or electronic records without human oversight, authorisation and traceability.
- Maintaining integrity and security of sensitive data in compliance with relevant legal frameworks.

7. Use AI to decide the right of access to services

AI should not determine access to services or cause discrimination. This includes:

- Prohibiting the use of AI to decide the right of access to services provided by the organisation.

8. Use AI to engage with individuals without human oversight

AI decisions and communications must involve human oversight. This includes:

- Avoiding the use of AI to engage with individuals or send communications of a personal, sensitive or confidential nature without human review.

9. Use AI to relay critical information without trained professional oversight

Ensuring that AI recommendations on care, legal advice, or critical information are overseen by trained professionals.

10. Scanning job applicant cohorts

Where protected characteristics have been used in the training set or the AI application is likely to cause discrimination.

- Ensuring that AI applications do not use protected characteristics in ways that lead to discrimination, especially in job applicant screenings.

11. Purely automated decision making

All AI decision making processes should have human oversight, meaning that suitably qualified human involvement and accountability is required.

12. Replicating a person's likeness

There is a high cybersecurity risk associated with creating virtual replicas of staff (either by replicating voice, image, or video), especially those in positions of seniority or budget holders. Staff are not permitted to utilise AI to create virtual replicas.

13. Obscurity

All members signing up to this agreement agree to the seven principles of Public Life (also known as the Nolan Principles - selflessness, integrity, objectivity, accountability, openness, honesty and leadership) and also accept a responsibility as part of openness, to ensure that their staff and the populations that they serve are well informed about how AI is being used in public service and have clear communication channels if they wish to raise concerns.