Guidance for the Responsible Use of Generative AI in Research

Scope & Purpose

With the evolution of Artificial Intelligence (AI) and Machine Learning (ML), there are rapidly increasing opportunities for the application of these digital assistance technologies in the conduct and dissemination of research. The purpose of this Guide is to outline some key consideration in support of the appropriate and responsible use of digital assistance tools in relation to research, in a manner that complies with our ethical, regulatory, and legal obligations. It also provides links to more detailed **resources for researchers and higher degree candidates, peer reviewers, grant assessors and thesis examiners**.

Use Cases and Opportunities for Generative AI in the Conduct of Research

Cornell University released a detailed report in Dec 2023 , entitled "Generative AI in Academic Research: Perspectives & Cultural Norms" [1] that identifies the key uses of generative AI in research across diverse research domains as:

- *To **automate** tasks that are tedious, time-consuming, or require large amounts of data, such as literature review, data collection, data augmentation, data analysis, and data visualization.*
- *To **augment** human capabilities and creativity, such as hypothesis generation, problem formulation, solution design, and content creation.*
- *To **disseminate** research findings and insights in various formats and languages, such as summaries, reports, presentations, podcasts, and videos.*
- *To **translate** research outputs into practical applications and products, such as prototypes, simulations, software, and patents.* [1]

That is, there is potential for Generative AI tools to increase productivity and efficiency by supporting high volume or complex analysis (of non-sensitive data), synthesis and design processes, including development of high-level outlines and structures for publications, and to stimulate creative ideas by providing new insights and perspectives on existing 'grey' and academic literature.

Critically, the Cornell report[1] also highlights the importance of using Gen AI in a responsible and beneficial way, and highlights the ethical, legal, and social implications of using Gen AI in research, such as data quality, human oversight, privacy, trust, and compliance.

Another useful resource is a recorded 2023 panel discussion that was hosted by the Centre for Research in Assessment and Digital Learning (CRADLE) and TEQSA: Generative AI: What do researchers need to know?

A number of implementations of AI in research at Griffith are identified on the Griffith's **AI Hub**.

Research and Commercialisation Activities

Generative AI can create novel content, including text, video, graphical and audio outputs, and software code, and may have broad applications in the conduct of research. Researchers must keep in mind both the opportunities and limitations of engaging with these tools.

Legal Services at Griffith produced a **fact sheet that outlines some key legal issues** to be aware of when considering the use of Gen AI in relation to research and commercialisation. Key considerations include:

- Ensure that you do not enter confidential information or personal information (of yourself or others) into publicly available Generative AI systems, such as Chat GPT.
- Recognise that the content used to train Generative AI might infringe on the IP rights of third parties or privacy laws.
- Recognise that using Generative AI to undertake research may produce incorrect, biased, misleading, outdated, discriminatory or defamatory outputs.
- Be aware that outputs of Generative AI might not be protected by current IP laws.

For the avoidance of doubt, the use of digital assistance tools in a manner that falsifies or fabricates or otherwise misrepresents data is a breach of the University's Responsible Conduct of Research Policy and the Australian Code for the Responsible Conduct of Research (the Code).

Data Inaccuracies and Bias

Generative models can propagate inherent biases that emerge from the model's design, training data and/or accumulated 'insights'. Many examples have also been identified of Gen AI errors in data representation, interpretation, authenticity, and attribution. Societal bias, structural racism and overt discrimination of underrepresented and marginalised groups can be amplified when using Gen AI tools. Further, there growing risks of deliberate misinformation being seeded (see for example, the Cyber Security CRC's 2023 report Poison the well - AI, data integrity and emerging cyber threats [2]).

Researchers should carefully assess generative AI system's outputs for unconscious or conscious bias, accuracy, relevance, and veracity. Researchers must take deliberate steps to identify, critically interrogate, and mitigate data inaccuracies and biases that may emerge though the use of Gen AI interfaces and tools.

Production of Research Writing, Visual, Audio and Software content

There is growing acceptance of use of Large Language Models (LLMs) and Generative AI in providing informative summaries of topics. While noting that the accuracy of these summaries cannot be assumed, they can serve as a mechanism to initiate deeper review and analysis of the area.

Gen AI can create novel text, images, audio and software code in response to questions or prompts, can change the style in which provided pre-existing content is expressed, or can produce summaries or expansions of outputs. Other tools offer rephrasing and rewriting, producing entirely new wordings of existing material. There are also digital assistance tools that can automatically generate images or alter images in various ways, some of which can be misleading.

University guidance on **preparing a Higher Degree by Research (HDR) thesis** clearly outlines the extent of professional assistance that can be provided when writing or editing a research thesis. This is limited to proofreading and copy-editing services.

Griffith University **researchers and HDR candidates are each responsible** for ensuring that the material provided in their research outputs, including theses, is their own work or, if derived from other sources, the use of those sources is transparent and appropriately acknowledged and cited.

Authorship & Acknowledgement

Authorship guidance under the Code and Griffith Research Integrity Resource Sheet #4 – Responsible Research Outputs each set out **Publication Ethics and Authorship expectations**. In broad terms, a research

output is attributed when a researcher has made a significant intellectual or scholarly contribution to a research output and is willing to take responsibility for the contribution.

All named authors must consent to being named and must be able to ensure the accuracy and integrity of the reported research. **Digital assistance tools cannot be named as authors**, as they are unable to provide consent or confirm the accuracy and integrity of the research output.

In order to quote the output of Gen AI or LLMs, we recommend that:

- Authors undertake the necessary due diligence to establish the veracity of the quoted material, where appropriate.
- A plagiarism check be completed, as LLM output can be closely paraphrase other sources.
- A footnote be used instead of an in-text citation. The footnote should state the tool used, the version (if available), the date on which it was used, and the exact text used as the prompt.
- For coursework subjects, the University has published advice that sets out, as for research writing, that all use of LLMs in work submitted for assessment must be fully acknowledged.

Griffith Library Services have published guidance on **How to Cite Generative AI Tools**.

Peer Reviewers, Grant Assessors and Thesis Examiners

Gen AI is not to be used as a peer review tool. Peer reviewers, assessors and examiners must not to use AI as part of their assessment activities (e.g. for manuscripts, HDR theses, grant applications, ethics proposals, or thesis examiners reports). The Australian Research Council has issued a statement on the **confidentiality obligations of assessors**, reminding all that under the Code, these activities must be conducted in a way that is fair, rigorous and timely and maintains the confidentiality of the content.

Academic Integrity

Separate guidance has been developed for students to consider how and in what instances AI or LLM tools might be used for learning, keeping in mind both the opportunities and limitations of engaging with these tools: **Why Academic Integrity Matters**. Further guidance for staff is also available at the **AI Hub**

Links to key external reports referred to in these Guidelines:

1. "Generative AI in Academic Research: Perspectives and Cultural Norms", Cornell University Task Force, 15 Dec 2023 Accessed 28 March 2024 at https://it.cornell.edu/sites/default/files/itc-drupal10-files/Generative%20AI%20in%20Research_%20Cornell%20Task%20Force%20Report-Dec2023.pdf
2. "Poisoning the Well- AI, Data Integrity and Emerging Cyber Threats" Rachael Falk & Anne-Louise Brown, 30 Oct 2023, Accessed 28 March 2024 via https://cybersecuritycrc.org.au/poison-well-ai-data-integrity-and-emerging-cyber-threats