DERBYSHIRE
County Council

Artificial Intelligence (AI) Policy

1 Version History details and author

1.0	17/05/2024	Approved by Information Governance Group.	Jo White
2.0	20/11/2024	Reference to AI agents added by SIRO.	Jo Williams
3.0	26/11/2024	Glossary updated, Use restricted to Copilot, No AI decision making,	Jo Williams
4.0	03/01/2025	ChatGPT and Copilot descriptions expanded.	Jo Williams
5.0	07/01/2025	CMT amendments.	Leonardo Tantari
6.0	12/08/2025	Amendments to sections 6, 8 and 11. “Microsoft Copilot 365” to “the Council provided Microsoft Copilot 365 applications”.	John Allen
7.0	10/03/2026	AI Risk assessment, AI governance framework and AI Register references added.	Information Givernance Group.

This document has been prepared using the following ISO27001:2022 standard controls as reference:

A.5.2 - Information Security roles and responsibilities
A.5.12 - Classification of information
A.5.14 - Information transfer
A.5.34 - Privacy and protection of PII
A.5.37 - Documented operating procedures
A.8.12 - Data leakage prevention
A.8.16 - Monitoring activities
A.8.20 – Networks security

2 Introduction

Derbyshire County Council recognises that Artificial Intelligence (AI) technology is already widely used in both commercial and everyday applications, and its influence is anticipated to grow exponentially, impacting almost all industries and job sectors including the public sector, particularly with the development of Generative AI, which is a specific type of AI.

Whilst AI was typically used by specialists to complete a specific task, Generative AI has become a widely used tool in a short period of time that is more accessible by non-specialists. Generative AI is a rapidly evolving and increasingly freely available technology which can generate new writing, audio, codes, images and video simulations. Whilst this offers opportunities for the Council and the residents that it serves in enhancing efficiency, decision-making and service delivery, it also increases risk.

There are many Generative AI tools that are freely available, including ChatGPT (owned by OpenAI), Google Gemini, and Microsoft Copilot 365. As with any other IT related technology, the Council needs to ensure that the use of AI, particularly Generative AI, is organised and controlled in a manner which will be beneficial to the safety, integrity, and reputation of the Council.

The Council recognises that this is a rapidly developing area and like many other organisations, continually reviews and adapts practice to changes in technology. The Council Digital Strategy includes an acknowledgement and commitment to exploring the opportunities offered by AI technology. This policy will be regularly reviewed by the Information Governance Group to ensure that it remains relevant and applicable in practice during this time of rapid change. The Council recognises that it is important to provide early appropriate guidance and advice to employees, to encourage users to be transparent about the use of AI, and to support a culture of responsible AI use as these rapid changes take place. This includes AI embedded within mainstream software and emerging AI agents that can support or automate tasks, which will be governed under this policy and the council’s AI governance framework.

Departmental Managers will be responsible for monitoring the use of AI within their service areas to ensure adherence to this policy.

3 Purpose

The purpose of this policy is to ensure users are aware of the controls and methods the Council has put in place to manage the use of AI. Users are expected to comply with the policy to ensure that AI tools are used appropriately. The policy is supported by the council’s AI governance framework, which sets out how AI systems will be identified, risk assessed, documented, registered and reviewed before and during use. There are a number of other relevant Information Security policies which this links with here. This policy is supplementary to existing policies and should be read in conjunction with them.

4 Scope

This policy applies to the use and configuration of all AI Tools and systems that have been provided or procured by the Council, including Microsoft Copilot 365, AI solutions developed in house, and AI capabilities embedded in contracted third party software. Other unsanctioned AI tools, which may be available for free or on a subscription basis and directly accessed by users via the internet must not be used. These rules apply when using both corporate-owned devices and any personal devices which may be used for Council business. The policy covers all employees, agency staff, elected members, contractors, volunteers, apprenticeships, student/work experience placements and partner agencies who have access to these tools, described as “users” within this document.

5 Responsibilities

There are a number of roles in the Council that form key contributors to AI policy and development:

- The Digital Director acts as a lead for the Council regarding the use of AI technology and works with departments to monitor compliance with and enforce the policy.
- Monitoring staff use of AI will, in the first instance, fall to line managers and Data Protection Liaison Officers (DPLOs).
- The Information Governance Group and the Digital Director will communicate, promote and regulate AI use, providing or arranging for training to be given where necessary.
- The Data Protection Officer is responsible for providing advice on data protection obligations in relation to AI use.
- Digital Services will provide technical support and guidance on the operation of AI ~~in lieu of the creation of the AI Centre of Excellence (COE).~~
- The Information Governance Group will be responsible for the Governance of AI and the review of this policy.
- The Caldicott Guardian provides a significant role in any ethical decisions around the use of AI and acts as the ‘conscience’ of the Council.
- The Council Digital Champions have a role to provide support to help people with the introduction of Microsoft 365 in the Council.

6 Policy Statement

Whilst there are a number of freely available Generative AI tools such as ChatGPT, Claude and Gemini, **users are permitted to use only the council provided Microsoft Copilot 365 applications as the default environment for AI,** as any data entered remains within the tenancy owned by the Council. In house AI solutions and AI capabilities embedded in contracted third party systems may only be used where they have been assessed and approved in line with this policy. This also allows the Council to have better Guardrails and oversight of the use of Generative AI by users, which is key to effective governance. If Microsoft Copilot 365 is not suitable for the task required, then users will consult with their line manager who will address the issue through the AI Governance Framework for further advice.

All AI tools and systems must be recorded in the council AI Register and assessed using the council’s AI risk model, with evidence retained in line with the AI governance framework.

All users of AI will comply with applicable laws, and Data Protection policies. There will be no unauthorised use of copyrighted material or creation of content that infringes on the intellectual property of others. Users will prioritise the safeguarding of stakeholders and will not knowingly use any AI technology that puts their safety or privacy at risk. Users will not allow or cause intellectual property, to be entered into Generative AI models without appropriate consent or exemption to copyright.

All users of AI will recognise that the technology is rapidly evolving and will be committed to adapting ways of working as necessary in line with this policy.

Users will be transparent and accountable about the use of AI technology so that stakeholders understand where and how AI is used and who is responsible. Key documents such as Privacy Notices will be updated where relevant to ensure that there is transparency for data subjects affected. Any stakeholder feedback or questions about the use of AI will be considered and responded to appropriately, in line with Council policy and processes.

By adhering to this policy, users understand and support the Council's aim to foster a responsible and inclusive environment for the use of AI by upholding privacy, fairness, and transparency for the benefit of all involved.

By combining the benefits of AI technology with professionals' expertise, experience, and professional judgment, users understand that they can create a collaborative and effective service that maximises the benefits of both human and AI capabilities.

7 Use of AI tools

Users are permitted to explore and utilise **approved** AI-based tools and technologies to assist in managing their work, subject to the restrictions in this and related Council policies. Examples of such tasks may include suggesting improvements to documents, call transcription, report writing, data analysis, summarisation of large or specialist documents, translation, drafting of communications materials, content creation, managing workflows and reviewing materials for accessibility. AI can provide valuable support while still incorporating users' professional judgment and expertise.

Users will not use AI to replace formal decision making, such decisions will continue to be made by Council, Cabinet, Committees, Cabinet Members and officers under delegated powers.

AI should not be relied upon to provide legal advice or interpret legislation, as this can lead to inaccuracies and potential legal risks. Legal interpretation should always be carried out by suitably qualified professionals.

AI tools will be used responsibly, ensuring they complement staff professional judgement and expertise, without replacing them. Users remain professionally responsible and accountable for the quality and content of any output generated by AI, however generated or used and staff should rely on their expertise to ensure that they review and tailor any AI output.

AI notetakers should never be used in a meeting without completing a data protection impact assessment and making all participants aware before the meeting starts. If a participant objects to the use of an AI notetaking app, the meeting organiser should take into account the nature of their objection and if their concerns cannot be overcome, the organiser should consider whether it is proportionate to continue the meeting without the use of AI. Although it is difficult to control the actions of external attendees at meetings which you organise, you should let attendees know upfront that they should not use AI notetakers and as the meeting organiser you will take responsibility for the taking and distribution of notes.

AI agents may only be used where they have been assessed and approved in line with the council's AI governance pathways. This includes completing the required AI risk assessment, confirming the intended use aligns with approved patterns, and ensuring the system is registered in the AI Register. All AI outputs must be subject to appropriate human oversight, with users reviewing, verifying and retaining responsibility for the final content or decision. Higher risk AI systems will be monitored in line with their review schedule as set out in the council AI governance framework.

As the Council continues to explore the potential of Generative AI, it is important to recognise that its implementation will be continuously reviewed to ensure it is aligned with the Council's core culture and values. The review process will assess how Generative AI influences the Council's ethos, including a commitment to transparency, inclusivity and ethical practices. This process will enable the Council to harness the benefits of this innovative technology while safeguarding the principles that define the Council's identity and interactions with its community and stakeholders.

Users will receive appropriate training and support to effectively integrate AI into their work, including professional development opportunities focused on AI tools and their effective integration into working practices. Training and support will be planned as part of user personal development reviews and appraisals or on an as-needed basis.

8 Data Protection implications of using AI

Users should be aware that any information entered into a Generative AI model may no longer be private or secure. Where entering personal data or private information (including information that has intellectual property implications or contains commercially sensitive information, such as contracts) users must only use the Council provided Microsoft Copilot 365 applications, or approved software.

Users who wish to utilise AI tools in a systematic way must consider if the potential new use is likely to be a data processing or profiling activity for which a Data Protection Impact Assessment is required. In these instances, users must follow the Corporate Data Protection Policy.

Where the use of AI is likely to result in a high risk to individuals' rights and freedoms it will be a legal requirement to carry out a Data Protection Impact Assessment (DPIA). Users should be aware that it is not always easy to recognise when AI tools are processing personal data and they should not presume that no processing is taking place.

Any DPIA or assessment of the data protection aspects of the use of AI may also include:

- What alternatives (both AI and non-AI) are there to the planned processing and what justification is there in choosing this method and how it is fair.
- A clear indication where AI processing may produce effects on individuals.
- Consideration of both individual and allocative harms (for example, where the harm results from a decision to decline a service to a particular person) and representational harms (for example, selecting groups of service users for different provision results in gender or racial bias).
- How the use of the AI tool is proportionate and fair by assessing the benefits against the risks to the rights and freedoms to individuals and/or whether it is possible to put safeguards in place.
- An analysis of any bias or Hallucination which may result in detriment to individuals.
- If the use of AI replaces human intervention, a comparison of the human and algorithmic accuracy in order to justify the use of the AI tool in the DPIA.

- Relevant variation or margins of error in the performance of the system, which may affect the fairness of the processing (including statistical accuracy) and describe if/when there is human involvement in the decision-making process.

9 Ethical use of AI

The use of AI systems, in particular Generative AI, will be carried out with caution and an awareness of their limitations. Users should be mindful of the following considerations:

Bias – data and information generated by AI will reflect any inherent biases in the data set accessed to produce it. This could include content which may be discriminatory based on factors such as race, gender, or socioeconomic background. Users must ensure they follow the requirements of the Equality Act 2010 as set out in the Council Equality Impact Assessment process. Particular care must be taken where AI could be used for profiling purposes (for example, identifying the most suitable candidate characteristics for a particular job role).

Accuracy – information may be inaccurate when generated so any content should be fact-checked. Tools may provide highly plausible and coherent results but may still contain errors (Hallucination).

Human oversight – a lack of human intervention may result in AI outputs going unchecked. Humans should ensure that outputs align with societal values, and users should be aware that Generative AI lacks flexibility, human understanding and compassion.

Currency – some AI models only collate data prior to a certain date so content generated may not reflect the most recent information.

Environmental issues – use of AI requires energy to run. Therefore it should only be used when relevant, appropriate and proportionate, where it is the most suitable and sustainable option.

10 Breaches of Policy

Breaches of this policy and/or security incidents can be defined as events which could have, or have resulted in, loss or damage to Council assets, or an event which is in breach of the Council’s security procedures and policies.

All Council employees, elected members, partner agencies, contractors, agency staff, volunteers and vendors have a responsibility to report security incidents and breaches of this policy as quickly as possible through the Council’s Incident Reporting Procedure. This obligation also extends to any external organisation contracted to support or access the Information Systems of the Council

The Council will take appropriate measures to remedy any breach of the policy and its associated procedures and guidelines through the relevant frameworks in place. In the case of an individual then the matter may be dealt with under the disciplinary process.

11 Glossary of terms

Agentic AI is the use of AI agents that may plan or sequence multiple steps, interact with tools or data sources, or act with defined autonomy subject to human oversight.

AI agent is an AI driven software component that receives inputs, interprets them, and produces outputs or actions that support or automate tasks.

AI Register is the council record of all AI systems in use or under development, including their purpose, risk classification, documentation and review dates.

Algorithm is a rule given to an AI machine to perform a task.

Artificial Intelligence (AI) is an umbrella term for a range of algorithm-based technologies and approaches that often attempt to mimic human thought to solve complex tasks, these may include, visual perception, speech recognition, decision making, and translation between languages.

Centre of Excellence is a group with a shared area of focus and subject matter expertise that they use to support others, usually by providing tips, insights, training and research.

ChatGPT is owned by OpenAI LP, an artificial intelligence research lab. GPT stands for 'Generative Pre-trained Transformer'. It means that the model has the ability to generate text or other forms of output. ChatGPT is primarily trained using public data from the internet.

Generative AI is a form of AI, which produces new content, such as images, text or computer code. It works by using large quantities of data, often harvested from the internet, to train a model in the underlying patterns and structures of that data. After many rounds of training the model is capable of generating new content. When a user provides a prompt or input, the AI evaluates the likelihood of various possible responses based on what it has learned from its training data. It then selects and presents the response that has the highest probability of being the right fit for the given prompt. That prompt and response then may be fed back into the model to provide further training.

Guardrails are restrictions and rules placed on AI systems to ensure they handle data properly and ethically.

Hallucination is when AI presents information as fact when it is not actual fact.

Large Language Model (LLM) is a huge database of language knowledge that can write articles, answer questions or create realistic dialogue and is pre-trained on large amounts of data.

Microsoft Copilot 365 is Microsoft's version of ChatGPT. The Council provides the Microsoft Copilot for Work Applications. These include an integrated application that works with Microsoft 365 applications such as Word, Excel and Outlook and a browser based version.

Natural Language Processing (NLP) understands written and spoken language eg translations.

Shadow AI is any AI tool or service that has not been approved for use by the council, including external web-based AI systems and AI features in software that have not completed the council's AI governance route.

This document is owned by the Information Governance Group and forms part of the Council's ISMS Policy and as such, must be fully complied with.