

The NHS Confederation

Charity number 1090329

Company number 04358614

AI Ethics and Compliance Policy

Owner of policy	Director of Finance and IT
	Director of People and Governance
Date of policy review	November 2024
Date of next review	November 2025
Approved by CEO	November 2024
Agreed with Joint Negotiating Committee	[N/A]
Approved by <i>[insert committee]</i>	[N/A]
Approved by Board of Trustees	December 2024

Version Control Table

Version	Date	Revisions
1	November 2024	New policy

Contents

Contents	3
Introduction.....	4
Purpose of this policy.....	4
Scope	4
Policy statement.....	5
Introducing new AI/ AI development	9
Record Keeping	9
Training and technical support.....	10
Reporting Concerns.....	10
Failure to Comply.....	10
Monitoring and Review	10
Responsibilities	11
Guidance notes	12
What is AI?	12
What are the risks that could affect me?.....	12
Authorised AI within the NHS Confederation	13
Requesting use of other AI	13
AI is constantly evolving	13
Annex 1: AI approval Form.....	15
Annex 2 Risk Assessment Framework.....	18
Equality Impact Assessment - Policies.....	19

Introduction

Purpose of this policy

This policy outlines our commitment to ethical, accountable, fair and compliant use of Artificial Intelligence (AI) within the NHS Confederation. It ensures that our use of AI systems is compliant with all relevant legislation and upholds our values and ethical principles.

Scope

This policy applies to any use by the workforce of AI for business purposes. More specifically, this policy applies to all AI systems commissioned, developed, or used by the organisation, including but not limited to machine learning models, automated decision-making processes, and AI-driven data analysis tools. Where AI is being used for business purposes by a member of staff of NHS Confederation, this policy applies to the use of AI both during and outside normal hours of work and whether or not use of AI is on an individual's own device or one of NHS Confederation's devices, and whether at home, in the office or from a remote working location.

This policy applies to everyone working for or with the NHS Confederation. It applies to:

- all staff, including directors and employees (whether permanent, fixed-term or temporary), seconded staff, agency workers and volunteers.
- consultants and contractors
- trustees and committee members.

Any employing or contracting manager must ensure that all temporary staff, consultants, and contractors are aware of this policy.

By the NHS Confederation we mean the NHS Confederation charity, any subsidiary companies and any hosted networked organisation.

The NHS Confederation has designated the **Director of People and Governance and the Director of Finance and IT** as the individuals who are responsible for ensuring that the NHS Confederation implements this policy.

This policy should be read in conjunction with the organisation's Values, Ethical Principles, Data Protection, Diversity, Equality and Inclusion, and IT Policies.

Policy statement

At NHS Confederation, we recognise the immense potential, and impact, of Artificial Intelligence (AI) on our operations. NHS Confederation is committed to ensuring that the development and deployment of AI within our organisation aligns with our strategy, our values, ethical principles, and compliance requirements under the UK General Data Protection Regulation (UK GDPR) and other relevant legislation.

Our core principles for the use of AI within the NHS Confederation are:

1. Inclusive, fair and non-discriminatory

- NHS Confederation is committed to ensuring that AI systems (including the datasets they have been trained on, or draw from), and their use, do not discriminate based on factors such as race, gender, age, or any other protected characteristic as defined by section 4 of the Equality Act 2010. AI will not be used to unjustly differentiate between individuals or groups.
- All users of AI should make sure they are aware of the potential for AI to be discriminatory and, where using AI, will actively build in appropriate fairness tests to avoid unfair bias and discrimination.
- Any new AI design or project involving AI will be designed with inclusivity in mind that will actively consider a wider range of users and proactive steps to ensure fair and equitable outcomes. Where appropriate, we will undertake an equality impact assessment in order to assure ourselves that there is no unfair outcome (see appendix 1).

2. Avoids harm to individuals

- NHS Confederation is keen to benefit from the use of AI in furthering its work and charitable aims. It, however, recognises that with those benefits comes potential risks including possible harm to individuals. Users must not use AI to create harmful, cruel or malicious content. This includes but is not limited to:
 - Generating demeaning, dehumanising, or otherwise harmful representations of people or their environments, cultures, religions, etc.
 - Intentionally promoting or propagating discriminatory content or harmful stereotypes.
 - Impersonating individuals without their consent.

- Sexual content
 - Mis- and disinformation
 - Representations of egregious violence and gore
- Whilst recognising that new EU laws do not apply in the UK, the NHS Confederation shall take account of the principles in the EU AI Act and shall adhere to the Act's list of prohibited uses of AI, not least social scoring¹; emotion recognition²; subliminal techniques, manipulation and deception and biometric categorisation (e.g., emotion recognition or inferring sensitive traits such as religion or sexual orientation)

3. Transparent and Accountable

- The NHS Confederation is committed to full transparency in its use of AI systems. Users must inform all participants in meetings, consultations, or decisions when AI tools are employed, specifying the purpose and data handling practices.
- All AI systems adopted must provide explainable and understandable outputs to ensure stakeholders can verify and comprehend AI-generated outcomes. Participants must be given the option to opt-out of their data being processed in AI.
- Where AI is being used to generate new content for external publication the user will ensure appropriate consideration of copyright and Intellectual property. This must include highlighting sources used and stating that the content was generated with AI support or obtaining appropriate advice on these issues as necessary.
- NHS Confederation recognises the need for human oversight and intervention to ensure control, accountability, and ethical decision-making. As such, we will always adopt a human-in-the-loop approach where all use of AI must include human intervention. Users must critically review AI

¹ Social scoring, also known as “social credit”, are systems that use AI to rank people based on their social and economic activity. The score is calculated using a variety of factors, such as their online activity, their spending habits, and their social interactions. It is found to be highly discriminative

² Emotion recognition in AI identifies specific human emotions (like happiness or anger) from biometric data, facial expressions, voice tones, or physiological signals, while sentiment recognition determines the overall positive, negative, or neutral attitude expressed in text alone. Sentiment recognition is therefore acceptable under the EU AI Act

outputs, address potential biases, and intervene when decisions or actions may negatively impact individuals or groups.

- Any adoption of new AI, or use of existing AI that uses personal data or influences decisions that impact on people or protected (or minority) groups of people must undergo an ethical screening process. Staff will be required to complete an ethical screening form (see Annex 1) and reviewed by the Director of People & Governance and Director of Finance and IT. They may also be required to comply with the Algorithmic Transparency Recording Standard if undertaking such work on behalf of a public body and must seek clarification on this at the time.

4. Data and Respect for Privacy

- Data is a critical component of AI development. NHS Confederation is committed to responsible data management practices to ensure ethical and legal use of data in AI development. All AI systems must adhere to robust data governance standards, ensuring training datasets are accurate, representative, and free of bias. Personal data processed by AI systems must comply with GDPR and NHS Confederation policies, with appropriate data minimisation and security safeguards.
- Where personal data will be processed as part of the use of AI, staff must undertake a data protection impact assessment prior to the implementation of that AI, and in doing so identify a lawful basis for processing that data and ensure that processing is fair and transparent as well as uphold all data protection principles, not least data minimisation, accuracy and security
- The use of AI systems for recording, transcription, or decision-making must be contingent on obtaining explicit informed consent from all participants. Consent must include clear explanations of the AI tool's purpose, data processing, and storage practices, with provisions for individuals to opt out at any time.
- The use of personal data is not permitted in any AI that is external to NHS Confederation systems, unless NHS Confederation has entered into an appropriate contractual arrangement (including a data processing agreement within the meaning given in Article 28 of the UK GDPR) with

any relevant third party providing external systems upon which AI is used. In doing staff must use the standard contract template which sets out appropriate measures for compliance.

5. Compliant and upholds integrity

- In all AI use, staff must be clear on data ownership and must ensure that intellectual property rights are clearly understood and communicated.
- Our AI practices must comply with all applicable laws, including the UK GDPR and the Data Protection Act 2018 and any relevant guidance issued by the Information Commissioner's Office³ as may be updated from time to time.
- Staff will not use AI in the sharing of copyrighted or licensed material in violation of its terms of use. Similarly, users will not use AI to share content that is an alteration of copyrighted or licensed material in violation of its terms of use.
- The organisation will stay appropriately informed about evolving legal standards and good practice related to AI and adjust our practices accordingly.

6. Safe and secure

- All AI systems must be secure against unauthorised access and malicious use. AI adopted for use by NHS Confederation shall only be those that have undergone adequate security reviews and maintain all data within NHS Confederation systems, unless NHS Confederation has entered into an appropriate contractual arrangement (including a data processing agreement within the meaning given in Article 28 of the UK GDPR) with any relevant third party providing external systems upon which AI is used, and through which has undergone an AI risk assessment.
- The organisation's approved AI product is Microsoft Co-Pilot. We will maintain a list of other systems which are authorised for use (see 'Record Keeping' below).
- The organisation will not adopt any high-risk AI systems (as recognised by the EU AI Act).

³ [What are the accountability and governance implications of AI? | ICO](#)

- Staff must not download, create, or access any other AI without adhering to this policy and obtaining the appropriate approval as outlined below.

Introducing new AI/ AI development

The NHS Confederation will establish a comprehensive ethical screening process for all AI systems. This process will evaluate potential impacts on equality, privacy, and data security, ensuring compliance with NHS Confederation and EU AI Act standards. High-risk systems, such as those impacting individuals' rights or processing sensitive data, will not be permitted.

Should any part of the organisation want to introduce a new AI into their operation, they must submit a request form to Director of Finance and IT and also refer to the Software Policy. This request shall detail:

- The purpose of the AI to be adopted
- How the AI works (explainable AI)
- How bias will be mitigated
- IT security arrangements
- Privacy and data protection arrangements, including:
 - Preparation of a data protection impact assessment (as necessary);
 - Identification of whether NHS Confederation (and any other entity involved in the AI process) is a controller or processor for specific processing operations in the development and deployment of AI systems and the resulting impact for NHS Confederation's responsibilities;
 - How to assess the risks to the rights and freedoms of individuals and how these should be addressed when designing and deciding to use an AI system; and
 - How to justify, document and demonstrate the approach including the decision to use AI for processing personal data.
- Other Risks and mitigations
- Any costs involved.
- Scope to scale the use of this new AI across the organisation.
- If the AI will process data or be used to inform decisions that directly affect people an ethical approval will also need to take place (see Annex

Record Keeping

To ensure compliant use of authorised AI applications and enhance usage by the workforce, we will document any use of authorised AI applications for business purposes.

We will implement an audit system to monitor and document all prompts and output received from authorised AI applications.

Training and technical support

All staff shall be provided with AI literacy training to understand responsible use of AI and the ethical considerations. Where a new AI is being introduced the staff lead will work with IT to ensure appropriate training is provided to any users before they use the new AI. Failure to complete the training may be treated by us as a breach of this policy.

For technical support in accessing and using authorised AI applications, please [contact [[POSITION] OR the Head of the IT Department] OR refer to our staff guidance which can be found on the [intranet OR [LOCATION]]].

Reporting Concerns

Staff must report any concerns with regards use of AI and any breach of this policy to the Director of People and Governance and the Director of Finance and IT immediately.

Failure to Comply

Failure to comply with this policy may result in disciplinary action up to and including dismissal or termination of your employment or engagement with us. Where disciplinary action is appropriate, it may be taken whether the breach is committed during or outside normal hours of work and whether or not use of AI is on an individual's own device or one of our devices, and whether at home, in the office or from a remote working location.

Monitoring and Review

This policy will be monitored and reviewed six monthly to reflect the rapid changes in technology, law, and societal expectations.

Responsibilities

The Board of Trustees: The board of trustees is responsible for ensuring appropriate policies and procedures are in place to safeguard the organisation and to comply with all with data protection and other relevant laws.

The NHS Confederation **Chief Executive** is responsible to the board and accountable for ensuring the organisation maintains high standards of governance and compliance through the adoption and implementation of appropriate policies and procedures.

Directors are responsible for implementing the policy and ensuring adherence by all users. They are responsible for ensuring responsible AI practices are embedded in the organisation's operational activities.

The **Director of People and Governance and Director of Finance and IT** are responsible for overseeing the effectiveness of the policy and safe use of AI in the organisation. This includes ensuring effective training is in place.

Employees are responsible for ensuring they comply with the requirements of this policy and all related policies, and for prompt reporting of any concerns. All employees are responsible for ensuring they participate in any mandatory training to support responsible AI use.

Guidance notes

What is AI?

Artificial Intelligence (AI) is a branch of computer science that enables machines and systems to simulate human intelligence and problem-solving capabilities. AI technologies can perform tasks that typically require human intervention, such as recognising speech, making decisions, translating languages, and more. It encompasses various subfields, including machine learning and deep learning, where algorithms are designed to analyse data, learn from it, and make predictions or decisions. While AI can range from simple automation to complex decision-making, it is continually evolving, aiming to augment human abilities and improve efficiency across numerous applications.

What are the risks that could affect me?

While AI has the potential to drive innovation and efficiency, it also presents various risks that must be managed responsibly. The key risks for users to consider include:

AI Hallucinations occur when an AI system, such as a large language model (e.g. ChatGPT) or computer vision tool, generates a response that includes false or misleading information presented as fact. This happens when AI perceives a pattern that does not exist, leading to nonsensical or inaccurate information being presented. *This is why we adopt a human-in-the-loop approach, where a human is always critically appraising the outputs of AI.*

Bias (algorithmic bias) usually due to skewed training data or a flaw in the algorithmic design:

- AI systems learn from data. If that data is biased, then that will reflect in AI. Recent examples include facial recognition having learnt from data from predominantly white-faced images. This has led to AI performing better for the over-represented group. AI bias can unfairly discriminate against certain groups for individuals. In the USA, police AI was found to be discriminating against BAME communities. *It is therefore important to understand the data sources for any AI.*
- Sometimes the AI algorithm design may inadvertently introduce bias. This can happen when the algorithm's parameters are set in a way that favours certain outcomes over others. *This is why we ask for explainable AI – so that we can describe what the algorithm sets out to do.*

Copyright infringement. AI systems that create content, such as text, images or music raise questions about who owns the copyright to the generated work. The legal framework is still catching up with this and there are a number of high-profile court cases taking place over copyright. *We must therefore be careful about any new content we are creating for external publication through AI.*

Misinformation and fake news

- AI can be used to spread misinformation and fake news. Similarly AI can predict and filter what information users see online, potentially isolating them by limiting their exposure to contrasting viewpoints.
- Deepfakes are the creation of very realistic but fake audio or video content, created to help spread misinformation and manipulate public opinion.

This is why it is important for staff to critically appraise any on-line information.

Privacy. Some AI can be opaque, making it difficult to understand where the data has come from, or where the data you have entered goes. If data is personal data, it is subject to UK Data Protection laws and individuals may not have consented for their data to be used in the way the AI algorithm is processing it. *We do not allow personal data to be processed by any AI model that sits outside of the NHS Confederation IT environment, unless NHS Confederation has entered into an appropriate contractual arrangement (including a data processing agreement within the meaning given in Article 28 of the UK GDPR) with any relevant third party providing external systems upon which AI is used t. We urge caution over the use of AI for any information that is also confidential or commercially sensitive.*

Authorised AI within the NHS Confederation

The NHS Confederation permits the use **Microsoft CoPilot** as this sits within our own IT infrastructure and therefore all data is retained under our control. Microsoft CoPilot is currently the only AI permissible by the NHS Confederation. Please note there are other AI's using the name CoPilot – these are not permitted.

Requesting use of other AI

To ensure that we are responsible in the use of AI, any new AI to be adopted must go through an approval process. You will need to submit a proposal to the Director of Finance and IT which outlines the AI and how you intend to use it, and how you are managing the risks (See Annex 1).

AI is constantly evolving

The world of AI and the legislation and best practice around this are evolving rapidly, all employees are asked to keep up to date with the latest policy.

Annex 1: AI approval Form

Please use this form if you are proposing to one or more of the following (please tick)

To adopt new AI (please also refer to the Software Policy)	
To use AI to process personal data (please also refer to the Data Protection Policy and you may also be required to complete a Data Protection Impact Assessment)	
To use AI to inform decision making that will impact on people	

Date completed	
Completed by	
Review by	
Reviewed Date	
Agreed actions:	
Proposal	Approved/Declined/Deferred

Please outline the purpose
Please outline source of AI and supporting details about how it works
What are the security arrangements of the AI (e.g. hosting, data etc)
<i>Include where it was developed (internally or externally)</i>

Is the result of the AI system to be used for decision making?	
Yes/No	
How will you mitigate against bias?	
What ethical considerations have you identified and how will these be managed?	
<i>e.g. transparency, discrimination</i>	
Please outline your processes to highlight where there will be human oversight and scrutiny; as well as defining clear accountability (roles and responsibilities)	
Risks and Mitigations	
Risk	Mitigation
EU Compliance Checker score: EU AI Act Compliance Checker EU Artificial Intelligence Act we use this to check compliance against EU AI Act – whilst not covering specifically the UK it provides a good steer on emerging legislations and ethical position on AI	
Cost of AI – please detail cost implications.	

Privacy and data protection arrangements – <i>please complete if personal data being used</i>	
Who is the data controller?	
Who is the data processor	
Have you completed a data protection impact assessment?	Y/N
Please explain why AI is needed for processing personal data	
Risks to rights of freedom of individuals	
Risk	Mitigation

Annex 2 Risk Assessment Framework

To be completed once AI identified as part of the workstream around this.

AI (with description)	Risk (High, Limited, Minimal)	Mitigations

Equality Impact Assessment - Policies

The following guidance and checklist provides a framework for Equality Impact Assessments (EIA). It should be used when carrying out equality impact assessments (EIA) in relation to any new or revised policy. The checklist will help in considering the impact of the policy in relation to equality and diversity (E&D).

The Checklist is to be used for any new or revised policy, not just those that appear to have high relevance in relation to equality and diversity issues. Completion of the Checklist does not need to be a time-consuming or difficult process but should raise some important questions as you carry out the process.

Name of policy being assessed	AI Ethics and Compliance Policy
Policy Owner	
EIA completed by	Head of Governance and Compliance
Date Completed	20/08/2024
Summary of purpose of the policy	To set out how the Confederation will approach the use of AI in an ethical and compliant manner
Who are the main stakeholders and what involvement and consultation have they had in the policy development. Include staff groups, trade unions and board committees as applicable.	Staff. Policy has been written with support from Capsticks legal team
Who is affected by the policy	All staff
What are the arrangements for monitoring and reviewing the actual impact of the policy	At least annual review due to changing nature of AI

Please indicate against each of the following protected characteristics, what the impact of the policy would be and actions that will be / have been taken to mitigate any negative or adverse impact identified.

(Where the policy is found to have either a positive or negative impact on a particular group it will need to be reviewed or justified within the permits of the law.)

Protected Characteristics	Impact Y/N	Action(s) you will take to mitigate or remove the negative or adverse impact if identified?	Action Owner
Age <i>Consider impact on young people, older people etc.</i>	y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Disability <i>Consider people with physical disabilities, hidden disabilities and neurodiversity.</i>	Y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Gender Reassignment <i>Consider people undergoing or have undergone gender reassignment</i>	Y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Pregnancy and Maternity <i>Consider those who are pregnant and those on pregnancy and parenthood leave. Consider those wishing to take parenthood leave</i>	Y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Race / Ethnicity <i>Consider potential impact on people from different ethnic groups and nationalities.</i>	y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	

Religion or Belief <i>Consider people with different religious, faith and non-beliefs</i>	Y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Gender <i>Consider all genders.</i>	Y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Sexual Orientation <i>Consider LGBTQ+ people.</i>	y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Marriage and Civil Partnership <i>Consider marriage and civil partnership in respect of the due regard to the need to eliminate unlawful discrimination in employment.</i>	y	AI has the potential for significant bias and therefore the policy has set out key principles for mitigating against bias. Any user of new AI will need to complete an approval form which includes how they will mitigate any bias	
Does the policy promote fairness and equal opportunities? Provide details.	The aim of the policy is to ensure proper ethical use of AI and avoidance of discrimination/bias as well as ensuring legal compliance.		
Manager Signature:		HR Review Signature:	
Date:		Date:	

