Private Bag X6001, Potchefstroom
South Africa 2520

Tel: 018 299-1111/2222
Web: http://www.nwu.ac.za

Office of the Deputy Vice-Chancellor:
Teaching-learning
Tel: (+27)18-285 2106

Guidelines for the Utilisation of AI in Teaching and Learning at NWU

Prof RJ Balfour, DVC Teaching and Learning

1) Background to the Matter

Senate, at its second meeting of 2023, took note of developments concerning Artificial Intelligence and its impact on higher education generally, and teaching and learning in particular. While it was noted that faculties organised, or were in the process of organising, workshops and other forms of engagement as regards AI, Senate expressed a view that an institutional opportunity for engagement be organised through the Office of the Deputy Vice Chancellor Teaching & Learning, and that a report be submitted through SCTL for the further attention of Senate both on the phenomenon of, and suggested University responses to AI. The Centre for Teaching and Learning thus assisted the DVC's Office with the conceptualisation of a programme of panels and presentations, also featuring a pre-programme competition for any member of the University to develop an accessible and user-friendly podcast featuring a selected form of AI.

Central to the themes explored in the course of the NWU Symposium[i] is the importance of ethical considerations concerning the use of artificial intelligence that is not transparently accountable to its users, and barely accountable to its creators[ii]. While this is not a new ethical challenge, it is an enduring human problem that takes the form of dilemmas to not only act principally in relation to knowledge, but also to know what principled and ethical action entails.

2) Continuity and Change: the quest for accessible knowledge configuration, AI, and the ethical dilemma of academic accountability

Artificial intelligence is not new to this century. Rather, it is a further iteration of a human desire to achieve more efficient access to, and to (re)configure existing knowledge based on large natural language datasets. The artificiality of AI lies not in it *creating* 'artificial' knowledge, but rather in relation to its capacity to generate "knowledge" from across language data, provide intelligent responses to human questions – without human agency (even though, of course, the datasets are derived from very real human artifacts). The intelligence of AI is also predicated on its capacity for responses to be generated and configured in an automated manner; for example, in its latest configurations, AI can generate 'new' images and perhaps more problematically, 'new' knowledge. At present, some of the scholarly literature that critiques AI points out how such "new" knowledge is not reliable, valid or credible (that is, it has no verifiable basis when checked). In terms of the foreseeable future, it is anticipated that new versions of existing AI, and new AI (such as generative versions), will "rectify" even those problems.

The question remains as to whether, ethically, AI should be enabled to function in a disembodied manner from its human agency foundation (pointing to machine learning). Such an approach risks ethical disconnect, because there is increasing concern globally regarding the ethical frameworks for research (involving humans, the environment and even texts) which AI needs to either demonstrate an awareness of, or become sufficiently transparent as to be open to auditing. At present AI does not at present offer the semblance of possessing a worldview of any kind, is not made explicit in the reconfiguration of human know through AI, but we know it exists and, more importantly, we know that it configures power relationships in particular ways. However, it is possible to insert a worldview by means of prompts to the AI platform. While it is acknowledged that ChatGPT does not express a particular worldview, a "worldview" could be constructed or influenced by this platform.

The essay mills before they became part of 'the internet things', were notes generated and sold by students to each other and by teachers to students. Simply put, there is profit to be made through possibly unethical means. Teachers and academics have, more recently, joined the fray and so EssayHub, EduBirdie, and PapersOWl

came into being, giving rise in turn, to the need for both prognosis and 'cures'. The prognosis took the form of a focus on the development of academic integrity scholarship in which the problematic nature of plagiarism, wrongful text usage and non-acknowledgement of sources were made clear, whilst online 'cures' took the form, amongst others, of software packages like Turnitin, Plagiarism Check, and CopyLeaks. Despite these interventions, the capacity of AI to generate images or knowledge independently of agency is a concern. This includes data, which can be configured as recognisable knowledge, and being responsive to human questions (in ways that seem to be conversation-friendly), mean that ChatGPT, Bard, and Vicuna have captured the imagination in ways that essay mills could not: cheap, personalised access to expertise, and expert knowledge configured automatically to respond to the bespoke needs of the student and the teacher alike.

3) Higher education: selected affordances and risks associated with the AI value-proposition

The literature on the affordances of AI is extensive. It is also linked to the perceived detrimental effects of the internet at large, and addresses matters such as the development of critical thinking and critical exploration of material, synthesis of ideas, and other skills. This literature was referenced in several presentations associated with the Symposium and is also referenced in the summary.

3.1) Selected affordances linked to the responsible use of AI

What does AI enable in terms of access and success? The Symposium featured a number of inputs as described in the sections below. Expertise has traditionally been 'scarce and expensive', but AI is developing to such an extent that everyone will have access to expertise cheaply (without having to be a graduate, or professional, or having been trained for years in a field). Another challenge is that AI will enable one to filter discourses, news and other sources in terms of one's preferences, leading to what is described as bubbles: confirmation bias and affirmation bias are already features of human thinking, and there is no reason to expect that these will not feature in AI - not unless ethical attributes are built in. Such attributes should include scepticism about findings, awareness of biases, and self-reflection (McCabe & Dzogang, 2023). Websites and sources that can be access through websites are transparent mostly in terms of traceability: there is sufficient detail to trace sources for a reader / viewer to assess bias in a number of different ways. LLMs (Large Language Models) present information without the underlying factuality being evident.

For example, a conventional approach towards a machine learning assignment would have entailed giving students a dataset, instructing them to use a machine learning technique, and report on their results. This approach would lead to the student very easily finding the answer using an AI. Instead, given the present realities and abilities of AI, the assignment should be changed to introduce students to the kinds of problems that can be solved by collecting datasets, the machine learning workflow, and the tools that can be used to report results. The student then uses this knowledge to find a dataset and problem they are interested in, finds techniques through exploration that could be useful, actions the work, and then conveys their results as well as reflections back to the lecturer in a way that focusses on what was learned and experienced.

One of the affordances of AI is that it can personalise the learning journey (it considers pace, level of interest, the success of a student's response, as well as responsiveness, and thus can guide students in their learning). The downside is that dependence on AI likely discourages independent thinking, or the skills development associated with independent thinking. In short, the risk of undue dependence is that it diminishes confidence in one's own capacity to learn and display critical skills and, in so doing, may actually inhibit learning. Another affordance concerns the capacity of AI to absorb indigenous languages' datasets. At present, the availability of AI in African languages is scant, but there may well be changes in this regard in the very near future[iii]. The more languages are shared on AI platforms, the more sophisticated the capacity becomes, in any language. Thus, what is anticipated is a fair(er) balance between participation in AI generation between the Global North and the Global South. AI also provides valuable assistance to students in a manner that could be enabling in terms of learning. From the student panel at the Symposium, an important perceptual difference was noted: Artificial Intelligence should be seen as Assistive Intelligence: it can provide a deeper engagement than what is offered in a classroom situation where the teacher has to adopt a one-size fits all approach. Thus, not only is AI depth-enabled, but also saves time: the time taken to find sources and the time needed to digest information. AI also attends to grammar issues (particularly important to second language speakers); structure issues and quality issues, leading to better marks. This is potentially transformative, because it provides the student with an almost instant means to achieve fluency and in a world where fluency in English is popularly perceived as linked to intelligence. Furthermore, the capacity of AI to simplify information makes it easily understandable; it is experienced as a means of obtaining access, and assisting towards success. Also, automated feedback (on plagiarism) and grading is personal, and it takes less time to get feedback.

AI has given rise to new pedagogic forms like Prompt Engineering (which define the role of the AI, define the content or the audience, provide sufficient information, breaking up complex sentences) (Ng, 2023). AI is not simply useful for derivative material (i.e., looking for source material or collate insights from source material), but

also useful for creative pursuits where entirely new knowledge or new imagery is generated. This is evident, for example, when it emerges that AI can generate the semblance of authenticity, and turn out (at present) images that are yet to be proven false (what is referred to as an instance of 'hallucination'). Another area of usefulness is student co-creation of assignments: in collaborative assignments, the search for problems of a certain type can be accompanied by a requirement for students to reflect on the types and varieties of solutions (previously, the problem would be provided by the academic; now students have to find a problem and engage with it, guided by the lecturer). As noted above, in a machine learning assignment, a traditional approach would be to give students a dataset, instruct them to use a machine learning technique, and report on their results. This approach would lead to the student very easily finding the answer using an AI. Instead, the assignment should be changed to introduce students to the kinds of problems that can be solved by collecting datasets, the machine learning workflow, and the tools that can be used to report results. The student then uses this knowledge to find a dataset and problem they are interested in, finds techniques through exploration that could be useful, actions the work, and then conveys their results as well as reflections back to the lecturer in a way that focusses on what was learned and experienced.

With AI, academics can also set up simulations that would take a month to set up normally, but using the AI takes little time. The tools are available, but the University lecturer needs exposure to the tools to use them effectively. It is perhaps necessary to educate on and authorise the use of the tools through the CTL in collaboration with the School for Computer Science and Information Systems. Future collaboration should also include the Faculty of Engineering's MuST with a focus on deep learning. In this regard, a community of practice (CoP) should be established to support awareness of and sharing of skills associated with different forms of AI. A critical issue here is making sure that the expectations in assessments/ assignments meet the levels of the associated module outcomes. This leads to a fundamental reassessment of assessment: do we need it? Perhaps we need more than assessments; we need students' evidence of learning. The use of AI for authentic assessment is possible: AI is only one form of technology use- and so a community of practice (where all faculties are represented) would serve this additional purpose, under the leadership of the NWU Centre for Teaching and Learning. In this community pf practice, collaboration and cooperation is key (including portfolio, project and collaborative and cooperative learning experience). On another and related note, service learning and community engagement can also benefit from the uses of AI. This is the case given that, through the uses of prompts and questions, AI is becoming more useful in relation to community issues (for example health advice), for the uses of AI for SL, WIL and CE. The holistic application of AI across all areas of University work (not only teaching and learning or the broader administration) is thus relevant not only to staff, but also to students in relation to CE, SL, and WIL opportunities.

3.2) Selected perceived risks associated with AI

On the downside, the accuracy of AI generated information (i.e. factual correctness, credibility and bias) has still to be improved. Cautions and critiques of AI are well-documented. For example, the Future of Life Institute's (2023) much cited insight is relevant here: "AI research and development should be refocused on making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal", while the Centre for AI Safety stated that "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war" (2023).

There remains a need, however, for social interaction between academics and students, and between fellow students, to ensure relationship-building and learning (counsellor, guide, friend). What is evident is that AI is part of what it means to learn online: its usefulness as part of pedagogy is thus clear as part of (not apart from) the development of critical thinking skills. The other problem is that of the bubble: algorithms channel information, based on their analysis of interests, language, preferences, and there is further channelling of information that feeds these same or similar interests, preferences and beliefs. The abuse and potential abuse of these technologies (and the techniques of surveillance underpinning them) has long been known and is well-documented. That noted, it should be recognised that large language models (LLM) will lean towards a 'global', rather than heavily contextual perspective. Thus, unsurprisingly, topics that are discussed globally surface most in AI outputs. Local knowledge (inclusive of indigenous knowledge) becomes more prominent, however, the more it becomes used through AI. In light of this, it is important to show students how to access indigenous or localised knowledge, and in this regard, context becomes critical.

The educational and social risks associated with algorithm-based information channelling not only include the risk of bias and prejudice, but also the risk of anti-social behaviour that does not accord with the need for collaboration and cooperation with like as well as other, to mutual and ethical benefit: aloneness, loneliness and the 'bubble'. Simply put, learning to engage may be enhanced through practice with a chat bot, but a chat bot is not the future of the human species, as a species in which scarcity of resources and mutual need compels us to learn better those skills of collaboration and cooperation, and complex interdisciplinary problem-solving that

requires more than knowledge memorisation, or individual test performance scores, to demonstrate readiness for work, and for survival.

4) An emerging NWU perspective on AI and the Curriculum

Rather than excluding AI from the ambit of education experience (which seems to be the experience necessary for survival, including 'workplace readiness' in the 21st century), AI needs to feature as part of teaching and learning (seen as an assistant rather than an aberration): Educational approaches should teach the ability to identify AI, explore variants, use AI, acknowledge AI, and reward the responsible use of AI, but also correct incorrect or unethical use. The approach adopted by NWU is thus not an 'uncritical embrace'.

All of the above suggests that it is imperative to integrate the guided and responsible use of AI as part of the curriculum, recognising that it can play a role in realising the transformation of teaching and learning at NWU. Most obviously, this presents an opportunity to develop 21st century skills (see the NWU TL Strategy). It is important to signal awareness of examples of a number of institutions that 'banned' AI (University of Oxford: 2023), whilst others (for example, the University of Johannesburg, 2023) have signalled the special place of AI by making it part of a compulsory course for all University students. Treating AI as a single practice beyond the context of academic integrity risks missing the opportunity to explore best practice about how it can be used, and also how and when it should not be used. The teaching and learning industry (much of which is driven by for-profit organisations) has seen a blossoming and prominence accorded to proctoring and regulation of online assessment. These forms of assurance veer closely to surveillance, and cannot become the focus of the higher education assessment experience associated with AI. Assessment should, instead, focus less on the outputs and more on the journey towards obtaining or displaying the outputs. Students should be asked to show how they use AI. Therefore, how to assess evidence of good AI use should be explored and celebrated by the University; and these it should be linked to a strong ethical centre that is linked explicitly to the NWU graduates' attributes and values.

5) Academic Integrity and accountability for the uses of AI

The 'cure' for the ethical problem of accountability does not lie in regulation, but in an ethically informed perspective on the demonstration of accountability. This, in turn, requires sound guidance by the institution, such that instances of right and wrong practice can be identified and recognised appropriately. In other words, that heart needed by the Tin Woodman. Dishonesty is not new in academic life. The use of AI that occurs without guidance and without accountability is indeed a form of academic dishonesty. Aside from the implications for Academic Integrity this situation presents, there are many implications in terms of AI in the classroom (as part of the pedagogy associated with the 2021 NWU *Policy on Academic Integrity* and the 2023 SOP on Academic Integrity). For one, AI allows for the introduction of new knowledge and this open flexibility in the curriculum, should be encouraged.

From a teaching point of view, AI can also assume an assistant role for academics, making it easier to generate quizzes and texts. However, the downside is that if AI is used by academics and also by students, what will anyone learn, either as an academic or as a student? Salient areas that risk being diminished include creativity, reading and writing – suggesting that the NWU needs to reflect on changing the curriculum (as exemplified by the NWU PQM project) with implications for how we teach and assess. This will help to explore creativity, as well as collaborative and cooperative learning, in which reading and writing are only two of a larger skills sets students need to have. What emerges from the present situation, also, is concern about the authenticity of the source, the user (and possibly also, the assessment done by AI on behalf of an academic). In summary: it is essential for the University to have guidelines in place for the uses of AI in the curriculum and the activities associated with it (assessment, research and engagement).

6) NWU Guidelines as regards the place of AI in the Curriculum and NWU Values

Integrity is a key value embedded in the Curriculum as well as in the Values of the NWU. The NWU considers AI not only as some sort of elaborate search engine; rather, it is also potentially a source of new knowledge. Admittedly, the challenge with LLMs is that no single response is exactly the same, because much depends on the preferences, the nuances and the profile of the person/s asking the question. In light of the above, the importance of standardised guidelines, specific to the needs to the discipline, or entity or administrative unit, needs to be made explicit.

1) Academics have a responsibility to train students to use AI appropriately, responsibility and frequently enough, within their disciplines and within the guidelines provided by the University.

    1.1) Part of that responsibility is to be able to illustrate an awareness of the limitations of AI. What is evident is that the human agent cannot be removed from the technology at the present time. Research still has to be undertaken by humans, and thoughts have still to be expressed by humans as part of learning, development and growth.

1.2) The use of AI should always be within and subject to the framework of statutory prescriptions and university policy and rules, especially those regulating intellectual property rights and privacy of information.

2) The University's regulation of AI (in terms of how, for example, Bard, ChatGPT and others are used/should be used) must be sufficiently flexible to allow for a changing recognition of implications (rather than an uncritical embrace, or equally uncritical ban, of AI in the teaching learning and research environment) to ensure that the gap between guidance and application is not too wide.

3) The University has a responsibility to make awareness of the forms as well as the uses of AI, part of the University's staff development programmes.

3.1) While it might be safe to assume some students will have a nuanced appreciation of the technology, not all academics and students will know enough, and there will likely be different levels of knowledge and understanding of AI. To this end, a platform for sharing among students and staff is necessary, and an educational offering for staff is required.

3.2) Academic units (subject groups, schools, research entities) should provide a well-articulated stance on the uses of AI (and AI detectors) tailored to the needs of the discipline and/or the nature of research in such a manner that students and staff are clear regarding expectations.

4) The relationship between academic integrity practice and values needs to be informed by a clear understanding of the forms and uses of AI. This relationship needs to feature in the declarations made by students about work submitted for assessment, which should allow for AI assistance to be acknowledged.

4.1) Students who use AI must not only to declare such usage on an institutional template to be developed by CTL, but also to provide in-text referencing of online sources (an example of which is provided here).

While the use of generative artificial intelligence (AI), particularly large language models (LLMs) like ChatGPT, is often integral to modern research and development, it is important to note that direct in-text citations of these tools are not possible. This is primarily because the generated information is inherently instance and context-specific, and dynamically formulated in response to each unique prompt. Consequently, to appropriately acknowledge and reference the use of such software, it is recommended to use the following citation formats:

APA referencing formula:
Format: Author. (Year). Title of software or model (Version date if known) [Format]. URL.
In this format:
   • "Author" refers to the organisation or individual that developed the software or model.
   • "Year" refers to the year the software or model was published or updated.
   • "Title of software or model" is the official name of the software or model.
   • "Version date if known" refers to the version of the software or model, if applicable.
   • "Format" is the description of the type of model as provided by the publishers.
   • "URL" is the web address where the software or model can be accessed.
For example:
For ChatGPT:
   OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. https://chat.openai.com/chat.
 Or link to a specific chat, for example:
   https://chat.openai.com/share/651dcdc4-679b-4ca5-a41c-a6ee3edaf753

For DALL-E:
   OpenAI. (2023). DALL-E [AI image generator]. https://labs.openai.com

Harvard referencing formula:
   Format: Author(s). (Year) 'Title of software or model [Type of Model]'. Available at: URL (Accessed: Day Month Year).
In this format:
   • "Author(s)" refers to the organization or individual that developed the software or model.
   • "Year" refers to the year the software or model was published or updated.
   • "Title of software or model" is the official name of the software or model.
   • "Type of model" refers to a brief description of the type of software or model.
   • "URL" is the web address where the software or model can be accessed.
   • "Accessed: Day Month Year" is the date when the software or model was last accessed by the person citing the source.
For example:

For ChatGPT:
   OpenAI. (2023) 'ChatGPT (Mar 14 version) [Large language model]'. Available at: https://chat.openai.com/chat (Accessed: 12 June 2023).
For DALL-E:
   OpenAI. (2023) 'DALL-E [AI image generator]'. Available at: https://labs.openai.com (Accessed: 12 June 2023).

5) Flexible decision-making, broad sharing and support for AI awareness and its uses, should be a salient feature of management engagement as well as classroom engagement, in the context of the University's emphasis on the development and infusion of academic integrity practice.

References:

Baum, L.F. (1900). *Wizard of Ozz*. London: Penguin Books.

Boston University. (2023). *Policy on the Use of AI Text Generation*. https://www.bu.edu/files/2023/02/GAIA-Final-2023.pdf (accessed on 2023/06/06).

Centre for AI Safety. (2023). *Statement on AI Risk: AI experts and public figures express their concern about AI risk.*(https://www.safe.ai/statement-on-ai-risk) (Accessed on 2023/06/08).

Future of Life Institute. (2023). *Pause Giant AI Experiments: An Open Letter*. (https://futureoflife.org/open-letter/pause-giant-ai-experiments/) (accessed on 2023/06/08).

McCabe, C. & Dzogang, F. (2023). *Addressing LGBTQ+ bias in GPT-3*. https://medium.com/asos-techblog/addressing-lgbtq-bias-in-gpt-3-93e556a1b0fe (accessed on 2023/06/07).

Ng., A. (2023). *ChatGPT Prompt Engineering for Developers (Short Course in collaboration with Open AI).* https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/ (accessed on 2023/06/08).

North-West University. (2021). *Policy on Academic Integrity*. https://www.nwu.ac.za/sites/www.nwu.ac.za/files/files/i-governance-management/policy/2021%20Update/2P_2.4.3.2/2P_2.4.3.2_2021_e.pdf

Stanford University. (2023). *Generative AI Policy Guidance*. https://communitystandards.stanford.edu/generative-ai-policy-guidance (accessed on 2023/06/06).

University of Oxford. (2023). *Unauthorised use of AI in exams and assessment*. https://academic.admin.ox.ac.uk/article/unauthorised-use-of-ai-in-exams-and-assessment (accessed on 2023/06/06).

University of Johannesburg. (2020). *Teach Artificial Intelligence to all Students*. https://www.uj.ac.za/news/teach-artificial-intelligence-to-all-students/ (accessed on 2023/06/06).

University of the Witwatersrand. (2023). *Approach to the use of AI in teaching and learning at Wits*. https://www.wits.ac.za/media/wits-university/learning-and-teaching/cltd/documents/AI-in-teaching-and-learning-at-Wits.pdf (accessed on 2023/06/06).

UNESCO. (2021). *AI and education: guidance for policy-makers*. Miao, Fengchun; Holmes, Wayne; Ronghuai Huang; and Hui Zhang. https://unesdoc.unesco.org/ark:/48223/pf0000376709 (accessed on 2023/06/06).

Endnotes

[i] The Symposium Programme featured on the leaflet, an image of a robot-like figure (replete with very human tattoos on its arms) engaging with a smartphone: ostensibly a humanoid representation of AI engaging with a device which features AI. This representation is often the subject of science fiction fantasy and sometimes horror (the creation of machines that come to function without reference to humans). In the *Wizard of Ozz* (L. Frank Baum, 1900) the Tin Woodman observes "You people with hearts have something to guide you and need never do wrong; but I have no heart and so I must be very careful". In this narrative about a journey to self-discovery (embodied by the Emerald City), Dorothy (a human) is accompanied by characters made animate from the inanimate (the Scarecrow, whose journey entails an aspiration to wisdom) and the Tin Woodman (who hopes to obtain lost heart); as well as animate characters like the Cowardly Lion (who seeks courage) and Toto, Dorothy's dog. Whilst this is a children's fantasy story, the desire for awareness, the need for empathy based on wisdom, a recognition of the interdependence between the animate and inanimate, the human and animal and both dependent on a sometimes hostile and sometimes friendly environment, affirms key values: trust, respect, recognition, diversity, solidarity in the face of adversity, self-knowledge and knowledge about the world. In the *Wizard of Oz*, the heart of the Tin Woodman is signalled not so much as the seat of sentiment, but rather that the ethical compass which makes sense of what is valuable, and makes sense of why connections of various kinds, matter. "Having the heart" is thus the facility necessary to be able to make right choices; to distinguish between right from wrong.

[ii] Speakers and experts external to the University were also invited and by the symposium date on 5 June, a 179 registrations had been received (of which 80 were face to face registrations for the 40 open spaces and 99 were online registrations). Inclusive of this "Summary and Way Forward" document, thirteen presentations featured on the day (please see programme for 5 June 2023 attached) with guests from the University of Lincoln and the University of the Witwatersrand, respectively. At the outset key guiding documentation had been shared by the DVC T&L Office with members of the NWU Teaching Learning Leadership Forum. These sources included reference to UNESCO's framing document for policy makers on AI, as well as selected institutional responses to AI as featured in the list of references to this document.

[iii] Recent tests of ChatGPT-4 using Afrikaans (84.1%) ran a close second to English. Swahili has also been tested using ChatGPT-4 (at 78.5%): https://openai.com/research/gpt-4