



North East Museums

Artificial Intelligence Policy and Guidance

Name of governing body: North East Museums Strategic Board

Other Bodies signed up to policy:

NEM Development Trust

TWAM Enterprises Ltd

Exhibitions by TWAM Ltd

Date of approval: 30 09 2026

Date for review: September 2027

Contents

Purpose / Introduction	3
What is Artificial Intelligence?	4
Generative AI	4
Generative AI: Microsoft Copilot.....	5
The 'Golden', 'Governing' and 'Data' Rules.....	6
Risks and Limitations of AI	7
Key Principles	8
Risk Assessing possible use of AI.....	12
Appendix 1 - Risk Assessment Template for possible use of AI	13

Purpose / Introduction

The purpose of this policy is to provide clear guidelines for the use of artificial intelligence (AI) at North East Museums (NEM). It aims to give anyone working or volunteering for, or on behalf of NEM (Including NEM Development Trust, TWAM Enterprises and Exhibitions by TWAM) a better understanding of the potential benefits of using AI and how to mitigate its risks, so that we can ensure effective, secure, responsible and transparent use of AI technologies when delivering our mission.

The policy has been designed to provide:

- An overview of AI, and in particular 'Generative AI'
- An overview of Microsoft Co-Pilot, an AI tool which NEM staff have access to
- The 'Golden' and 'Governing' and 'Data' rules staff must follow when using AI
- The risks and limitations to be aware of when using AI
- 10 Key Principles staff should apply when using AI
- A Risk Assessment template (Appendix 1), for assessing the potential risks and benefits of using an AI tool.

It also contains links to several external resources throughout, which staff may find useful for furthering their understanding of AI.

The policy has been informed by the following publications, which you may find useful to refer to in full:

- UK Government's [AI Playbook](#)
- Newcastle City Council's guidance on [AI Governance: Ensuring Compliance and Transparency](#) and [Managed Copilots & Generative AI Tools](#)
- Watford Borough Council's [Artificial Intelligence Policy](#)
- Leicester City Council's [Generative Artificial Intelligence \(GenAI\) policy](#)
- Arts Marketing Association's [AI Sector Support Example Ai Policy](#)
- Royal National Lifeboat Institution (RNLI) [Generative AI Fair Usage Policy](#)
- The Wildlife Trust's [AI Risk Assessment](#)
- The Audience Agency's [Let's Get Real: AI Report](#)

Due to the fast-paced evolution of guidance and regulation in the field of AI, this policy should be reviewed every 12 months.

What is Artificial Intelligence?

The UK Government uses the [definition of AI adopted by OECD countries](#):

“An AI system is a machine-based system that... infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.”

Put more simply: it is the ability of computer systems to perform tasks that have historically required human intelligence, such as learning, problem-solving, and decision-making.

The use of AI can bring significant benefits for organisations, including delivering efficiency savings, improving how data is used, producing more accurate or higher quality outputs, and ultimately delivering a better service for customers.

This policy does not seek to describe every different type of AI, as the vast majority do not apply to the work carried out by North East Museums.

However, it does focus on one particular type in more detail: ‘Generative AI’, since this is currently the most widely used and popular type of AI, and also likely to be the most used by North East Museums’ staff.

The UK Government’s [AI Playbook](#) provides a good overview of some of the other main fields of AI.

Generative AI

Generative AI uses machine-learning models to generate content (such as text, images, video or audio), perform analysis and evaluation, develop code or aid creative processes.

Generative AI uses large quantities of carefully selected data to train models. Once trained, generative AI models are capable of generating new content consistent with what they have learned from the training data. When a user provides a ‘prompt’ or ‘input’, the model evaluates the probability of potential responses based on what it has learned from its training data. It then selects and presents the response that is likeliest to be the right fit for the given prompt.

Whilst lots of modern applications and programmes use AI to perform tasks and solve problems, Generative AI is different, in that it produces ‘new’ or ‘original’ content, which comes with both benefits and risks.

Generative AI: Microsoft Copilot

As part of our standard Microsoft licensing package, North East Museums staff have access to a Generative AI Tool: 'Microsoft 365 (M365) Copilot'.

M365 Copilot is a Generative AI tool designed to help you work more efficiently within Microsoft apps like Word, Excel, PowerPoint, Outlook, and Teams, as well as interacting with the web.

By providing suggestions, generating content and offering insights, M365 Copilot can assist staff in writing and editing documents, analysing data, creating presentations, managing emails and meetings, writing code, generating images and answering general questions.

NCC has produced an article on the [top 10 ways M365 Copilot can help us become more productive.](#)

Logging in to Office 365 using a work device, ensures you are working within the NCC managed version, and ensures any information you enter into M365 Copilot has 'enterprise data protection', is protected from external access. [You can find out more about this online, here.](#)

Because of this, M365 Copilot is the preferred Generative AI tool to use unless there is very good reason to use a different tool.

If you do wish to use a Generative AI tool other than M365 Copilot, you must email Corp.Gov@northeastmuseums.gov.uk outlining your reasons, for approval first.

Note: This does not apply to 'AI overviews' provided by web search providers, which can be used alongside M365 Copilot, in-line with the rules, key principles and risks and limitations outlined in this policy.

The 'Golden', 'Governing' and 'Data' Rules

NEM has three clear rules when using Generative AI, which all those using it must follow:

Golden Rule: You can't say "AI did it". You are responsible for your AI outputs and you should always check the output before sharing or saving.

Governing Rule: AI technology can be used to help replace 'task-based' processes and for process automation but should never be used to fully replace a 'knowledge process'. For knowledge processes, AI technology must only augment the decision-making process empowering enhanced informed data-centric decisions.

Data Rule: You must not submit any personal or confidential information into any Generative AI tool, other than M365 Copilot. As with any non-managed online tool, we have no control over how or where the information is used or processed.

- Personal data can be defined as any data that would allow the identification of an individual.
- Confidential information can be defined as any information about NEM that should not be made publicly available (including but not limited to, financial, HR, legal or security information).

If you have any questions, or you believe any of the above rules have been broken, please contact Corp.Gov@northeastmuseums.gov.uk immediately, in line with NEM's GDPR processes.

NCC has also produced some other important guidance [on using Generative AI tools](#) responsibly and effectively. **Please ensure you read these.**

Risks and Limitations of AI

The capabilities and potential uses of AI are evolving all the time. However, they must still be treated with caution. Some of the main risks and limitations of AI systems are:

- **Accuracy:** AI systems heavily rely on the quality and quantity of data to drive accuracy. Insufficient data can lead to the model failing to generalise effectively, and, similarly, poor quality data can lead to inaccurate outputs.
- **Bias:** AI systems lack consciousness and their outputs can replicate any bias present in the data they were trained on.
- **Hallucination (confabulation):** AI systems are vulnerable to creating content that appears plausible but may actually be factually incorrect, as they generate content by returning the most likely output for a given input based on the patterns learnt from the data they were trained on, without an actual understanding of the content.
- **Lack of critical thinking, personal experience and judgement:** although some AI systems can give the appearance of reasoning and their outputs may appear as if they come from a person, they are in no way sentient.
- **Sensitive or inappropriate content:** AI systems can generate offensive or inappropriate content if not properly guided, since they can replicate any bias or toxic material present in the data they were trained on.
- **Domain expertise:** AI systems are not experts in any one area. They are not a substitute for professional advice, for example legal, financial, HR, or collections research, where precise and contextually relevant information is essential.
- **Dynamic real-time information retrieval:** some AI systems such as Microsoft Copilot are now able to include access to real-time internet data in their results. But there are still many that do not have real-time access to the internet or data outside their training set and are therefore at risk of their output being out of date.
- **Sustainability:** Data centres that house AI servers are large consumers of water, rely on critical minerals and rare elements, and use huge amounts of electricity. A [2023 study](#) by researchers in the USA estimated that for every 20 to 50 questions ChatGPT is asked, it “consumes” a 500ml bottle of water.

Newcastle City Council have produced [guidance on some limitations specific to M365 Copilot](#)

Because of the above risks, it remains **the responsibility of every member of staff** to critically review, validate and verify any AI-generated output before it is:

- uploaded to any organisational system;
- shared with customers, partners or third parties;
- used in reports, documentation or decision-making;
- issued as an official organisational communication.

Key Principles

North East Museums encourages its staff to engage in responsible use of AI by applying the following 10 key principles:

1. Rationale for using AI

Any use of AI tools at North East Museums should be done with the express purpose of helping us deliver our mission and strategic goals.

While recognising the potential of AI, North East Museums believes in the enduring importance of human oversight and creativity.

AI systems should be used as a tool to complement and enhance human decision-making and creativity, not replace it.

Key questions to consider

- Does the proposed AI use help deliver our mission?
- What criteria is being applied to establish success?
- Are there any risks that need to be assessed before going ahead?
- Does anyone need to be consulted before AI tools can be used?
- How do we draw the line on what enhances human creativity and what replaces it?

2. Data Protection and Privacy

When using AI systems that process personal data, we will comply with all relevant data protection legislation, including the UK GDPR.

You must not submit organisational, personal, confidential, or commercially sensitive information into any online AI tool. As with any non-managed online tool, we have no control over how or where the information is used or processed.

However, there are two exceptions to this:

- if you are securely logged into M365 Copilot,
- if another tool has been pre-approved for use.

You can use AI tools with non-sensitive / publicly available information.

Please see Newcastle City Council's [guidance for how to use M365 Copilot securely](#) and [guidance on protecting data when using other Generative AI tools](#).

Key questions to consider

- Am I entering organisational, personal, confidential, or sensitive information into the AI Tool?
- Am I using M365 Copilot, or a different AI Tool?
- Do I need to seek pre-approval for a use of different tool?
- Do I need consent? How will I get it?
- Do I need to carry out a data protection assessment?

3. Intellectual Property and Cultural Ownership

We respect the intellectual property rights of artists, creators, and communities whose works or cultural heritage may be used in AI tools.

It is prohibited to use AI to generate content that infringes upon the intellectual property rights of others. For example, using AI to produce a design, could produce something based on a copy of a design that is copyrighted.

We will obtain necessary permissions and give appropriate credit where it is within the scope of our control.

We will carefully consider the cultural and ethical implications of using AI in relation to objects or knowledge of cultural significance.

Key questions to consider

- How do we ensure AI outputs have not infringed upon the rights of others?
- How/when do we assess if our use of AI impacts culturally sensitive material?
- When do we need to consult relevant community stakeholders?
- How do we ensure appropriate attribution?

4. Human Oversight and Quality Assurance

We are committed to ensuring that any content created by AI is of the highest quality.

We recognise that AI systems cannot be relied on to provide 100% accurate outputs and can amplify biases present in training data and design. As AI tools become ever more commonplace and integral to our work, it will be more important than ever that humans have the final say.

We will take proactive steps to identify and mitigate biases by ensuring all AI-generated content is reviewed by staff to ensure accuracy, as well as alignment with our brand voice, mission and values.

We will be alert to the risk of AI-generated content creating a misleading or unbalanced interpretation of art, history, or culture.

We will be clear about what measures the AI outputs are being assessed against.

Key questions to consider

- What human checks can we put in place?
- Who is responsible for reviewing AI outputs / do we need to escalate decision-making?
- What criteria is being applied to ensure quality?
- Do we need to consider AI content coming in from external sources?

5. Transparency and Accountability

We will be transparent about our use of AI systems both internally and externally.

Staff must be transparent about the use of AI in their work, ensuring that any audiences are aware of the technology's involvement in the process.

When publishing AI-generated content or using AI in applications, we will clearly label or state this. For example, you could add a footnote such as:

“Note: This document contains content generated by Artificial Intelligence (AI), using Microsoft Copilot. AI generated content has been reviewed by the author for accuracy and edited/revised where necessary.”

Staff are responsible for the outcomes generated by AI systems and must be prepared to explain and justify those outcomes.

Audiences are more likely to trust in the use of AI at North East Museums, if we are able to demonstrate that the AI solution has been developed and used competently, responsibly and ethically.

Key questions to consider

- How do we ensure we have been transparent about when AI has been used?
- Are we able to audit what AI has been used for, the rationale behind the decision to go ahead and who approved its use?

6. Stakeholder Engagement and Ethical Review

We will proactively engage with our audiences, local communities, cultural stakeholders, academic experts, and policymakers to inform our approach to AI.

We will encourage collaboration and knowledge sharing, including on ethical guidelines, and responsible AI practices.

Where appropriate, we will carry out ethical reviews to assess AI projects or processes and ensure they align with our values and those of our partners.

Key questions to consider

- Which parts of our AI use should be informed by stakeholders?
- Who needs to input into designing the ethical review process to reduce bias?
- Does an ethical review need to take place? How will we do this with our partners/artists?

7. Resource Management

If not available for free, any use of AI must be resourced appropriately and sustainably.

Key questions to consider

- What budget is required?
- What staff training is needed?
- How do we measure Return on Investment?

8. Environmental Sustainability

Recognising the potentially significant environmental footprint of AI, we will aim to use AI tools efficiently and avoid unnecessary energy consumption wherever possible.

Key questions to consider

- How do we balance the opportunities of AI with our Environmental policy?
- Can we assess whether the impact of an AI tool represents a reasonable trade-off between benefits and energy consumption, and whether a less energy-intensive system might be able to achieve the same or similar results?
- Can we evaluate the environmental credentials of potential model providers and wider partner organisations – and select those with sustainable practices, or low(er) carbon emissions?

9. Social Impact and Job Displacement

We will strive to use AI in ways that promote cultural understanding, inclusion, and accessibility. We will be mindful of the potential impact of AI on our workforce and commit to supporting staff in developing the skills needed to work effectively with AI.

Key questions to consider

- Will the use of AI risk a negative impact on our staff or audiences?
- How do we assess any risk?
- How do we plan the development of skills?
- How do we ensure that unintended bias is not influencing results?

10. Use of AI in Recruitment

We are committed to using AI in recruitment processes responsibly to ensure fair and equitable hiring practices.

AI tools may assist in the recruitment process, but final decisions will be made by human recruiters to ensure a holistic evaluation of candidates.

Candidates will be informed if AI tools are being used in the recruitment process, and we will provide explanations of how these tools influence decision-making.

Candidates clearly using AI in applications will not be rejected automatically but this may harm their chances of being invited to interview if they fail to personalise their application.

Key questions to consider

- What selection criteria will we use and how will we score this?
- Do we need to weight answers about specific experiences, attitudes or skills more highly than subject knowledge?
- Should we make ability to use AI a selection criterion for some roles?
- Can we assess how the candidate used AI, where they added to it, and how they applied judgement?

Risk Assessing possible use of AI

It should be clear from the information provided so far in this policy, that AI systems and applications can have both a positive and negative outcome.

If not used in responsible ways and with appropriate safeguards in place, AI tools may risk generating harm to the organisation.

When considering the use of an AI tool, you should identify both the potential positive impacts of new technologies as well as any negative impacts or unintended consequences.

If you are in any doubt as to whether to use an AI tool, a Risk Assessment template is available at **Appendix 1** of this Policy, which you can use to help you assess the risks and benefits of using an AI tool, make a decision, and record the reasoning behind this.

The Risk Assessment template **MUST** be used when considering whether to use a non-approved AI tool.

If, after completing the Risk Assessment, the residual risk rating is Amber or Red, you must contact NEM Corporate Governance team at corp.governance@northeastmuseums for further advice.

The matrix for assessing whether a risk is Red, Amber or Green rating is here:

RISK PRIORITY MATRIX		Impact			
		Min	Low	Medium	High
Likelihood	High	4	8	12	16
	Medium	3	6	9	12
	Low	2	4	6	8
	Min	1	2	3	4

Appendix 1 - Risk Assessment Template for possible use of AI

North East Museums Risk Assessment Template for use of AI	Venue / Team:	Employee Group:
Officers/Manager involved in assessment review:		Date of assessment:

1. Activity description:		<i>e.g. Write fundraising application</i>			
AI Tool being used	Type of AI being used	Potential Benefits	Potential Risks		Control Measures
<i>E.g. Microsoft Co-Pilot</i>	<i>E.g. Generative</i>	<i>E.g. Reduced workload for staff</i> <i>Ability to meet deadline faster</i>	<i>E.g. Accuracy, bias, lack of compelling narrative, failure to relate to NEM mission, values, Investment Principles</i> <i>Failure to secure funding, reputational damage</i>		<i>E.g. Implement human review before submitting.</i> <i>Rigorously validate accuracy and tone to ensure accuracy, as well as alignment with our brand voice, mission and values.</i>
Initial Risk Rating for the above activity before control measures are implemented =>			High, Med, Low Insert RAG colour	Residual Risk Rating for the above activity if all control measures are implemented =>	High, Med, Low Insert RAG colour
Decision to go ahead (yes /no + justification)		<i>E.g. Yes – controls in place will mitigate risks. Tight deadline means we would be unlikely to be in a position to submit application, if not able to utilise AI.</i>			