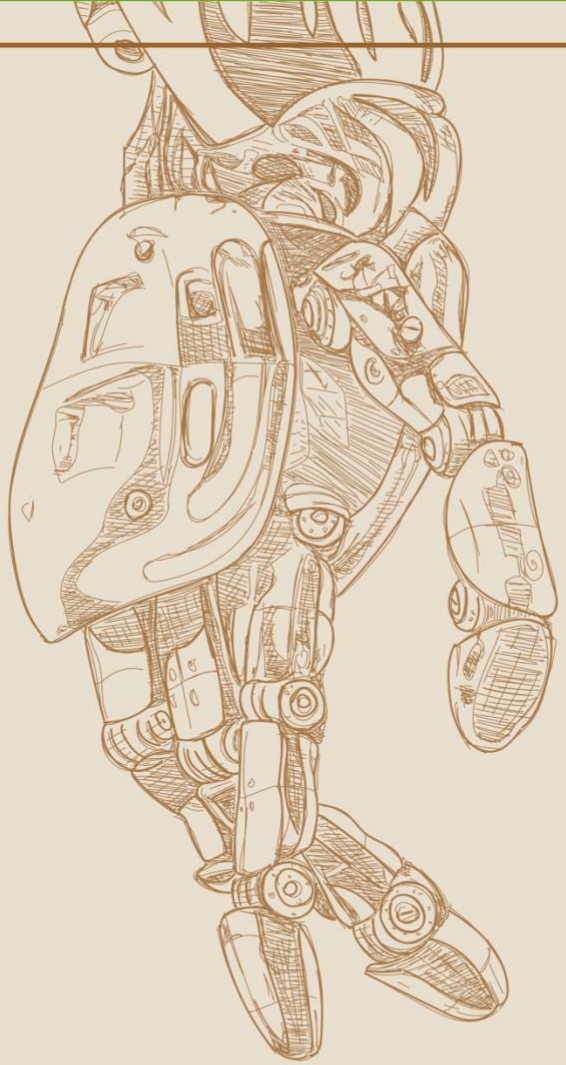
CLM Faculty
TEACHING AND LEARNING UNIT



Researcher Guidelines for the Ethical Use of Generative AI Tools (Draft)

September 2024

COLTStudentSupport@wits.ac.za

COLT
CLM Online Learning and Teaching

CLM Faculty: Researcher Guidelines for the Ethical Use of Generative AI Tools

Version: September 2024

Adapted from: Centre for Innovation, Learning and Teaching (2023). Researchers' Guide: Ethical use of generative AI for research purposes. University of Cape Town. CC-BY.

1. Introduction

The aim of this faculty guideline document is to suggest responses for how staff and student researchers could think about the ethical use of Generative Artificial Intelligence (GAI) tools for research purposes. This guideline document should be read in conjunction with the companion guidelines:

- CLM Staff Guidelines for Learning and Teaching with Generative AI Tools
- CLM Staff Guidelines for Generative AI Tools with Assessment
- CLM Guidelines for Effective Prompting with Generative AI Tools
- CLM Student Guidelines for Generative AI Tools in Higher Education

The guide should also be read in conjunction with other Wits and CLM policies and frameworks, and guidelines, including but not limited to:

- Wits Research Integrity Policy and Wits Research Integrity Procedures
- Wits Student Academic Misconduct Policy
- Wits Staff Plagiarism Policy
- Wits Framework for Academic Integrity
- Guiding principles on the outsourcing of editing and statistical work by postgraduate students in CLM

These policy documents stress how important it is for students / staff / researchers to uphold and adhere to the principles and values of integrity and ethics. These documents do not make specific mention of GAI (yet), but the current value system should be applied to the use of GAI. Key concepts from these documents should be kept in mind when using GAI during research:

1. The core work of research – such as the design of research instruments, the analysis of data, the generation of statistics, the writing up of the literature review – cannot be outsourced and must be completed by the researcher (see CLM Guidelines on outsourcing).
2. The advancement of scholarship is associated with the development of thought and adding knowledge through discovery, as a result Wits policies and frameworks emphasise the principles of integrity and honesty.
3. The Wits Framework for Academic Integrity draws on the principles of integrity (ICAI, n.d.) from the International Center for Academic Integrity (ICAI) and call on all academic community members to be honest, trustworthy, respectful, responsible, and courageous.
4. The Wits Research Integrity Policy forefronts the need for intellectual honesty, researcher responsibility, and the protection of research participants and their data.

These concepts should be used as guiding principles when considering the use of GAI in research. This list is not exhaustive, and researchers should engage with institutional documents to ensure that they adhere to the rules of the University. Please note that when the use of a GAI tool contravenes any of the ethical or integrity principles laid out in institutional documents, they should not be used.

2. Using GAI Tools For Research

The rapid proliferation of GAI tools poses both challenges and opportunities for researchers. This guide offers suggestions for how these tools can be used productively and ethically for research processes and implications for research integrity. Examples of interest to researchers, student researchers, supervisors, and others are included. It is recommended that at the start of any research project, discussions are held with co-researchers and students / supervisors about what ethical GAI use would look like (drawing on the guidelines below). It should be clear to co-researchers and students whether GAI tools can be used and how they can be used for research and writing.

Balancing productivity and ethical concerns require careful consideration. The use of GAI in research remains a changing and uncertain space requiring those involved to develop shared understandings. It is noted that standard word processors already make use of GAI in spelling and grammar checkers. Therefore, attempting to ban any use of GAI in academic writing is implausible. In summary, the use cases guide when GAI can (and cannot) be used to support academic writing (Bekker, 2024):

Possible uses (to check first with research colleagues/supervisors/intended journal or publisher)	**Non-acceptable uses**
• Proofing tool: Fix spelling, grammar, tone and style issues. • Language editing: Rephrasing text, expanding/shortening sentences, reviewing citations. • Drafting consultant: shaping and improving ideas and writing based on conversations, counsel and correspondence (acting as an informal review by a peer who may or may not be an expert in the field). • Technical help: as would be available in other forms such as word processors, spell checkers, software programmes and human proofreaders and language editors.	• Authoring or co-authoring: Any process that outsources the thinking, reasoning and originality that goes into the research process. • Plagiarism: Presenting work as if it is your own when it has been generated by others or a tool and/or not attributing sources. • Reliance on data analysis: Outsourcing GAI tools to interpret results and draw conclusions.

Three principles underline the considerations for if/when to use GAI tools in academic writing (Bekker, 2024):

- Human authors should not contract out or **outsource thought** to (non-human) algorithms.
- Any researcher (human author) remains the **owner** of any submitted or published work and the author is responsible and accountable for any mistakes or implications of the

written work. Any mistakes cannot be blamed on a GAI tool if a GAI tool was used at some point during the research process.

- **Transparency** in the use of GAI tools in the form of when and how they are used (if at all). Refer to the 'CLM Staff Guidelines for Generative AI Tools with Assessment' for how to cite the use of a GAI tool using APA referencing. Authors may wish to provide a way for readers to access GAI interactions such as prompts used and responses received, such as providing a supplementary file or hyperlink to a repository (Bekker, 2024).

Bekker (2024) suggests the following possible tiers of GAI use in academic writing. Note that this guide further explains possibilities for Tiers 2-4 as **Tiers 1 and 5 are considered unsuitable for research writing**:

Tier	Possible use in academic writing	Associated benefits and risks
1 – Ban GAI use	N/A	• Benefits: Ensures human authorship and authenticity. • Risks: Cannot be guaranteed unless writing under 'exam' conditions – unlikely to apply to research projects. **This tier use is not recommended for research writing.**
2 – Proofing	After writing	• Benefits: Increases efficiency and may reduce proofing costs. • Risks: May subtly alter meanings or obscure intentions.
3 – Editing	During writing	• Benefits: Improves writing for readers. • Risks: Language may be bland or may foster reliance on tools for writing efficiencies.
4 – Drafting Consultant	Starting or during writing	• Benefits: An alternative to human partnerships, makes interpretive and improvement suggestions, checks for errors, etc. • Risks: High risk of introducing hallucinations and biases, loss of autonomy and reproducibility, possible opaque authorship.
5 – No limits	Throughout	• Benefits: speeds up academic writing, support for inexperienced researchers. • Risks: Loss of critical reasoning and outsourcing of thought. **This tier use is not recommended.**

3. GAI Tools: Functions and Limitations

This guide deals primarily with GAI tools such as ChatGPT, Gemini, Copilot, and Claude, which are based on large language models. This natural language styles of GAI invites one to have a conversation, however these are not human conversations nor is the information factually correct necessarily. The large language models respond to prompts by predicting a likely next word. By carefully crafting prompts one can direct what is generated so that it will be more useful. Please see the CLM Guidelines for Effective Prompting with Generative AI Tools for more information. When it comes to research, bespoke research GAI tools (see Section 5 and 6 in this guide) may better serve researchers looking for specific uses, rather than the general GAI tools listed above.

Providing greater context to the GAI tool will often improve the response. For example, it may be useful to include your role, your knowledge area, or any constraints on the output. In

ChatGPT for example, you can include this under “Customize ChatGPT” a function associated with your login (bottom left corner on the ChatGPT page using the website). These would be used each time you use ChatGPT. Otherwise include this information when you start your session to help make the results more relevant.

Knowing how GAI functions can help with crafting prompts and anticipating cases where it will be useful:

- **Language generation**: GAI models are trained on natural language text and can generate coherent and grammatically correct content. This is useful for tasks like summarising articles, changing styles and proofreading writing. It is useful particularly for those writing in a language that is not their first language (Tiers 2-3).
- **Large-scale information base**: GAI is trained on diverse sources, which provides a broad information base, however there may be gaps in the information base. Researchers could use these models to gain insights into various subjects, identify or find relevant literature, explore discourses and discover interdisciplinary connections (Tiers 3-4).
- **Speed and efficiency**: GAI models can quickly process large amounts of data and generate outputs immediately. Researchers could save time and effort on tasks and even automate some tasks such as identifying possible sources for literature reviews, creating artificial data sets or getting assistance with learning how to undertake types of data analysis (Tiers 3-4).
- **Pattern recognition**: GAI models can help identify patterns, trends, and relationships within data. Generating code to perform additional analysis and outputs on your data is also possible, such as in Python or R. Drawing on these outputs can lead to new insights (Tier 4).
- **Customisability**: Researchers can tailor GAI models to their specific needs. GAI can be asked to adopt specific genres for creating a variety of outputs suitable to different audiences/purposes. GAI also offers researchers the opportunity to fine-tune models (Tiers 3-4).

Considerations and limitations

Being aware of the considerations and limitations helps identify where unreliable responses are more likely and where more detailed checking is needed:

- **Hallucinations:** While GAI outputs may come across as authoritative and convincing, the responses are based on next word predictions. The software has no real understanding of the information it is providing. This leads to what are known as ‘hallucinations’ or ‘credible untruths’, thus resulting in misinformation. When using GAI tools, the outputs must be verified for accuracy and reliability.
- **Lack of transparency**: The inner workings of many GAI models are complex and opaque (“black box algorithms”), making it difficult to understand how they generate their outputs. This can be challenging for researchers to be transparent about their intellectual processes and ownership (Bekker, 2204).
- **Plagiarism**: GAI tools tend not to include attributions of their sources or training data. There are concerns about how data sets have been sourced to be used in the training of

these tools, which may contravene principles of academic integrity. This means that ideas could be misattributed or used out of context. It is important to fact check all information generated by GAI, and to find credible sources for information.

- **Data protection and privacy**: Be aware that data entered into GAI tools, for example, OpenAI stores and can use data stored in ChatGPT3.5, could lead to privacy issues. ChatGPT4.0 (the paid version) allows you to change this setting (uncheck the setting that data is saved). For general GAI tools, we recommend the use of Copilot instead of ChatGPT as when you login with your Wits credentials the user and company data are protected and not shared back to the tool for training. You might have to update consent forms to ask participants for permission to use their data for analysis in GAI.
- **Inherent biases or errors in training dataset**: GAI models learn from the data they are trained on, which may contain bias or particular perspectives. Researchers need to critically evaluate outputs for bias/hegemonic perspectives.
- **Misuse of generated data**: GAI can be used to create fake data or manipulate existing data. Researchers need to critically evaluate all data they work with.
- **Reproducibility issues**: GAI models may not be reproducible, as the same model may produce different outputs each time it is run due to the nature of the generative process. This can make it difficult for researchers to replicate results.

4. Research Integrity Implications

Authorship and Publications

It is not acceptable to use GAI to write research publications for you. Some journals may allow GAI for conceptualisation or copyediting but not for discussion or conclusions (original contribution). GAI can also not be listed as a co-author by many journals. Most journals provide instructions on how to acknowledge the use of GAI in research, which differs from disciplines to discipline. Where GAI is legitimately used to generate text, this should be referenced (e.g., APA style) with further declarations (i.e., including the prompts used). See examples for citing ChatGPT and generative AI. See some examples of major publisher statements:

Authorship policies of three large publishers (from Rahman et al., 2023, p. 3)

Publisher	Authorship Policy Extract	Comment
Springer (2023)	LLMs, such as ChatGPT, do not satisfy the authorship criteria. However, if the researchers use these tools, s(he) must mention their use in the appropriate section of their academic paper, such as the 'methodology' or 'acknowledgements' section.	GAI cannot be an author or co-author. If used, this should be clarified in the appropriate section.
Taylor & Francis (2023)	Authors must be accountable for their research work per the publishing agreement. As AI tools do not take this accountability, thus: AI tools cannot be co-authors in an academic paper. However, if a researcher uses these tools, s(he) must mention their use in the appropriate section.	GAI cannot be an author or co-author. If used this should be clarified in the appropriate section.

Elsevier (2023)	Though AI and AI-assisted technologies help you to enhance the quality and readability of the language of the work, they do not replace key researchers. Thus, the researchers are not allowed to list AI and AI-assisted technologies as an author or co-author nor cite AI as an author.	GAI cannot be an author or co-author. GAI cannot be cited as an author.

Research Participants and Data Handling

The Wits Research Integrity Policy and Research Integrity Procedures should guide any research processes and the use of data.

5. Supporting Writing Stages with GAI Tool Examples

The research stages and possible examples GAI tool usage are explored below. Please note that tools can change quickly and may become outdated.

Drafting consultant examples at the start or during writing (Tier 4 use)

An author can interact with a GAI tool to plan a research project or shape or develop an argument as in the back-and-forth of ideas with a fellow researcher or student. GAI can offer suggestions and criticisms about writing. Note that for students any suggested outputs should not take the place of supervisory feedback (Bekker, 2024).

Usage Possibility	Tools to Use	Opportunities and Concerns
Refine purpose, problem statement and rationale: Once you have a research topic and questions, ask for help to refine your problem statement, research questions or rationale.	ChatGPT, Gemini, Copilot, Claude	Follows a typical format. What is generated is likely hypothetical and without any references.
Abstract writing: Create an initial draft of an abstract or compare a generated one to yours. Provide bullet points and ask it to write as a paragraph.	ChatGPT, Gemini, Copilot, Claude	Helps with starting or rethinking an abstract. But cannot represent your ideas fully.
Evaluate academic writing samples: Provide feedback on strengths or weaknesses in writing pieces like a paragraph.	ChatGPT, Gemini, Copilot, Claude	Needs to be carefully evaluated as risk of making arguments less clear or poorer.
Speed up structured tasks: Given the appropriate text, can be asked to format references or format text in a specified format. Example prompt (Bekker, 2024): *Check the citations and reference list in the following text for accuracy and style using the [APA referencing style version 7]. List any corrections pointwise.*	ChatGPT, Gemini, Copilot, Claude	Can have errors, so requires critical editing.

Drafting consultant examples for reviewing literature (Tier 4)

GAI tools could assist in summarising research, finding additional sources, helping to understand specific discourses and assisting with literature reviews (Bekker, 2024).

Usage Possibility	Tools to Use	Opportunities and Concerns
Assist with literature reviews: Provide specific research topics or questions to generate related search terms for searching literature databases and web searches. Example prompt: *"Suggest keywords and search strategies that I can use to find relevant sources on technologies and teaching in higher education?"*	Elicit Perplexity Consensus Scopus AI (available to Wits users, go to SCOPUS > SCOPUS AI)	Elicit does full paper search, but only searches the semantic web. Additional functionality requires payment (limited credits for free version). Consensus uses GAI to find insights in research articles. Scopus AI draws on abstracts in the Scopus database for outputs. It also provides the references for articles used.
Create literature maps: Generates a map of relevant articles related to a specific seed article. Show the top citations and references related to a seed article. Selecting any article one can see who the article cites, and which articles are citing it.	LitMaps inciteful.xyz ResearchRabbit Connected Papers	Visual view of research field using citations. No made-up references. But some papers are behind paywalls.
Evaluate references: Helps with evaluating scientific articles by looking at citations. Provides the context of the citation and a classification describing whether it provides supporting or contrasting evidence for the cited claim.	Scite (Paid tool)	Contextualises article citations. But mostly for medical sciences field.
Summarise articles: Allows one to upload PDFs and summarise articles.	Scispace (Free version is limited) Scholarcy ChatPDF	Aims to help understand research papers better. Has some limitations and requires assessing the responses critically.
Identify key parts of article: Use for reading and to understand document structure and merge it with information via tooltips and other overlays. Category labels include Goal, Method, and Result.	Semantic reader	Can customise what is shown. Works better for structured articles.

Drafting consultant examples for data collection and transcription (Tier 4)

Usage Possibility	Tools to Use	Opportunities and Concerns
Refine interview and survey questions: With some context, GAI can help refine, evaluate, and test interview protocols or survey questionnaires.	ChatGPT, Gemini, Copilot, Claude	Useful for generic or standard types of questions. May not be appropriate for your target group.
Speeding up transcriptions: Allows live edits of transcriptions while listening to the transcript. Supports analyses of texts by, for	Otter.ai (free version is limited)	Can learn different accents and improves transcription, identifies speakers. Currently only English transcription

example, providing a list of key words/terms used.		possible, no other South African languages. Most features require payment.
Manufacture fabricated data: Creation of synthetic data for analysis. For data that is not available or impossible to gather (due to ethical or legal issues) or to be used for teaching or illustrative purposes.	Statice (paid only)	Provide instructions of what vast data sets should contain.

Drafting consultant examples for data analysis (Tier 4)

Usage Possibility	Tools to Use	Opportunities and Concerns
Python code generation: Using the code interpreter option, you can provide data in a file, then ask for Python code to be generated that will analyse the data. The Python code can be executed. Follow-up prompts can be used to refine the analysis.	ChatGPT4.0 Plus, Claude, Turbo (code interpreter option) (paid version)	The saved Python code can be used to analyse other datasets without uploading any data to ChatGPT. ChatGPT Plus/Turbo is subscription based. Claude is a free alternative. Python code should be checked before reporting. Uploading sensitive data would not be appropriate.
Text analysis and advanced statistical analysis: Perform text analysis or analyse large multimodal models. It generates text outputs (natural language, code, etc.) given inputs consisting of interspersed text and images.	ChatGPT4.0 (paid version) LangChain Voyant	Text analysis supports scholarly reading and interpretation of texts or corpus. Data analytics generates text outputs (natural language, code, etc.) given inputs consisting of interspersed text and images. May generate harmful advice, buggy code, or inaccurate information.
Support qualitative thematic analysis: Summarise texts, coding of themes, and identifying common themes across your codings. Can help to explain the coded content.	Maxqda (paid tool) and related data analysis tools Atlas.ti has incorporated ChatGPT.	Can choose the language and the length of the summary. Can modify the summaries as needed.

After writing: Proofreading examples (Tier 2)

Usage Possibility	Tools to Use	Opportunities and Concerns
Fix writing errors: Helps fix spelling, punctuation, grammar, and style errors and improves the clarity and flow of your text. *Example prompts (Bekker, 2024):*	Grammarly, ChatGPT, Gemini, Copilot, Claude	Grammarly works in Word and Google Docs via a plugin. Free version provides basics.

Proofread the following piece of writing and suggest corrections for spelling, grammar, tone and style errors. Ensure text is clear. Review the section for tone and style variations and note any, pointwise. Identify problematic or misused words, list each and provide recommendations for replacement words. *Make recommendations to improve the overall readability of the following text [insert text].*		

During or after writing: Language editing examples (Tier 3)

Usage Possibility	**Tools to Use**	**Opportunities and Concerns**
Rephrase sentences: Paraphrases sentences. Example prompt (Bekker, 2024): *Edit the writing piece below to shorten to 1000 words while preserving content, intention and clarity.*	ChatGPT, Gemini, Copilot, Claude, Quillbot (Free version of Quillbot keeps the sentence structure and replaces key words with synonyms. This is regarded as plagiarism according to the university's definition of academic misconduct. Recommend not to use).	Suggest to use as a first attempt to help engage with and understand a dense text, but do not submit Quillbot output in a final draft.
Reorganisation of writing: Improve structure and flow of writing. Example prompt (Bekker, 2024): *Make suggestions for reorganising the paragraphs in the writing piece below to improve the overall structure and flow of the argument. Briefly explain each modification.*	ChatGPT, Gemini, Copilot, Claude	Can have errors, so requires critical engagement with the suggestions before accepting suggested changes.
Dissemination: Help make your outputs (after publication) more accessible, i.e., to summarise or generate social media posts.	ChatGPT, Gemini, Copilot, Claude	

6. Bespoke Research GAI Tools

Researchers are also developing their own tools based on GAI platforms to support research processes. An example is the suite of tools that give feedback on your drafts, developed by the Academic Insight Lab based on the YouAI platform. The free version requires you to sign-up:

- **Purpose statement reviewer**: This GAI tool will provide feedback on your purpose statement to ensure it succinctly describes the study population, approach, and setting of the study.

- **Problem statement reviewer**: This GAI tool will provide feedback on your research problem to ensure it appropriately frames the research problem you plan to address.
- **Lit search**: Designed to help researchers and academics in their literature search process by identifying keywords, synonyms, search strings, and more, based on the key concepts and variables of your problem statement.
- **Writing diagnostic tool**: Evaluates academic writing samples, offering a holistic assessment through the lens of SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis.

7. Conclusion

GAI tools can be used in the research process to make your research and writing process more efficient. It is important to remember the limitations of GAI and ensure that any use of GAI does not contravene Wits ethical, research, and academic standards. You need to remain in control of the project, in other words, do not outsource the core thinking work of your research to GAI.

8. References

Academic Insight Lab (2023). HybridHorizons: exploring human-ai collaboration in research writing. YouTube. https://www.youtube.com/watch?v=nRLKLGsz6bc

Bekker, M. (2024). Large language models and academic writing: Five tiers of engagement. *South African Journal of Science*, 120(1-2), 1-5. http://dx.doi.org/10.17159/sajs.2024/17147

Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Hosseini, M., Rasmussen, L. M., & Resnik, D. B. (2023). Using AI to write scholarly publications. *Accountability in Research*, https://doi.org/10.1080/08989621.2023.2168535

ICAI. (n.d.). Fundamental Values of Academic Integrity. https://academicintegrity.org/resources/fundamental-values

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614. https://www.mdpi.com/2071-1050/15/7/5614

Nelson, J. (2023). The ethics of using ai chatbots for academic assistance: What you need to know. Linked In. https://www.linkedin.com/pulse/ethics-using-ai-chatbots-academic-assistance-what-you-james-nelson/

Rahman, M. M., Terano, H. J., Rahman, M. N., Salamzadeh, A., & Rahaman, M. S. (2023). ChatGPT and academic research: A review and recommendations based on practical examples. *Journal of Education, Management and Development Studies*, 3(1), 1-12.

Wits University (2020a). Research Integrity Policy. https://www.wits.ac.za/media/wits-university/research/documents/Wits%20Research%20Integrity%20Policy%20Approved%202020-04-03.pdf

Wits University (2020b). Research Integrity Procedures. https://www.wits.ac.za/media/wits-university/research/documents/Wits%20Research%20Integrity%20Guidlines%20%20Procedures%20Approved%202020-04-03-1.pdf