UEA ETHICS REVIEW OF RESEARCH INVOLVING GENERATIVE AI TOOLS

Authors: Helen Brownlee (Research Integrity Manager, RIN) with the University Research Ethics Committee (UREC); Information Compliance Team - Sue White (DPO), Emma Gedge (Information Compliance Specialist), and RIN - Anna Collins (Open Access and Research Data Officer), Josh Ludkin (Research Integrity Officer)
Date: 29th April 2025 (version 2)

This Ethics Guidance Note is part of a series of documents issued by the University Research Ethics Committee to support the University's Research Ethics Policy. For the full set of Guidance Notes, please visit the Research and Innovation Services' Research Integrity my.UEA page for Research Ethics: Policy and Guidance

This Ethics Guidance Note should be read in conjunction with the University's *Generative AI Policy for Research and Innovation*.

INTRODUCTION

Artificial Intelligence (AI) is a novel technology and is constantly developing, but generally it:

- involves machines using algorithms to find patterns in large amounts of data.
- is the ability to perform repetitive tasks with data without the need for constant human intervention or guidance.

The field of AI has witnessed remarkable advancements, particularly in the realm of generative AI. Generative AI describes any type of artificial intelligence system that identifies patterns and structures in data / information / material and generates content, including: audio, code, images, text, simulations, and videos in response to instructions ('prompts') that resembles human-created content. Generative AI tools can be trained on large language models (LLMs). These models are the algorithmic basis for generative AI tools such as ChatGPT and Gemini. Alongside the many benefits of generative AI tools in research, there are also risks associated with this technology.

Researchers must take full responsibility for the use of generative AI tools in their research and any data / information / material they have entered into those tools, and "must take all reasonable precautions to prevent or diminish loss destruction or damage or any occurrence or cease any activity which may give rise to liability" in order to ensure their activities are covered by the University's Insurers.

Researchers also need to be aware of the potential for inaccurate/biased outputs to be generated by the tool.

The University is committed to providing support and protecting researchers in their use of generative AI tools. As stated by the Information Commissioner's Office (ICO), "Applications of artificial intelligence (AI) increasingly permeate many aspects of our lives. We understand the distinct benefits that AI can bring, but also the risks it can pose to the rights and freedoms of individuals".

Areas of UEA Research Involving Generative AI Tools

Generative AI tools can be involved in many ways during the project lifecycle from design to method to analysis to dissemination. In the UEA setting, we are becoming more aware of these different ways. For example, generative AI tools could be involved in a UEA project to:

- challenge an assumption or to provide an instruction ('prompt').
- collect data to challenge creative ways to use AI.
- contribute to participatory research or co-creative research.
- learn about how people interact with AI.
- help build AI models.
- understand how generative AI tools make decisions and the broader implications of the decisions made by new algorithms.
- tag an image.
- interview a participant.
- analyse primary and secondary datasets.
- undertake a literature review.

In line with UEA's high ethical and research integrity standards, and the UK GDPR, when working with generative AI, **UEA researchers** have a responsibility to ensure that any potential harms must be mitigated against as far as practicable, for example:

- contravening consent and permissions;
- invasion of privacy;
- data inaccuracy;
- bias and discrimination;
- unexpected or unknown transfers or storage to other data controllers and / or out of the UK, and
- non-transparent, unexplainable, or unjustifiable outcomes.

UEA ETHICS REVIEW REMIT

UEA staff and students (or involving UEA), where a generative AI tool will be used, built or developed, and are seeking ethics review anyway as outlined in the University's *Research Ethics Policy*, will have the ethical considerations of using these tools reviewed in their ethics application form. The UEA Ethics Review Committees have a responsibility to review the researcher's documented ethical considerations in their ethics application form to ensure that a study involving generative AI has:

- appropriate participant information to enable informed consent (the UREC has provided a template of the Participant Information Sheet (PIS) and Consent Form to use for projects involving generative AI tools);
- adequate protection of the rights of individuals and groups, in particular children and other vulnerable people, and respect of their dignity;
- transparency and will be carried out with integrity;
- a clearly defined line of accountability;
- considered the potential future reuses of data in connection with any generative AI tool;
- addressed the broader ethical and societal impact of the research, and

- considered the safeguarding required to mitigate against the risks the generative AI tool could potentially introduce, for example, unfairly biased decision making, discriminatory outcomes, or unexpected data linkages (more details are provided below).

As best practices on the involvement of generative AI tools in research are formed at UEA, and generative AI tools are vetted by UEA; the University's governance processes in this area will also be further streamlined. Currently, it remains that the University strongly advises researchers not to enter any personal[1] (including special category data),[2] confidential,[3] third party,[4] or UEA business critical[5] data / information / material into a generative AI tool until the general data privacy and security aspects of generative AI tools are more established. However, there may be exceptional circumstances where this is appropriate, and prior UEA approval would need to be obtained from the Information Compliance Team (to ensure compliance with Data Protection legislation) and the relevant IT Team (to approve the security of the tool) and relevant IT Team following these steps**.** Note that MS Teams is not classed as a Generative AI tool.

Researchers planning to involve generative AI as part of their NHS health research or social care study should refer to the NHS document 'Understanding Regulations of AI and Digital Technology in Health and Social Care' and read the University's Ethics Guidance Note on NHS Health and Social Care Research for advice on how to do so appropriately.

UEA INFORMATION SECURITY, DATA PROTECTION, INTELLECTUAL PROPERTY PROTECTION AND OTHER SIGNIFICANT CONSIDERATIONS WHEN INVOLVING GENERATIVE AI TOOLS

It is crucial to exercise caution and adopt responsible practices when utilising generative AI tools. Researchers should be mindful of the following: security risks, legal aspects of data protection, intellectual property protection, and other considerations such as biases and inadvertent data linkages when employing generative AI tools.

The University is trialling the use of a Self Assessment Form to assist researchers to:

- consider potential security, legal and ethical issues; and
- undertake other significant considerations before using a generative AI tool.

Further details are presented below:

[1] As defined within the UK GDPR Article 4(1), as 'any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'.

[2] See the definition of special category data provided by the Independent Commissioners Office (ICO). You can only process special category data if you can meet one of the specific conditions in Article 9 of the UK GDPR.

[3] Any information disclosed by one legal entity to another and either a) identified as confidential before or at the time of disclosure, or b) which, by its nature or from the circumstances of its disclosure, should reasonably be presumed to be confidential. It includes commercially sensitive data, which is defined as any information which could damage commercial interests.

[4] Data belonging to or being provided by a person / organisation that does not belong or is not controlled by UEA (staff or students).

[5] Data that UEA deems essential for its success and that could cause UEA harm should it be published.

Information Security

i. UEA Approved Software:

Currently, UEA does not have a centrally supported generative AI tool that is officially recommended, supported, or for which staff training is provided. The University currently **prohibits** researchers to input any personal (including special category data), confidential, third party, or UEA business critical data / information / material into a generative AI tool, without prior approval from the Information Compliance Team (to ensure compliance with Data Protection legislation) and the relevant IT Team (to approve the security of the tool) following these steps**.**

Note that while many external generative AI tools offer basic functionalities for free, more advanced features often require a fee. There is no central UEA funding for this and any associated licenses.

ii. Cybersecurity: Generative AI tools may be susceptible to cyberattacks, potentially exposing sensitive information. Researchers must employ robust cybersecurity measures to protect data and systems, and where inputting personal (including special category data), confidential, third party, or UEA business critical data / information / material into a generative AI tool, will need to check with ITCS that these measures are in place before using the generative AI tool. **For further advice, contact the Information Security Team.**

Data Protection

A DPIA is required by law if you will input personal data where the processing of the personal data is likely to result in a high risk to the rights and freedoms of individual data subjects. You can complete UEA's DPIA Screening Form to determine if a DPIA is required. **For further advice, contact the Information Compliance Team.**

As stated in the Information Commissioner's Office (ICO) Guidance on 'How to Use AI and Personal Data Appropriately and Lawfully, "If you use AI, you need to assess the risks and implement appropriate technical and organisational measures to sufficiently mitigate them. You cannot realistically remove every risk, and data protection law does not require you to do so. You should: (i) carry out a Data Protection Impact Assessment (DPIA) to help identify and minimise the risk of non-compliance with data protection legislation that your use of AI poses, as well as preventing harms to individuals; and (ii) consult with different groups who may be affected by your use of generative AI to better understand the risks.

Researchers must also carefully consider the following data protection aspects associated with using personal data etc in generative AI tools:

i. Privacy: Unless clearly stated in the privacy policy of the generative AI tool provider, there is no guarantee that the data / information / material provided to a generative AI tool will remain confidential. Thus, when sharing any personal, confidential, third party, or UEA business critical data with a generative AI tool, especially data that is legally protected, for example copyrighted material, it is essential for researchers to review the privacy policy of the generative AI tool provider and ensure compliance with relevant data protection regulations. It is also important for researchers to examine the data control settings of the generative AI tool to ensure they are happy with the privacy settings. **For further advice, contact the Information Compliance Team.**

ii. Data Accuracy: The UK GDPR requires personal data to be accurate and kept up to date. If inaccurate or outdated personal data is used, it could result in harm to individuals when wrong decisions are made based on it. If a researcher does enter personal data into a generative AI tool, they must take all reasonable steps to make sure the personal data they are using in the generative AI tool is not "incorrect or misleading as to any matter of fact". **For further advice, contact the Information Compliance Team.**

iii. Transferring and / or Accessing Personal Data Outside of the UK: Researchers need to consider any personal data which might become subject to a 'restricted transfer'. This means any personal data which is transferred outside of the EEA is now unlawful without a transfer risk assessment and suitable documentation in place. Researchers therefore must discuss entering any personal data into a generative AI tool outside of the UK with the University's Information Compliance Team. **For further advice, contact the Information Compliance Team.**

iv. Data Security: Personal data must be processed in a manner that ensures appropriate levels of security against its unauthorised or unlawful processing, accidental loss, destruction or damage. Researchers should always check the security settings of the generative AI tool and choose to have the data / information /material given to the tool automatically erased after a set period of time, or manually delete the data themselves, otherwise the generative AI tool may keep it indefinitely. **For further advice, contact the Information Compliance Team.**

v. Automated Individual Decision-making (such as algorithms) Including Profiling: The UK GDPR has provisions on:

- Automated individual decision-making (making a decision solely by automated means without any human involvement); and
- Profiling (automated processing of personal data to evaluate certain things about an individual). Profiling can be part of an automated decision-making process.

A key question for researchers to ask when using generative AI in the context of decision-making is whether the decision is wholly automated or not. The UK GDPR prohibits solely automated decisions that have legal or similarly significant effects on individuals, with certain exceptions. If there is meaningful human involvement (which varies depending on the circumstances) in the AI lifecycle, the decision-making might be deemed to be 'AI assisted' rather than wholly automated. This will be one of the UK GDPR provisions that is considered before a generative AI tool will be approved by ITCS and this will be assessed as part of the DPIA. **For further advice, contact the Information Compliance Team.**

Intellectual Property Protection

There may be intellectual property restrictions or regulations pertaining to the involvement of the generative AI tool in the research project.

i. Copyright and Patents: Researchers must carefully consider the copyright implications of involving a generative AI tool in their research. Copyright permissions when entering third party Intellectual Property in generative AI tools including copyright licensed under Creative Commons, and when entering material for which UEA holds the intellectual property need particular attention. In essence,

the researcher needs to ensure the necessary permissions are in place for them to input this information legally into the generative AI tool.

Generative AI's impact on copyright also extends to the AI-generated content. Researchers need to be aware that if they use a generative AI tool to create something, they will need to establish whose intellectual property is the outputted data: the researcher's, the generative AI tool supplier's, or the owner of the data / information / material that was inputted to the generative AI tool?

There are also considerations when using generative AI tools for any research that may be patentable. The University recommends no patentable research is input into a generative AI tool.

Further details are available in the University's *Generative AI Policy for Research and Innovation*. **For further advice regarding third party copyrighted material, contact the Information Compliance Team. For further advice regarding UEA copyrighted material, contact the Intellectual Property Team.**

Other Signficant Considerations

There are a number of other significant aspects of using a generative AI tool, that researchers must consider:

i. Biases and Discrimination: These can arise during the different stages of a research project using generative AI, for example, data collection, model development and deployment, leading to potentially harmful outcomes. It is important that researchers address sampling bias issues that could arise as the research project using generative AI progresses, as the underlying algorithm might be making decisions that are biased or discriminatory (see 'Training of Models' section below) and in breach of the broad fairness requirements of the UK GDPR.

ii. Inadvertent Data Linkages: It is also important for the researcher to consider what other data / information / material the generative AI tool they are planning to use, has access to and has accessed in the past. This can be very challenging, as researchers are unlikely to know exactly what other data has been put into the AI tool in the past. Generative AI tools can be potentially manipulative, resulting in researchers providing data / information / material they did not anticipate. For example, generative AI tools may create nebulous connections between data points, enabling re-identification of anonymous data records, known as a 'linkage attack'. Researchers need to be aware of a generative AI tool's ability to create indirect identifiers within a dataset (and that the data / information / material they put in may be used to create linkages in the future). It is therefore important to minimise the information that is entered as anonymised data into a generative AI tool so that only the anonymised data necessary for a particular purpose is released, and thus the risk of re-identification is sufficiently reduced.

iii. Training of AI Models: As generative AI models evolve; generative AI providers are adapting their terms and conditions to allow them to retain data / information / material entered into their products for model training purposes. Researchers must scrutinise the terms and conditions of the generative AI provider carefully to understand their rights and limitations before entering material into a generative AI tool. Also, when data is being used to develop and train a generative AI's model, if the data is flawed, impartial or biased in any way (this can be because the data is incorrect, excludes

certain groups, or was collected in bad faith) then the resulting AI tool will be biased as well. A generative AI tool is only as unbiased as the data and people training the AI model.

HUMAN PARTICIPANT INVOLVEMENT IN UEA RESEARCH INVOLVING GENERATIVE AI TOOLS

Informed Consent to Participate

When primary data is collected, participants should know what data will be collected and what will happen to it, in order for them to give their informed consent for this. UREC therefore requires that researchers notify participants of the use and context of AI in the PIS, so that participants can decide as to whether they wish to participate in the research. The PIS needs to explain in lay terms the generative AI's use in the study (for example, how it will be used, what tool will be used, who will be using the tool, and explain what data – if any – will be inputted into the tool). This explanation should include a description of what the generative AI tool will do with their data. Researchers must inform participants in the PIS **if** their data will be inputted into a generative AI tool and cannot be removed from it. For example, in some instances the generative AI tool may ingest the data into it, and in such cases, it would not be possible to remove their data from the tool.

UREC have provided a template (available in section 7. under the 'Guidance' button on the Homepage of Ethics Monitor) of the PIS and Consent Form to use for projects involving generative AI tools. Please ensure that the additional text shown in green is edited and covered in the PIS and Consent Form as required.

As many generative AI tools are still experimental and the results unpredictable, it can make the informed consent process difficult. The researcher must check the terms and conditions of the generative AI tool to be aware of the systems' intended purpose, capabilities, and limitations to ensure they are informed to make decisions and understand potential impacts on the inputted data / information / material before it is collected from human participants or used.

Some researchers may collect primary data for a current project that does not use generative AI. However, if there is potential for someone to take that data and reuse it in a generative AI tool in a later project, then the primary user of the data should consider and raise the potential reuse in a generative AI tool at the point of seeking the original consent from the participants, and make this clear in the terms and conditions for reusing the data in a generative AI tool.

When using secondary datasets,[6] researchers must check that they will not be infringing a copyright or license agreement (see above), when involving a generative AI tool. For example, some publicly available datasets may state that the data must not be reused in an AI tool, and publicly available datasets may also contain personal data that should not be entered into a generative AI tool. Where data has been sourced from a repository or data centre, researchers must follow any guidelines provided by that repository / centre on how data must be used and acknowledged when using this data with a generative AI tool. Researchers using third party material whose terms for reuse are governed by permissions given by research participants must make sure that the reuse of the material

[6] Secondary data is data that has already been collected by someone other than the primary user of the data, e.g., a researcher who is analysing survey data that has been collected by another researcher or organisation.

is in line with the original consent given by the participants, before using this material with a generative AI tool. **For further advice contact the UEA Ethics Committee Chairs**.

OTHER ETHICAL AND SOCIETAL RISKS OF INVOLVING GENERATIVE AI TOOLS IN UEA RESEARCH

In line with the UK's Trusted Research guidance for protecting the University's important assets and ultimately its global reputation, consideration needs to be given to the potential risk of a generative AI tool using the data for purposes that were not originally intended. As stated by the UKRI, 'Trusted research' is a research and innovation sector term for protecting the UK's intellectual property, sensitive research, people and infrastructure from potential theft, manipulation and exploitation, including as a result of interference by hostile actors. A generative AI tool could be classed as a form of 'hostile actor' under certain circumstances, and it is vital for researchers to assess how they will mitigate the risk of AI development and future unintended usage of their data.

Ethical and societal risks of generative AI research can manifest at different stages of the research. Generative AI research has therefore moved the singular moment of ethics approval at UEA to a dynamic ethics review process, requiring multiple amendment requests from the applicant(s). The UREC has agreed that currently, the maximum length of ethics approval for a study involving a generative AI tool is 12 months in the first instance. Thereafter, ethics approval will be required every 12 months up to a maximum of five years from the project start date. An amendment ethics application can be submitted before the ethics approval date expires to obtain an extension of the ethics approval end date.