LSE Guidance on the use of Generative AI for research

This guidance applies to all LSE staff and students[1] undertaking research.

Notes:

i) Due to the pace of technological developments, this document is subject to review a) every 3 months and b) as and when substantial new technologies or features emerge that directly impact the use of generative AI in research.

ii) All academic research, irrespective of tools used, should adhere to the School's principles of "honesty, accountability, transparency, research rigour and good stewardship"

Definition

Generative AI refers to AI systems that create new content, predominantly text but also images, audio and video, based on users' natural language prompts. It should be thought of as a supportive tool or assistant, with researchers always in the driver's seat and accountable for what they produce.

Overview

As an institution committed to promoting innovation and impact, the LSE understands and welcomes the enormous value generative AI can bring, and the School actively supports and encourages its responsible use by staff and students. Generative AI is still in its infancy but is already transformative, enabling unprecedented productivity advantages for knowledge exchange activities, ideation, learning new concepts and skills, planning, feedback, analysis and accelerating knowledge discovery. As the scale and quality of AI technologies improve, and as researchers learn to make the best use of them, the potential for enhancing research will only grow. For now, the guidance on effectively using generative AI tools is a wide-ranging and continually-updated resource for LSE researchers, including an introduction to the technology, good practice advice and numerous illustrative examples.

As with all technologies, there are risks. The most important are those with legal, regulatory or financial consequences, particularly around deliberate or inadvertent sharing of data with third parties. The School's Legal and Regulatory Guidance for using AI <link forthcoming> covers data governance and legal risks for all AI usage by staff including for administrative and education work. For research work specifically the main risks are as follows:

Primary Risks

Data Privacy and Security

Researchers should not share personal or sensitive data with 3rd party tools that do not provide assurances of privacy and security. Microsoft Copilot is available for staff and student use in the LSE, supported centrally, and is fully secure and private when logged in with an LSE email in the

[1] Where students are engaged in taught course provision (including BA/BSc/MA/MSc/MPhil dissertations and summative research projects, and PhD students taking 500-level courses with assessment components) they are subject to the School's guidelines on the use of AI in education, and subsidiary departmental policy/policies

Edge browser. No data is stored or shared with any third parties nor used to train models, so its use is strongly encouraged. Please read the full LSE guidelines on MS Copilot.

- If any kind of personally identifiable information (PII) is required for any research work, and MS Copilot isn't viable, and the researcher is unable to anonymise or pseudonymise (see the School's guidelines on data anonymisation and pseudonymisation), a Data Protection Impact Assessment (DPIA) must be submitted, as with any research project involving PII processing. The LSE AI Legal Checklist <link forthcoming> will also be required if using a tool other than MS Copilot.

Researchers should read and understand the Terms and Conditions (T&C) of any tool used. Wherever possible researchers should use Microsoft Copilot because it assures privacy and security. While the Chat GPT Team Plan provides enhanced security and privacy compared to the standard commercial licenses and is permitted within the LSE, the terms and conditions make clear that any data is temporarily stored in the US which risks breaching GDPR and therefore personal or sensitive data should never be uploaded to it.

Copyright and Intellectual Property

Researchers should avoid sharing copyrighted materials with third party AI providers. This may include licensed and open access works held by the LSE Library, as some publishers have policies against uploading content where there is a risk of the data being used to train future commercial AI models. Again, Microsoft Copilot when logged in with an LSE email address is the recommended tool for staff given that content is not retained to be used in training models.

Acknowledgement, Authorship and Accountability

Researchers should consult publishers' and funders' guidelines in advance as they can differ on policies regarding acceptable use of generative AI, for example requiring acknowledgement or audit trail of how it was used, or questions of confidentiality of submitted work or reviewer feedback. Here's a list of links to major publishers' policies relating to generative AI:

Cambridge University Press
Elsevier
Nature
Sage
Taylor & Francis
Wiley

The core themes in academic publishers' policies are as follows:

Category	Detail
Accuracy	Factual verification of any 'information' output by AI.
Acknowledgement	AI use must be acknowledged in the manuscript where used substantively.
Authorship	Author is always fully accountable; AI cannot be used or listed as an author.
Bias	Evaluate and critically reflect on inherent biases of AI outputs.
Copyright	Check AI outputs for inadvertent plagiarism (e.g. via Turnitin)
Privacy	Confidential information (including paper content or comms relating to submissions) must not be uploaded to AI tools by reviewers / editors.

Most publishers require acknowledgement of substantial generative AI use. The EU (2024: 6) cites the following uses as potentially constituting 'substantial' contribution: "interpreting data

analysis, carrying out a literature review, identifying research gaps, formulating research aims, developing hypotheses". But Elsevier requires disclosure even when used for readability improvements, so always check the publisher's policies. Acknowledgements may include citing the tool used, prompts, outputs and how outputs were used and/or adapted.

Note for taught undergraduate and postgraduate students: as dissertations are assessments that form part of a degree, the Guidance on the use of generative AI in education applies; students should consult their home departments to check on any specific policies relating to AI use.

UKRI's policy on the use of AI in grant application preparation and assessment takes a different approach on grant application content, including explicitly instructing applicants not to cite the use of AI. Instead the focus is on applicants' responsibility of any content, alongside strict rules regarding application assessment by reviewers. Here's a summary of the main themes from the UKRI policy:

Category	Detail
Integrity	Uphold values of honesty, rigour, transparency, and open communication.
Confidentiality	Avoid sharing confidential and personal data with AI tools that do not provide privacy guarantees, unless consent has been granted.
Applicant Responsibilities	Ensure AI-generated content is not falsified, fabricated, plagiarised, or misrepresented, and look to mitigate potential AI-generated biases.
Assessor Guidelines	Must not use AI tools for assessment; should not speculate on AI use in applications.
Disclosure	Applications should not cite the use of AI tools in developing content.
Compliance	All applications must comply with intellectual property and data protection laws.
Misconduct Consequences	UKRI may reject applications, prevent future submissions, or reclaim funding for upheld misconduct allegations.

Secondary Risks

Informed Consent

If your research involves human participants and you intend to use generative AI in any capacity on the data the participants produce, you should explain this in advance with them as part of the usual informed consent process.

Bias and Limitations

Generative AI outputs can:

- be **unreliable and inaccurate.**
    - Generative AI models can generate plausible-sounding but factually incorrect information, commonly referred to as 'hallucinations'. Generally, generative AI as it currently stands should not be used as any kind of reliable information source, and human verification steps should be incorporated wherever accuracy is important. The guidance on effectively using generative AI for research includes a section on Prompting and mitigating hallucinations, with suggestions on how to make the verification process simpler

- **replicate human biases from the underlying training data.**
  - Just like when engaging with any texts, researchers should critically evaluate generative AI outputs for potential biases and avoid perpetuating or amplifying them in their work. Where relevant, researchers should acknowledge these limitations.

- **prohibit reproducibility, due to their inherently probabilistic nature**
  - Generative AI models can generate different outputs even for the same prompts. Where generative AI is substantively used for research findings (e.g. classifying or coding texts), researchers should document their use of the tool including prompts and pre- and post-processing, to ensure transparency, but acknowledge the limitations for perfect reproducibility.

- **bring reputational risk if reproduced without due diligence**
  - Researchers are always fully accountable for all their content including anything produced with the support of generative AI, so careful review is needed prior to sharing any outputs.

Environmental Impact

Generative AI technologies consume substantial **computing resources** which increase carbon emissions. 16 Chat GPT queries are equivalent to boiling a kettle (see "Gen AI's environmental ledger: a closer look at the carbon footprint of Chat GPT" for more examples)