Session 5 – Swimming between the flags: Avoiding AI rips and snags (a workshop)

Dan Bendrups


Chat GPT
Nov 2022

ACGR Good Practice Guidelines for

Generative Artificial Intelligence Use in Graduate Research Training

December, 2024

The Guidelines…do not address the use of GenAI in all research or higher education activities.

Similarly, the Guidelines are not a comprehensive overview of AI technologies.

Accordingly, the advice should be supplemented with other frameworks that support research, teaching, learning, and other contexts within and beyond the institution.

🎓 Monash University – Graduate research AI guidance

- Requires **alignment with the Australian Code for Responsible Conduct of Research**
- Emphasises **disclosure of AI use in theses**
- Signals that examiners are informed about AI use expectations

➡️ Key idea: AI is allowed, but must be **transparent and accountable**

Monash U...

🎓 Australian National University – Best practice guide

- Prohibits presenting AI output as your own work
- Warns against using AI as a **substitute for research and writing**
- Requires **verification of AI-generated content against sources**

➡️ Key idea: AI is a **support tool, not a replacement for scholarship**

Australia...

🎓 University of New South Wales – Responsible AI use framework

- Introduces **“levels of AI assistance”** (none → editing → planning, etc.)
- Requires:
  - Maintaining academic integrity
  - **Not claiming AI-generated text as your own**
  - **Checking accuracy of outputs (“hallucinations”)**
  - Protecting **privacy and sensitive data**

➡️ Key idea: AI use should be **explicitly bounded and risk-aware**

UNSW Sites

🎓 Federation University Australia – AI policy

- Requires **formal acknowledgement of AI use** in research outputs
- Specifies including:
  - Tool name
  - Version
  - Date of use

➡️ Key idea: Treat AI like a **citable research tool**

Federation Unive...

1. Transparency

- Declare and acknowledge AI use (e.g. Monash, Federation)

2. Integrity

- Do not present AI output as original work (ANU, UNSW)

3. Verification

- Check accuracy and reliability of outputs (ANU, UNSW, TEQSA)

4. Accountability

- Researchers remain responsible for all outputs, even if AI-assisted

5. Data & privacy protection

- Avoid sensitive or proprietary data in AI tools (UNSW, TEQSA)

6. Methodological rigour

- Ensure appropriate datasets and monitor outputs (Newcastle)

7. Limits on use

- Some activities restricted (e.g. peer review at Macquarie)

LTU guidance also includes:

- Must be justified
- Must not be used on others
- Must be ‘proportional’
- Only humans can ‘own’ outputs

So where are we now?

- Planning vs. panicking
- Prescribing vs. proscribing
- Not *really* addressing the one thing graduate researchers ask for: specific rules

Guidance? Well, is it a discipline thing?

Acceptable and unacceptable cases

A sample of acceptable use cases of GenAI within CEUR-WS publications are exemplified here, alongside notable exceptions where their employment would be deemed unacceptable.

Use case	Acceptable	Unacceptable
Text Creation	While GenAI can assist with writer's block or retrieving definitions, its use should be contingent upon human critical thinking and judgment to ensure accuracy and originality.	Using GenAI to generate new text, such as paragraphs, or even entire sections of a paper is ethically unacceptable. Academic writing must be original and attributed to human authors.
Text Translation	This involves using GenAI to translate text from another language into English or vice versa.	Employing GenAI to translate a previously published work into English, without subsequent editorial refinement, raises ethical concerns about self-plagiarism.
Sentence Polishing	GenAI can be used to identify and correct grammatical errors, typos, and other writing mistakes. This helps improve the clarity and professionalism of your writing.	GenAI's ability to process large textual data allows it to seemingly process minor sentences to whole paragraphs. Refining whole paragraphs without human critical thinking and judgment can perpetuate biases from its training data and erode the author's narrative voice.
Image Creation	GenAI can create images based on text prompts. This is acceptable only when the paper's core topic is about automatic image generation.	Employing GenAI to create visual aids, such as diagrams, charts, and illustrations is unacceptable. The generated output is often nonsensical, and hinder the quality, trustworthiness, and value of scientific papers. See article on Science Integrity Digest.
Rephrasing	GenAI can help you rephrase sentences or paragraphs to improve clarity, conciseness, or style. This can be helpful for expressing ideas in a more sophisticated way.	As stated for sentence polishing, rephrasing without human oversight can perpetuate biases from its training data and erode the author's narrative voice.

https://ceur-ws.org/GenAI/Policy.html

Or is it a publishing industry thing?

The standard Springer Nature policy

Nature Communications follows the Springer Nature AI policy without any modifications:

1. **AI can't be an author.** LLMs and generative AI tools don't meet authorship criteria.
2. **AI use in manuscript preparation must be disclosed in Methods.** Name the tool, version, and describe how it was used.
3. **AI-generated images are banned.** No figures, graphical abstracts, or visual content from generative AI.
4. **Copy editing is exempt.** Standard grammar and spelling tools don't require disclosure.
5. **Authors bear full responsibility** for all content, including AI-assisted sections.

https://manusights.com/blog/nature-communications-ai-policy

The problem:

- Knowing what to do requires judgement
- Judgement is a product of development
- Development is a product of experience
- There is no short-cut to developing judgement

Our approach at LTU:

- Provide the broad guidelines
- Specify exclusions (including arbitrary ones)
- Targeted reflection (within discipline cohorts) on the tolerances for inclusion, noting that the expectations of graduate research are different to research in general because of the ‘training’ component

Schön's *Reflective Practitioner*:

- Reflection-on-action
- Reflection-in-action
- Case examples as 'safe', externalised surrogates
- Cases devised to suit specific cohorts

For your consideration:

- 3x cases, designed specifically for Deans/ PVCs/ Directors of graduate research
- Increasing complexity (3rd one is the fun one)
  - 5 minutes reading/reflection on all scenarios
  - Group decision about which one to discuss
  - 15 minutes group discussion
  - 15 minutes room discussion

Observations:

- The issues raised *may* be about AI, but *may* also about other things (developing judgement, attention to detail, legislative compliance, etc.)
- Concerns about writing *may* actually about assurance of learning.
- Existing checks and balances for good research governance *may* already equipped to deal with the challenges presented by GenAI in graduate research.