Guidance for the Responsible Use of AI in Research

The following document provides Guidance for the Responsible use of AI in Research, with this one page Do/Don't checklist as a **foreword:**

Quick Reference Do / Don't Checklist for Researchers & HDR Candidates

✅ DO

- Consciously explore opportunities for the ethical and responsible use of AI tools in your research.
- Researchers (including HDR Candidates) retain full responsibility for the accuracy, integrity and originality of all research that they conduct and outputs that they produce.
- Use AI to support, *not replace*, scholarship and original thought.
- Consistently apply human oversight and critically review all AI-assisted outputs.
- Verify accuracy and bias, and complete plagiarism/similarity checks where appropriate
- Protect data and information confidentiality, privacy and security. Assess risks carefully before using AI tools, especially for qualitative data analyses (e.g. interviews, surveys).
- Be aware of any restriction on the use of AI from research funders, partners and the University.
- Acknowledge AI use transparently in publications, submissions and theses (tool, version, date, and prompt(s) where required).
- Acknowledge all assistance and ensure the work reflects your own intellectual contribution.
- Institutional frameworks, policies and approvals for AI use exist to guide and protect your reputation and assure the integrity of you research – comply with them.

❌ DON'T

- Do not use AI for peer review, assessment or examination activities, or where confidentiality obligations apply.
- Do not assume AI outputs are correct; hallucinations, bias and errors can occur.
- Do not obscure AI use. Failure to disclose assistance is a breach of research integrity standards.
- Do not accept AI analyses or summaries that may fabricate, falsify or misrepresent data.
- Do not enter confidential, personal, sensitive or proprietary information into public or unapproved AI tools.
- Do not overlook IP and legal risks. AI outputs may incorporate third-party IP and/or introduce your own or others IP into the public domain.
- Do not list AI tools as authors or imply authorship; AI cannot consent or take responsibility.

Guidance for the Responsible Use of AI in Research

1. Scope & Purpose

As artificial intelligence (AI) continues to evolve, its potential applications in the conduct and dissemination of research are increasingly recognised. This Guide outlines key considerations to support the responsible use of digital assistance tools in research, in alignment with ethical, regulatory and legal obligations. It also directs readers to further resources **for researchers and higher degree candidates, as well as peer reviewers, grant assessors and thesis examiners.**

2. Griffith Context

Griffith University is committed to the highest standards of research integrity and to the conduct of ethical and responsible research. The University is also committed to capacity building and enabling high-quality research practices and outcomes. When used responsibly, AI tools can reduce burdens from repetitive and administrative tasks and enable greater focus on research and innovation – our researchers are encouraged to adopt secure AI tools for these purposes where it is ethical, safe and appropriate.

Under the *University's Artificial Intelligence Framework*, AI may be used at Griffith to enhance administrative and operational processes, support teaching and learning, improve staff and student experiences, and enable research and research management. While AI presents significant opportunities, it also introduces risks that require careful management. The Framework outlines requirements for proposing or modifying AI-based solutions, to ensure appropriate endorsements and approvals are secured before implementation.

Researcher must be alert to the risk that information provided to generative AI tools may enter the public domain or be accessible to third parties, and AI-generated outputs may contain inaccuracies or inadvertently incorporate the intellectual property of others.

3. Use Cases and Considerations for AI in the Conduct of Research

3.1 Generative AI

Generative AI is a type of artificial intelligence that generates new content—including text, images, video, audio, and code—by recognising patterns in existingg large datasets. Cornell University released a detailed report in Dec 2023 : *Generative AI in Academic Research: Perspectives & Cultural Norms* [1] identifying the key uses across diverse research domains as to:

- ***automate*** *tasks that are tedious, time-consuming, or require large amounts of data.*
- ***augment*** *human capabilities and creativity, such as hypothesis generation and content creation.*
- ***disseminate*** *findings and insights, such as summaries, presentations, podcasts, and videos.*
- ***translate*** *research outputs into practical applications and products and patents.*

Gen AI tools can greatly boost productivity and efficiency in complex or high-volume tasks such as analysis, synthesis, and design. They also help generate creative ideas and offer fresh perspectives on academic and 'grey' literature. However, Large Language Model (LLM) based tools can generate factually incorrect, logically inconsistent, or entirely fabricated outputs or "hallucinations". 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.

3.2 Agentic AI

Agentic AI is a collective term for AI systems that can plan, make decisions, and take actions toward defined goals with minimal human intervention. Agentic AI combines reasoning, planning, and execution capabilities to handle complex workflows in research. In Sept 2025, Elsevier released an organisation-level report *Agentic AI in academia: How to adopt for research, learning, and innovation* [2].

Use cases are seemingly boundless and constantly evolving, which is both exciting and risky. Henc the importance of keeping ethics and integrity at front of mind. Best practice considerations include:

- ***Human-in-the-Loop Design***: Maintain oversight to prevent errors and bias.
- ***Start Small, Scale Fast***: Pilot in low-risk areas before expanding.
- ***Continuous Monitoring***: Validate outputs and ensure ethical compliance.

4. Key Considerations for the Responsible Use of AI in Research

4.1 Research Ethics and Integrity

GenAI can create novel text, video, graphical and audio content and software code, while Agentic AI can serve as a virtual research assistant, planning, making decisions and progressing actions**. Researchers are ultimately responsible** for any potential ethical or integrity breaches during the conduct of the work or the outputs that are generated through the employment of these tools.

For the avoidance of doubt, the use of any 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).

4.2 IP and Legal Obligations

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

- Do not enter confidential information or personal information (of yourself or others) into publicly available AI systems, such as Chat GPT or unlicenced versions of M365 CoPilot.
- Recognise that content used to train AI might infringe on the IP rights of third parties or privacy laws.
- Know that AI outputs may be inaccurate, biased, misleading, outdated, discriminatory or defamatory
- Be aware that outputs of AI systems might not be protected by current IP laws.

4.3 Data and Information Privacy and Security

Before using of AI tools, researchers must exercise due diligence in relation to data and information privacy and security, including in relation analyses of qualitative analysis of survey or interview data. AI systems may infer personally identifiable information (PII) from within or beyond a data set: they are are frequently trained on data scraped from the internet that may include personal, visual or otherwise sensitive information. The use of AI tools may reduce transparency and accountability by obscuring how personal data are collected, processed and reused.

Privacy risks also arise from the large-scale, and often non-consensual, collection of personal data used to train AI models, creating potential for biased interpretations, data breaches, surveillance and inadvertent disclosure of sensitive information.

4.4 Data Inaccuracies and Bias

AI tools may perpetuate underlying biases that originate from AI model design, the selection of training data, or the synthesis of accumulated insights. Many instances have been identified of 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 AI system 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 interfaces and tools.

4.5 Production of Written, Visual, Audio and Software content

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

Researchers must recognise that the accuracy of AI generated/developed content cannot be assumed and must be validated. Nonetheless, such content can serve as a mechanism to initiate deeper review and analysis of the topic/field of interest. In each case, the use of AI assistance must be acknowledged when releasing/publishing/sharing the final product.

The use of AI in development of any material that is submitted for assessment must be fully acknowledged. University guidance on **preparing a Higher Degree by Research (HDR) thesis** outlines the extent of professional assistance permitted when writing or editing a research thesis. This is limited to proofreading and copy-editing services. Such supports, whether human or AI, must be acknowledged.

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.

4.6 Authorship

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, it is expected that:

- Authors undertake the necessary due diligence to establish the veracity of the material
- A plagiarism/similarity check be completed.
- 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.

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

4.7 Peer Reviewers, Grant Assessors and Thesis Examiners

AI tools are not permitted 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). Major funding bodies, including the Australian Research Council have issued statements 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. The University has published guidelines for the examination of a thesis/exegesis, requesting that examiners **do not** use generative AI in their assessment process.

5. AI Guidance Review Cycle

As the use of AI becomes increasingly normalised, the authors anticipate that this guide will require periodic revision to reflect developments in AI-enabled research, and welcome feedback on its relevance and currency. Feedback and suggested updates are invited by email to red@griffith.edu.au.

Acknowledgements

The author acknowledges the use of M365 Copilot (GPT-5) to support drafting and editing these Guidelines: *Microsoft (2026). Microsoft 365 Copilot (Mar 8 version) [Large language model]. https://copilot.microsoft.com/*

Key References & Resources

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. "Agentic AI in academia: How to adopt for research, learning, and innovation", Elsevier Report, September 2025, Accessed 16 Jan 2026 at https://assets.ctfassets.net/o78em1y1w4i4/4CJtu7gXT5EvhinZdS0OfD/663ece72f347df49d7/eaea7066072f95/eaea7066072f95/CS-1372_Scopus_AI_Agentic_AI_Buyer_Guide_Sept_2025_V7.pdf
3. "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
4. 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?
5. Examples of the rapidly expanding, diverse range of AI in Research uses at Griffith are profiled by the eResearch team at Griffith and through annual AI in Research Symposia.