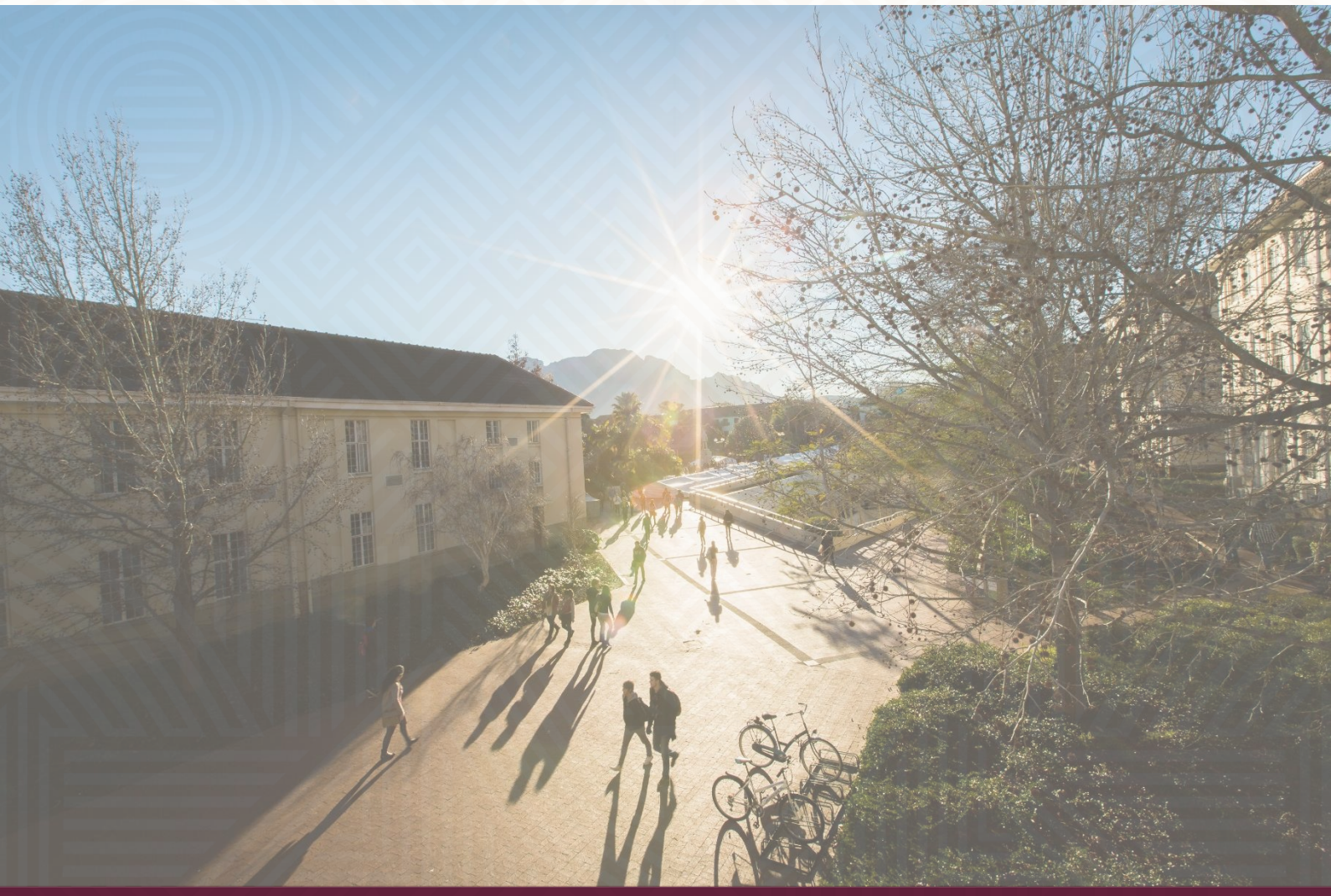
Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Ethical use of Artificial Intelligence in Research and Teaching-Learning-Assessment



Stellenbosch University's Position Statement

forward together
sonke siya phambili
saam vorentoe

Stellenbosch University's Position Statement:

Ethical Use of Artificial Intelligence in Research and Teaching-Learning-Assessment

Type of Document:	Position Statement
Purpose:	This Position Statement provides governing principles for the ethical use of AI in research and teaching-learning-assessment, assigns accountability and links up with specific governance documents to enable the required governance.
Approved by:	SU Senate
Date of Approval:	22 November 2024
Date of Implementation:	1 January 2025
Date of Next Revision:	Internally reviewed every year, and formally reviewed every three years or earlier if considered necessary based on the internal review.
Date of Previous Revision(s):	None
Owner:	Deputy Vice-Chancellor: Research, Innovation and Postgraduate Studies and Deputy Vice-Chancellor: Learning and Teaching
Curator:	Senior Director: Research and Innovation Senior Director: Learning and Teaching Enhancement
Keywords:	Artificial Intelligence, research, teaching, learning, assessment, ethics, principles
Validity:	The English version of this Position Statement is the operative version, and the Afrikaans and the isiXhosa versions are the translations.

SU governance documents are available at
www.sun.ac.za/policies

Stellenbosch University (SU) is firmly committed to pursuing knowledge, research and innovation, in service of society, and integrating our core values (namely equity, respect, compassion, excellence and accountability) through our actions. We provide a world-class centre for learning and development, which is globally relevant, yet rooted in the upliftment and transformation of our local communities. With this commitment, we acknowledge the institutional responsibility to provide our staff and students with the required principles and guidelines to navigate the profound impact of the development and application of Artificial Intelligence (AI) and more specifically, generative AI, on research and teaching-learning-assessment, and to uphold ethical principles for the promotion of responsible conduct of all research and teaching-learning-assessment activities undertaken under the auspices of SU.

It is widely acknowledged that AI presents unprecedented opportunities and benefits to revolutionise scientific enquiry and acts as a supportive tool for staff and students as part of their research and teaching-learning-assessment activities. It is acknowledged that AI can play a supportive role, particularly in an emerging economy context and where language barriers exist and provide positive learning opportunities. However, it is important to acknowledge the range of potential ethical challenges and risks posed when using AI in these activities. Consequently, a set of principles on how to use AI tools ethically and responsibly in support of research and teaching-learning-assessment are vital. Given the dynamic nature of AI, it is understood that there is a limit on the extent to which the institution can provide institutional guidelines and procedures that encompass all possible instances of these tools and their use. The principles apply to AI systems and -tools used in research and teaching-learning-assessment.

Governance:

As a responsible institution, the governance of the ethical use of AI systems and tools in support of academic endeavours - both research and teaching-learning-assessment - is critical. Due to the different nuances of the research vs. teaching-learning-assessment areas, as well as amongst disciplines, provision for the operational structures, guidelines, monitoring actions, guardrails and governance documents addressing the differences for various environments, will be provided in separate supporting documents, which will be updated from time to time as required. Faculties may provide additional guidelines and guardrails to support the ethical use of AI within their specific discipline. This Position Statement provides guiding principles for the ethical use of AI in research and teaching-learning-assessment, assigns accountability and links up with specific governance documents to enable the required governance. If the unethical use of AI leads to plagiarism, ethical misconduct, breach of intellectual property etc, it will be dealt with in terms of the relevant governance document regulating the specific matter.

Purpose:

In light of the above, this Position Statement sets out to provide a set of principles for staff and students regarding the integration of AI into research and teaching-learning-assessment and underlines the importance of academic integrity and taking responsibility and being held accountable for ethical conduct in research and teaching-learning-assessment. These principles draw on established frameworks and guidelines from leading national and international institutions and should be read and considered within the context of other relevant governance documents. The most pertinent governance documents are listed in Section A.

Definitions:

AI systems: "AI systems" refers broadly to any system, architecture, algorithm, framework, technology, etc., based on AI concepts and used for research or AI applications.

Example: Transformers are a type of AI system used in natural language processing and image classification.

AI tools: "AI tools" refers more specifically to software, libraries, packages or apps based on AI systems that deploy specific applications of some of these systems.

Example: ChatGPT is an AI tool based on transformers designed to generate text.

This distinction is not clear-cut, as AI tools can be considered AI systems, although not all AI systems should be termed AI tools. The term "AI tools" is used mostly when one deals with specific applications of AI systems.

Example: Self-driving cars are a form of AI system, which is too complex and technologically broad to be called an AI tool.

Generative AI: "Generative AI systems**"** are AI systems capable of generating text, images, videos, and other output by extrapolating from their training data.

Example: Examples of applications built on such systems include ChatGPT, Copilot, and Midjourney.

Principles:

SU adopts the following set of principles:

1. **Openness, Transparency, Responsibility, Accountability, Trust, Inclusion and Diversity:**

    a. **Openness:**

    Research: We encourage the promotion of openness in the use of AI systems and tools in the research processes, and the sharing of data, code and methodologies to facilitate reproducibility, collaboration and scrutiny.

    Teaching-Learning-Assessment: We further encourage the promotion of openness in the use of AI systems and tools in any teaching-learning activities, including assessments.

    b. **Transparency**:

    Research: Transparency is important to ensure that trust in the research endeavour is not eroded. The use of AI in any form must be declared during the scientific and ethics review processes (where applicable) and clearly explained in any research activities (e.g. research design, processes, protocols, reports, dissertations, theses, publications, conference/workshop/meeting presentations) and outputs (whether it is a partial or derivative of an output generated by AI), similar to what would be expected with the use of other software, tools or methodologies. Specifics regarding the AI systems or tools used, the extent of their use, and the nature of the contribution to research must be formally declared. Note that individual publishing houses of journals or authored books may have more specific requirements in their AI policies.

    Teaching-Learning-Assessment: The SU Assessment policy advocates for transparency, with "students receive clear information about the assessment requirements against which their performance will be measured for the various assessment opportunities and assessment methods". In the submission of assessment tasks, detailing whether AI use is allowed as well as the process of appropriate AI use can help safeguard students against unintentional wrongdoing and aid in suspected cases of wrongdoing (plagiarism or AIgiarism[1]). As a default rule during assessment: AI-generated content may not form a substantial part of any assessment or any portion of an assessment unless it is clearly indicated as permitted.

[1] The unethical, irresponsible, unauthorised or not-permitted AI use in teaching-learning-assessment.

c. **Responsibility**:

SU advocates for responsible research and upholds academic integrity in all its activities. The ability of critical thinking by staff and students remains an important part of the academic process and should remain intact. Hence, SU emphasises the importance of human-centred research design, critical thinking, risk assessment, originality and continuous evaluation of the output of AI systems. In the context of generative AI systems, staff and students must take full responsibility for analysing and verifying information and citing the original authors, as per the applicable referencing convention. It needs to be noted that while some AI tools can offer sources, these sources still need to be verified to ensure accuracy, quality and relevance (i.e. not all websites are acceptable as sources in the academic context). This is particularly important given the propensity of some generative AI tools to "hallucinate" or fabricate references.

d. **Accountability**:

As AI tools do not meet the requirements for authorship or inventorship due to the absence of human intellectual contribution and accountability, staff and students are always accountable for the accuracy, integrity and originality of their work, which includes any content generated through the use of AI tools. It is the responsibility of the author or creator of a piece or product to ensure that their work is factually correct and not likely to cause harm, i.e., through spreading false information, misappropriation or sharing personal information. Generative AI tools may not be listed as an author on any scholarly work. Use of such tools may be included in the Acknowledgements section of manuscripts submitted for review to journals or in accordance with the specific guidelines of the journal.

e. **Inclusion and Diversity**: Staff and students who use and develop AI systems or tools for research or applications, should be mindful of the possibility of bias (such as gender, race, worldview, and other biases) in such systems and should take measures to mitigate and address potential biases. Similarly, staff and students who make use of generative AI systems should be mindful of the biases (including racial, gender, worldview, and other biases) inherent in such systems and take steps to correct for such biases in the system outputs. In this regard, the National Department of Health emphasises that diverse stakeholder or participant groups should be actively involved in the design and testing phases to ensure fairness and representation and that algorithms be tested for bias and discrepancies.

2. **Academic/Scientific Integrity:**

a. **Originality and Plagiarism:**

Research: Research must be the original work of the author(s) with appropriate citation and referencing of any material used, produced or influenced (including through AI). The use of AI must not breach the university's plagiarism policies or research misconduct procedures.

Teaching-Learning-Assessment: To validly determine whether students have learned and achieved the outcomes of a module or programme, lecturers need to know that the work they are assessing is a student's own as per the SU Assessment policy (2022). The original contribution to work presented by a student as part of an academic activity can only be evaluated if it can be distinguished clearly from the contributions of others or the author's own earlier work. Where AI tools have been used, it should be declared what tools were used and how, when and where they were used. The student should also indicate why the work still qualifies as their own, especially if AI tools were used in producing the substance of the work being assessed (or any substantial part thereof).

b. **Ethical Use and Misconduct:** Where AI systems or tools are used during research, teaching and learning, the staff and students must ensure that their use does not compromise ethical standards or result in any form of research or academic misconduct (such as the fabrication of results or breach of confidential or sensitive information).

c. **Review process**:

To protect academic integrity and confidentiality, staff and students must take full responsibility and accountability for the accuracy, integrity and originality of their reviewer feedback. Staff, students and Research Ethics Committee (REC) members are discouraged from using generative AI to review research proposals, manuscripts, theses, dissertations, grant applications, assessments or any other form of research or teaching-learning-assessment output. Currently, most scientific journals do not permit the use of generative AI for the review of manuscripts due to the potential for confidentiality breaches and errors. Most major international funders (such as the NIH) do not permit the use of generative AI in the review of grant applications. AI platforms are known to retain content, which will lead to confidentiality breaches. It is further known that AI tools are not trained or validated to critically review academic content and may lead to inaccurate and biased reviews.

d. **Differentiating between when to declare the use of AI tools and when not to:** Although the preference is for declaring AI use, it is acknowledged that declaring the use of some forms of AI tools is not essential. To distinguish instances where staff or students must declare the use of AI tools vs where it is not essential, we can consider AI tools as falling into two categories, namely:

i. Use of Natural Language Processing (NLP)-based AI tools to assist in writing or correcting text or computer code: The tools can be generative, but they are not used to generate a substantial part of the relevant work or computer code. One example is Grammarly which assists a person in improving text that is already written. Another is Copilot which assists a person in writing and improving computer code. AI tools falling in this category do not necessarily need to be declared, unless explicitly required by the lecturer.

ii. Use of generative AI tools to create novel content such as text, images, music, or computer code, where users have little to no input: One example is the use of ChatGPT to generate large pieces of text, summaries of works or reviews. The use of AI tools falling in this category should be fully disclosed, including the purpose for which it was used (for instance, was it used to fast-track a literature review and summarise your literature?). However, any declaration of such AI use does not absolve staff or students from the responsibility of ensuring the accuracy and validity of the work produced.

These distinctions are not clear-cut, and where there is any doubt, treat it as the latter (d. ii) and declare the use of AI. Where possible, a dynamic list of AI tools falling under the first category will be made available and updated from time to time. For further guidelines, advice and monitoring controls, use cases, and a list of tools falling into these categories, please refer to the following:

- For Research refer to Research Development (sun.ac.za)
- For Teaching and Learning refer to https://www.sun.ac.za/english/learning-teaching/learning-teaching-enhancement
- For general use of AI refer to IT Service Catalogue - Home (sun.ac.za)

3. **Ethical and Legal Considerations:**

    a. **Safeguarding individuals and communities:** Research and teaching-learning-assessment that involves AI must prioritise transparency, accountability and fairness, and safeguard individuals and communities against bias, discrimination and other potential harms. Staff and students should be aware of the potential biases in AI-generated data and outputs and mitigate these risks. A further critical consideration is that privacy can be compromised through AI, and therefore staff and students should clearly state how privacy will be preserved when using AI during the research and teaching-learning-assessment.

    b. **Legal considerations**: As with all research and teaching-learning-assessment, the relevant regulatory requirements must be considered and potential risks must be mitigated. The use of AI in research and teaching-learning-assessment must comply with data protection regulations and respect the privacy and confidentiality of the research participants or subjects involved. Staff and students should further be aware that intellectual property (most notably, copyright) infringement can occur with the use of generative AI systems and should seek guidance from the Division for Research Development or Innovus where there are concerns. Although SU supports the drive to open science, the protection of personal- or confidential information should never be compromised.

4. **Education and Awareness:**

    a. **AI Literacy**: It is vital for staff and students (and where appropriate, for research participants) to become AI-literate and to understand the capabilities, as well as the potential risks and limitations, of AI tools. There is an institutional responsibility to provide training and support, in this regard, including training for staff, students and REC members. In particular, as part of the supervision of students, training on the ethical use of AI tools for research and teaching-learning-assessment must be offered to ensure ethical behaviour is promoted across all disciplines and categories of staff and students. The University is in the process of developing course material to optimally support our staff and students.

    b. **Guidance and Resources**: Staff will be provided with resources and training to guide students in the appropriate use of generative AI in research and teaching-learning-assessment, including the understanding of the ethical implications and consequences of using AI and the importance of compliance with Stellenbosch University policies. Fostering an understanding of taking responsibility and accountability in decision-making during the research and teaching-learning-assessment process will contribute to reinforcing ethically responsible behaviour amongst all staff and students. The University is in the process of developing guidelines and resources to optimally support our staff and students.

5. **Social Responsibility:**

    a. **Broader Implications:** SU recognises its social responsibility to conduct research, teaching and learning for the public good and to address societal challenges. It is imperative that staff and students consider the broader implications and potential dangers of their use of AI systems or tools or the creation of AI systems or tools, especially where it amounts to the protection of personal or confidential information, or abuse as weapons, or have a potential detrimental impact on the environment.

    b. **Contribution to Ethical Debates:** We encourage staff and students to contribute to ethical debates surrounding AI's impact on society, the economy and governance. Open, unprejudiced dialogue should be fostered to provide an opportunity for staff and students

to discuss the ethical use of AI in research and teaching-learning-assessment and address any concerns.

6. **Ethical Governance:**

    a. **Governance Documents and Guidelines:** As far as possible, the institution will continuously review and update its governance documents and/or provide guidelines on the use of AI in research and teaching-learning-assessment, to reflect technological advances and associated emerging ethical challenges.

    b. **Collaborative approach:** It is acknowledged that collaboration between universities, staff and students, funders and regulatory agencies is required to navigate the ever-changing landscape and share best practices. SU is committed to actively participating in this regard.

Conclusion:

AI is here to stay, and, therefore, our university commits to fostering an environment where AI can be used to responsibly enhance (rather than compromise or undermine) research, teaching and learning by upholding the integrity of academic work and ensuring the credibility and integrity of research findings. This commitment extends to staff, students and other stakeholders who collaborate in these activities with and on behalf of SU.

A. Relevant Policies and Guidelines (and subsequent reviews):

1. Policy for Responsible Research Conduct at Stellenbosch University. https://www.sun.ac.za/english/research-innovation/Research-Development/policies-guidelines
2. Stellenbosch University's (SU) procedure for the investigation of allegations of breach of research norms and standards. https://www.sun.ac.za/english/research-innovation/Research-Development/policies-guidelines
3. Policy on Plagiarism (in Support of Academic Integrity). https://www.sun.ac.za/english/research-innovation/Research-Development/policies-guidelines
4. Stellenbosch University ("SU") Procedure for the investigation and management of allegations of plagiarism. https://www.sun.ac.za/english/research-innovation/Research-Development/policies-guidelines
5. Draft interim SU guidelines on allowable AI use and academic integrity in assessment. Interim SU guidelines on allowable AI use and academic integrity.pdf (sun.ac.za)
6. Disciplinary Code: Rules regarding Disciplinary Action against Staff Members. Microsoft Word - DISSIPLINERE PROSEDURE AFR met veranderinge soos deur die Raad en Anton aanbeveel. finaal aan AP aanbeveel__46 (sun.ac.za)
7. Disciplinary Code for Students of SU. Disciplinary Code For Students Of Stellenbosch University_2021.pdf (sun.ac.za)
8. Stellenbosch University Code 2040. Code 2040 (sun.ac.za)
9. Teaching-Learning Policy[2]

[2] The revised Policy will serve at the Senate in 2024. The 2018 Policy will apply until the revised policy is approved.

B. References:

1. ASSAf and SciELO Guidelines or the Use of Artificial Intelligence (AI) Tools and Resources in Research Communication. 17 September 2024. https://assaf.org.za/wp-content/uploads/2024/09/Final-ASSAf-and-SciELO-Guidelines-for-the-Use-of-Artificial-Intelligence-AI-Tools-and-Resources-in-Research-Communication_17-Sept-2024-.pdf.

2. Bekker, M. Large language models and academic writing: Five tiers of engagement. S Afr J Sci 2024. 120(1/2). https://doi.org/10.17159/sajs.2024/17147.

3. Cambridge launches AI research ethics policy. 14 March 2023. Cambridge launches AI research ethics policy | Cambridge University Press & Assessment

4. Guidelines and recommendations for the use of generative artificial intelligence (AI) tools in research, University of Cape Town. 31 October 2023.

5. "I can't help falling in love with AI": chatbots and research integrity. 25 July 2023. AI in research - UK Research Integrity Office (ukrio.org)

6. Nature Editorial. Tools such as ChatGPT threaten transparent science: here are our ground rules for their use. Nature **613**, 612; 2023. *doi: https://doi.org/10.1038/d41586-023-00191-1*

7. Russel Group principles on the use of generative AI tools in education. 4 July 2023. New principles on use of AI in education (russellgroup.ac.uk)

8. South African Ethics in Health Research Guidelines: Principles, Processes and Structures 2024. 3rd Edition. Department of Health. South African Ethics in Health Research Guidelines: Principles, Processes and Structures NDoH 2024 3rd ed.

9. University Policies on AI. 2 October 2023. https://www.researchprospect.com/university-policies-on-ai/