



TEQSA Request for Information

Artificial Intelligence

The University of Sydney Action Plan

Contents

Introduction	2
The University of Sydney action plan	3
Principle 1: Establishing rules	3
Principle 2: Providing equitable access	8
Principle 3: Building familiarity	9
Principle 4: Fostering trust	16
Appendices	17
Appendix 1 – Generative AI Governance Roadmap.....	17
Appendix 2 – Generative AI Steering Committee Terms of Reference.....	18
Appendix 3 – Guardrails for Generative AI use at the University of Sydney	20
Appendix 4 – Generative-AI: guidelines for researchers	22
Appendix 5 – Dynamic Generative AI Roadmap	26
Appendix 6 – The AI x Assessment menu	28
Appendix 7 – Guidelines for determining academic integrity breaches involving AI	29
Appendix 8 – Generative AI essentials for educators workshop slides.....	30

Introduction

The University of Sydney has recognised both the risks and opportunities presented by generative AI in the context of higher education. Recognising that this new technology cannot be outrun, the University has responded carefully and positively to generative AI, rather than imposing ineffective bans. The University has centred its approach around four guiding principles: establishing rules, providing equitable access, building familiarity, and fostering trust. We have encouraged an environment for staff and students to work together to develop productive and responsible ways to engage with generative AI – respecting the risks and also embracing the opportunities. All initiatives have been guided by strong governance leadership and oversight to ensure innovations are balanced and aligned with compliance obligations.

The University recognised the urgency required to respond to generative AI. As such, the University's action plan includes many initiatives already completed or in progress, with key future actions articulated throughout. The Higher Education Standards Framework (Threshold Standards) 2021 have been embedded within all aspects of the University's action plan. The relevant HESF standards are represented via the following shorthand:

- [OP] 1.3.1: Orientation and Progression
- [LOA] 1.4.3-5(b): Learning Outcomes and Assessment
- [DEI] 2.2.1: Diversity and Equity
- [CD] 3.1.3: Course Design
- [SD] 3.2.3b: Staffing
- [LRS] 3.3.1, 3.3.4: Learning Resources and Educational Support
- [RT] 4.2.1a, 4.2.4-5: Research Training
- [ARI] 5.2.1-4: Academic and Research Integrity
- [MRI] 5.3.2, 5.3.4b: Monitoring, Review and Improvement
- [CGA] 6.1.4: Corporate Governance, 6.2.1(h, k): Corporate Monitoring and Accountability, 6.3.1(a, d)-2(a, d, h): Academic Governance

The University of Sydney action plan

Principle 1: Establishing rules

Governance structures

- The University’s action plan for generative AI has been crafted in close collaboration with the Office of General Counsel (OGC) to ensure a safe and responsible pathway that provides a robust framework from which to apply innovative approaches. A Roadmap ([Appendix 1](#)) was created, outlining critical governance milestones required including key policy updates, establishment of a steering committee and working groups, establishing guardrails and determining risk appetites.
- The University Executive established the Generative AI Steering Committee in August 2023 [CGA, MRI]. Co-chaired by the DVC (Education) and DVC (Research), membership included representatives from OGC, the Operations Portfolio and Information and Communications Technology (ICT) and is tasked with the broad oversight and strategic direction of generative AI across the University. This committee meets monthly, reporting regularly to the University Executive [CGA, MRI]. The Steering Committee’s terms of reference are in [Appendix 2](#).
- The Generative AI Steering Committee is supported by the Generative AI Coordinating Group. Membership includes staff involved in the operational delivery and impacts of generative AI and includes representatives from Research, Education, ICT, Library, Student Administration Services, Operations and OGC. The Coordinating Group meets monthly to share updates and ideas, as well as flag significant issues that require attention from the Steering Committee. A critical focus for this group is to monitor and review outcomes of initiatives responding to generative AI across key parts of the University [MRI], with outcomes tracked via an ‘in flight initiatives’ tracker. A future goal for the University is to expand this monitoring to capture a more holistic inventory of data. Within this Coordinating Group there are three focus groups: AI in Education Working Group (Education), Architecture Process Uplift (Operations) and Research Data Advisory Group (Research).
- The Education Portfolio established the AI in Education Working Group [MRI, CGA] in May 2023. This Working Group brought together staff and student representatives from across the University to discuss the implications of generative AI and share and review practices and policies across faculties. Representation from the Research Portfolio was an important feature of the Working Group, discussing implications for Honours and Higher Degree Research Training, as well as research data security impacts [RT, ARI]. Initially the AI in Education Working Group met monthly and has since transitioned to a reduced meeting schedule. The Working Group reports regularly to both Academic Board and to Senate on developments in the technology, coursework assessment and integrity policy and training [CD, SD, ARI, CGA].
- The Research Data Advisory Group (RDAG) advises on research data operations and governance. Although not specifically dedicated to generative AI, discussions surrounding this, and security of research outputs are a regular feature of the RDAG. It is chaired by the Pro-Vice-Chancellor

(Research) and reports to University Executive Research Committee. Membership includes all Associate Deans (Research), Pro-Vice-Chancellor (Research Training), Pro-Vice-Chancellor (Research Integrity) and senior leaders in Professional Service Units.

Guidance

- An important part of the University’s action plan was the establishment of a set of guardrails to help inform and guide students and staff when interacting with generative AI. These guardrails were developed by the Generative AI Coordinating Group and endorsed by the Generative AI Steering Committee [ARI, CGA] ([Appendix 3](#)). These guardrails encourage the University community to learn and experiment with generative AI whilst ensuring that data, privacy, intellectual property, and other information is appropriately protected.
- The Research Portfolio has established guidelines ([Appendix 4](#)) for generative AI in research [ARI, CGA]. These guidelines focus on handling sensitive information, research integrity, research ethics, and how to use generative AI safely for research purposes. Important aspects include advising users not to upload unpublished data or confidential/sensitive information into generative AI tools, dealing with data sovereignty considerations, validating generative AI outputs, and aligning with publisher and funder requirements. These guidelines are provided to all researchers and research students and are available on our internal intranet site [OP].
- The University has updated its [Academic Integrity Policy 2022](#) [ARI] to explicitly include *inappropriate* use of AI. This also includes requirements for unit coordinators on specifying permissible and prohibited tools that students can use within assessments.

Strategy, policy, and principles

- The AI in Education Working Group produced an AI Strategy Green Paper in July 2023, which was later adopted and endorsed by the Steering Committee. The Green Paper formed the basis of the ‘Dynamic Generative AI Roadmap’ [CGA] ([Appendix 5](#)), which was provided to University Executive. The paper is grounded in Australia’s AI Ethics Principles, and establishes an aspirational position on generative AI for the University, including:
 - AI has applications in all facets of University work
 - Human agency, expertise, and accountability are central
 - AI must benefit the University and its community
 - We engage productively and responsibly with AI
 - Where AI is used, it is transparent and documented
 - Our staff and students will model the use of AI
 - AI-human collaborations are normalised
- In consultation with faculties, the Education Portfolio established the ‘[two-lane approach](#)’ to assessment design [LOA, CD, SD] (Table 1). The University’s new approach to assessment is closely aligned with the [TEQSA guidelines](#) on assessment design in the age of AI, balancing the need to help students engage productively and responsibly with AI, whilst assuring attainment of learning outcomes.

Table 1: Overview of the two-lane approach to assessments in the age of generative AI.

	Lane 1	Lane 2
Role of assessment	Assessment <i>of</i> learning.	Assessment <i>for</i> and <i>as</i> learning. Emphasis is on the <i>process</i> of learning instead of the product.
Level of operation	Mainly at program level. May be must-pass assessment tasks.	Mainly at unit level.
Assessment security	Secured, in-person, supervised assessments.	Not secured.
Role of generative AI	May or may not be allowed by the examiner.	As relevant, use of AI is supported and scaffolded so that students learn how to productively and responsibly engage with AI.
Alignment with TEQSA guidelines	Principle 2	Principle 1
Examples	In person interactive oral assessments; viva voces; contemporaneous in-class assessments and skill development; tests and exams.	Leveraging AI to provoke reflection, suggest structure, brainstorm ideas, summarise literature, make multimedia content, suggest counterarguments, improve clarity, provide formative feedback, learn authentic uses of technology, etc.

- The Education Portfolio has developed an AI x Assessment Menu [SD, CD] to support faculties to develop their lane 2 assessments ([Appendix 6](#)). This menu approach to assessment was chosen over a ‘traffic light’ or ‘assessment scale’ approach, recognising that instructing students to only use AI for certain purposes is untenable and that it is impossible to restrict AI use in unsecured assessment. This method also accepts that any unenforceable restriction damages assessment validity¹. The AI x Assessment Menu provides many options for students to apply generative AI to assessments and introduces the idea that an educators’ role is to help scaffold and support students to engage productively and responsibly with AI to enable learning within lane 2 assessments.
- Assessments are being redesigned towards either lane 1 or lane 2 as unit coordinators and program directors reconsider unit and program learning outcomes. Lane 1 assessments are being considered to assure attainment of learning outcomes, especially at the program level, and in many cases as ‘hurdle’ (must-pass) tasks. For unit coordinators of units where many or even all assessments are in lane 2, a cultural shift is required to assure attainment

¹ <https://www.tandfonline.com/doi/full/10.1080/02602938.2023.2209298>

- of program learning outcomes through lane 1 assessments. Targeted Strategic Education Grants have been made available in 2024 to assessment leaders in faculties to pilot new approaches to assessment in the age of AI.
- Work has commenced to configure the curriculum management system to capture coordinators' current practice relative to the two-lane approach.
 - The University updated its [assessment principles](#) as part of its [Coursework Policy 2021](#) (parts 75-79) [CGA]. In addition to revising existing principles, two new principles have been added that respond to AI and assessment:
 - Assessment practices must be integrated into program design
 - Assessment practices must develop contemporary capabilities in a trustworthy way

The University updated its mandatory academic honesty induction module for students to reflect these changes [OP] as well as providing training workshops, resources for tutorials and a website co-designed with students to broaden understanding in ethical and effective use of AI in learning and assessment.

- In consultation with faculties, the AI in Education Working Group developed guidelines for how AI tools may be used for [marking and feedback](#) [CGA]. The marking guidelines were endorsed by the University Executive Education Committee and will be refined and solidified into formal policy as AI increasingly forms part of our productivity and educational technologies [MRI]. Through these guidelines we aim to ensure that educators continue to be at the centre of marking and feedback decisions, while allowing for responsible experimentation with AI.

Academic Integrity

- The University's [Academic Integrity Policy](#) underwent a major policy review and was approved in November 2022 [ARI, CGA]. The updated Academic Integrity Policy includes the inappropriate use of generative AI as a form of breach and has been monitored since the introduction of ChatGPT. Faculty sessions were held for unit coordinators on the new policy, which covered the new form of breach and recommended reporting process. The Office of Educational Integrity developed decision making guidelines ([Appendix 7](#)) to outline how decisions relating to AI breaches will be handled.
- Related data has been reported to governing bodies via formal annual reporting every April to the Academic Standards and Policy Committee, and through to Academic Board and Senate. A bi-annual report on AI trends capturing unapproved use of AI and has also been reported on since 2023. Data is captured within the academic integrity reporting system and integrity data will be released to faculties ahead of semester 2, with a breakdown of integrity breaches and any areas of concern.
- The [Academic Integrity Policy](#) and [Academic Integrity Procedures](#) underwent further revisions ahead of Semester 2, 2024 to explicitly cover acceptable use and requirements for coordinators on specifying permitted and prohibited digital tools within their unit and individual assessment items. This refreshed the requirements for undertaking a risk assessment for each iteration of a unit of study and updated mitigation strategies. Individual unit coordinators will be provided with template wording that can be included in their unit sites about how to phrase appropriate use and acknowledgement of generative AI as part of completing assessments. Additionally, improved resources for academics to help determine common academic indicators for identifying breaches in written work remains a continued priority for the University.

Future action plan

- Understanding that the University cannot prevent AI use within unsecured assessment, the [Academic Integrity Policy](#) will be updated so that the default position is that AI is permitted unless otherwise specified within the assessment instructions [ARI]. This will increase assessment validity and drive change in assessment design towards an integrated lane 1 and lane 2 assessment approach [LOA].
- The University will continue to update the academic honesty induction material for students to include AI [OP]. This will be developed in line with the updated assessment policy changes mentioned above, to be completed by January 2025.
- Aligned with the above policy changes, the University aims to update its curriculum management system in 2026 to accommodate a new configuration of assessment categories and types. This will require staff to adopt assessment designs that are valid and aligned with the two-lane approach. For example, assessment categories can only be classified as ‘secured’ lane 1 assessments that have the security features outlined in Table 1.
- Following the pilot program in 2024, the University will undertake a two-stage approach to transition all assessments to be aligned with the updated Policy. Stage 1 will see 60% of assessments compliant with policy and is expected to be complete by January 2026, with the remaining assessments to be transitioned to the two-lane approach by January 2027. A budget request is currently being developed to ensure this work is supported across all programs.
- The University recognises the need to incorporate the use of AI and emerging technologies into the course review process. The assessment strategy evaluation, including the diversity of assessments and potential risks of academic integrity, are already incorporated within the current course review process in place [LOA, ARI, MRI].
- The course review process is currently undergoing a continuous improvement evaluation. The implications of AI and other emerging technologies, in line with the other University-wide initiatives, will form part of this review. Any recommendations resulting from this evaluation process will be implemented within the 2025 course review cycle [LOA, CD, MRI] and will be supported through additional training to relevant stakeholders.
- The University is an active participant in community practice discussions across the sector, where AI and similar topics are discussed in the context of the course reviews and continuous improvement initiatives (e.g., ATEM Reviews Network, discussions and benchmarking activities with other HE providers) [MRI].
- From 2025, all new proposed academic offerings will be required to consider the use of AI and emerging technologies throughout the approval process. This will be considered under the evaluation of the proposed pedagogical approach with reference to the delivery modes and assessment strategy to ensure attainment of learning outcomes and graduate capabilities across diverse student cohorts [CD].
- Professional accreditation bodies will be consulted on their requirements regarding the use of AI and emerging technologies [LOA] [CD].

Principle 2: Providing equitable access

- Equitable access to state-of-the-art generative AI models (like GPT-4) is essential for students and staff to develop familiarity. The University has worked with Microsoft to make its AI-powered tool, Copilot for Web, available across the institution for free [DEI]. Adobe's generative AI, Firefly, is also available free to all staff at the University.
- The University has developed an in-house generative AI platform called Cogniti². Cogniti allows teachers to design their own 'AI agents' which are powered by AI models like GPT-4. Crucially, it allows them to steer the behaviour of AI by providing custom instructions and curating the resources available to it, as well as monitoring student usage. Cogniti allows educators a self-serve mechanism to make AI tools available to students in a controlled environment to support pedagogy [DEI, CD, LRS]. This is fostering [new approaches](#) to learning and teaching, including providing formative feedback, personalised support, and experiential learning.
- AI-based AI detection tools (from Turnitin) have been made available to staff in the central Office of Educational Integrity. Due to its limitations, this tool has not been made available to other staff, or students. Our approach to AI detection software acknowledges that it has inherent issues, and only serves as one piece of data in academic integrity investigations. Students are aware of the presence of AI detection software, which serves as a deterrence [ARI].

Future action plan

- The Library, in collaboration with the Research Portfolio and Education Portfolio, aims to pilot research-specific AI tools. They are collecting feedback from key stakeholders and considering longer-term investments [RT].
- ICT will work in partnership with the Education Portfolio and faculties to consider discipline-specific AI tools, such as image generation AI for architecture. This will take into account the requirements of industry and professional bodies in terms of graduate capabilities around AI tools [LOA, CD].
- The University will monitor the use of University-provided AI tools, such as Copilot for Web and Cogniti, to understand how students and staff engage with these resources [MRI]. For example, we will monitor patterns in use across different faculties and student cohorts to identify areas of stronger and weaker engagement and review the pedagogical impacts of these tools. University Executive will have strong oversight of these initiatives through annual reporting of third-party educational technologies report provided by the eTools Review Committee.
- The University will continue to engage with TEQSA, Universities Australia, the Group of 8, accrediting bodies, and international partnerships such as Universitas 21 to ensure we are continually informed of best practice and respond to industry needs [LOA].

² <https://cogniti.ai/>

Principle 3: Building familiarity

'Building familiarity' is a concept which comprises the level of understanding students and staff have of generative AI, and their comfort levels using generative AI. We intentionally use the word 'familiarity' instead of 'skill' to reflect the fact that not all users will develop specific generative AI skills. The term 'familiarity' also emphasises an awareness of the broader context of generative AI outside of its use, including ethics, privacy, and safety. This includes how AI impacts assessment and how the University responds to it. There is currently a wide range of levels of familiarity among the current and incoming student and staff cohorts. This is particularly significant for incoming students with varying levels of generative AI experience in school or workplaces, exacerbating socioeconomic inequalities. These equity implications are a priority for the University and our commitment to equal access and building familiarity together aim to address these issues.

Staff

- As generative AI has moved rapidly, so too has the University's response and accompanying support for staff. The Education Portfolio has designed training sessions for staff to help build familiarity with generative AI tools [SD]. Launching in January 2023, these workshops also communicated the urgency and scale of action required to respond to generative AI. Workshops were initially designed to build familiarity with AI and its implications for assessment design ('Prompt engineering for educators – making generative AI work for you'; 'Responding to AI and assessments'). These workshops have now been consolidated into the 'Generative AI essentials for educators' ([Appendix 8](#)). This workshop currently focuses on:
 - i. The current state of generative AI
 - ii. Demystifying generative AI technology
 - iii. Introducing the two-lane model, assessment menu and assessment redesign
 - iv. Demonstrating available AI tools and prompting basics
 - v. Outlining possible uses of generative AI in education
 - vi. Encouraging a 'possibilities' mindset

Appendix 7 – Guidelines for determining academic integrity breaches involving AI

Guidelines for determining breaches related to the use of generative AI tools

When determining academic integrity breaches involving AI, it is crucial to consider the severity of the breach and the specific circumstances surrounding each case. The severity of the findings and penalties will vary depending on the specific details of the AI-related integrity breach, including the affected proportion of the student's work and their own contribution and the student's academic record and previous breach history. Penalties should be applied consistently and fairly, with discretion for individual circumstances and the opportunity for students to present their responses. Feedback, informal warnings and penalties should educate on appropriate use of AI.

The Office of Educational Integrity will conduct a preliminary assessment to gather any supporting information and determine the appropriate decision maker to refer the case. In these circumstances, the interview will form a key component of the investigation to determine the use of AI tools on the balance of probabilities.

Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
Inappropriate use of digital technology to complete an assessment task (incl. use of generative artificial intelligence tools, paraphrasing or translation software)	<ul style="list-style-type: none"> Use of generative AI tools Student acknowledged the tool's but unpermitted by the unit coordinator 	<ul style="list-style-type: none"> Turnitin AI detection software confirms use of AI Minimal content generated by AI with substantial contribution from student 	<ul style="list-style-type: none"> Refer to faculty. 	Minor breach	<ul style="list-style-type: none"> Feedback or informal warning on permitted use within unit. Assessment to be marked on merit – unit coordinator to resolve within mark of assessment.
	<ul style="list-style-type: none"> Suspected use of AI tools Lack of proper citation or attribution of AI-generated content. 	<ul style="list-style-type: none"> No content identified by the Turnitin AI detection software Authorship report finds no evidence of a discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student provides a clear explanation of the process for completing their assignment and demonstrates an understanding of the content. Unable to substantiate allegation of AI use. 	No breach	<ul style="list-style-type: none"> No further action.
	<ul style="list-style-type: none"> Instances of unintentional or minor misuse of AI tools Lack of proper citation or attribution of AI-generated content. First-time breach with no significant impact on academic work or others. 	<ul style="list-style-type: none"> No or limited content identified by Turnitin AI detection software Unit coordinator's assessment of the submission suggest content is indicative of AI usage Authorship report finds limited evidence of discrepancy in the submission under investigation 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or cannot adequately explain their process of completing their assignment. 	Minor breach	<ul style="list-style-type: none"> Undertake development activity (AH-EM) Submit corrected submission to rectify the breach; or Minor mark reduction for the assessment and mark on merit for the genuine component of the assessment.

Guidelines for managing AI-related academic integrity breaches – 7 July 2023

Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
	<ul style="list-style-type: none"> Deliberate misuse of AI tools to gain unfair advantage Submission of AI-generated content with limited evidence of the student's original contribution and without proper acknowledgement or permission Substantial similarity to AI-generated content without adequate citation 	<ul style="list-style-type: none"> Moderate content identified by Turnitin AI detection software Authorship report finds no evidence of discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or is unable to adequately explain their process for completing their assignment or the content of the submitted work. If the student admits to engaging a third party or there is evidence to suggest contract cheating, these should be referred as potential misconduct. 	Major breach	<ul style="list-style-type: none"> Mark reduction for the assessment and mark on merit for the genuine component of the assessment Undertake development activity (AH-EM)
	<ul style="list-style-type: none"> Submission of content generated solely or primarily by software, which is presented as the student's own original content. Systematic or repeated use of AI tools to cheat or deceive Creation or distribution of AI models designed to facilitate academic dishonesty Significant impact on the academic work of others or the institution's integrity 	<ul style="list-style-type: none"> High volume of content identified by the Turnitin AI detection software; and/or Authorship report finds evidence of a discrepancy in the submission and/or potential indicators of contract cheating 	<ul style="list-style-type: none"> Refer to Registrar 	Potential Misconduct	Refer to Registrar

Guidelines for managing AI-related academic integrity breaches – 7 July 2023

Appendix 8 – Generative AI essentials for educators workshop slides

Appendix 7 – Guidelines for determining academic integrity breaches involving AI



Guidelines for determining breaches related to the use of generative AI tools

When determining academic integrity breaches involving AI, it is crucial to consider the severity of the breach and the specific circumstances surrounding each case. The severity of the findings and penalties will vary depending on the specific details of the AI-related integrity breach, including the affected proportion of the student's work and their own contribution and the student's academic record and previous breach history. Penalties should be applied consistently and fairly, with discretion for individual circumstances and the opportunity for students to present their responses. Feedback, informal warnings and penalties should educate on appropriate use of AI.

The Office of Educational Integrity will conduct a preliminary assessment to gather any supporting information and determine the appropriate decision maker to refer the case. In these circumstances, the interview will form a key component of the investigation to determine the use of AI tools on the balance of probabilities.

Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
Inappropriate use of digital technology to complete an assessment task (incl. use of generative artificial intelligence tools, paraphrasing or translation software)	<ul style="list-style-type: none"> Use of generative AI tools Student acknowledged the tools but unpermitted by the unit coordinator 	<ul style="list-style-type: none"> Turnitin AI detection software confirms use of AI Minimal content generated by AI with substantial contribution from student 	<ul style="list-style-type: none"> Refer to faculty. 	Minor breach	<ul style="list-style-type: none"> Feedback or informal warning on permitted use within unit. Assessment to be marked on merit – unit coordinator to resolve within mark of assessment.
	<ul style="list-style-type: none"> Suspected use of AI tools Lack of proper citation or attribution of AI-generated content. 	<ul style="list-style-type: none"> No content identified by the Turnitin AI detection software Authorship report finds no evidence of a discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student provides a clear explanation of the process for completing their assignment and demonstrates an understanding of the content. Unable to substantiate allegation of AI use. 	No breach	<ul style="list-style-type: none"> No further action.
	<ul style="list-style-type: none"> Instances of unintentional or minor misuse of AI tools Lack of proper citation or attribution of AI-generated content. First-time breach with no significant impact on academic work or others. 	<ul style="list-style-type: none"> No or limited content identified by Turnitin AI detection software Unit coordinator's assessment of the submission suggest content is indicative of AI usage Authorship report finds limited evidence of discrepancy in the submission under investigation 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or cannot adequately explain their process of completing their assignment. 	Minor breach	<ul style="list-style-type: none"> Undertake development activity (AHEM) Submit corrected submission to rectify the breach, or Minor mark reduction for the assessment and mark on merit for the genuine component of the assessment.

Guidelines for managing AI-related academic integrity breaches – 7 July 2023



Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
	<ul style="list-style-type: none"> Deliberate misuse of AI tools to gain unfair advantage Submission of AI-generated content with limited evidence of the student's original contribution and without proper acknowledgement or permission Substantial similarity to AI-generated content without adequate citation 	<ul style="list-style-type: none"> Moderate content identified by Turnitin AI detection software Authorship report finds no evidence of discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or is unable to adequately explain their process for completing their assignment or the content of the submitted work. If the student admits to engaging a third party or there is evidence to suggest contract cheating, these should be referred as potential misconduct. 	Major breach	<ul style="list-style-type: none"> Mark reduction for the assessment and mark on merit for the genuine component of the assessment Undertake development activity (AHEM)
	<ul style="list-style-type: none"> Submission of content generated solely, or primarily by software, which is presented as the student's own original content. Systematic or repeated use of AI tools to cheat or deceive Creation or distribution of AI models designed to facilitate academic dishonesty Significant impact on the academic work of others or the institution's integrity 	<ul style="list-style-type: none"> High volume of content identified by the Turnitin AI detection software; and/or Authorship report finds evidence of a discrepancy in the submission and/or potential indicators of contract cheating 	<ul style="list-style-type: none"> Refer to Registrar 	Potential Misconduct	Refer to Registrar

Guidelines for managing AI-related academic integrity breaches – 7 July 2023

- Appendix 8 – Generative AI essentials for educators workshop slides
The University’s learning and teaching blog, [Teaching@Sydney](#), is home to several critical articles that have been published since early 2023 to help staff understand and respond to generative AI [SD]. These include:
 - [What teachers and students should know about AI in 2023](#)
 - [Student-staff forums on generative AI at Sydney](#)
 - [Prompt engineering for educators – making generative AI work for you](#)
 - [Ten myths about generative AI in education that are holding us back](#)
 - [ChatGPT is old news: How do we assess in the age of AI writing co-pilots?](#)
 - [How can I update assessments to deal with ChatGPT and other generative AI?](#)
 - [Where are we with generative AI as semester 1, 2024 starts?](#)
 - [What to do about assessments if we can’t out-design or out-run AI?](#)
- The Education Portfolio created an AI in Education Community of Practice, engaging with over 250 members to build organic networks across the institution and share practice and experiences around teaching and assessing with generative AI [SD].
- The Education Portfolio and educational design teams within faculties run school- and faculty-wide sessions to introduce academic staff to the realities of generative AI and assessment. These include outlining the current state of AI technology that students have access to, the risks to academic integrity, the implications for learning outcomes and curriculum design, and the need to rethink assessment. These have played a critical role in shifting the culture away from denial or banning, towards acceptance and stronger appetite for assessment/curriculum redesign.
- In February 2024, the Education Portfolio hosted a sector-wide ‘AI in Education Symposium’ with 2,000 registrations from Australia and overseas. AI in Education is also now a main feature in the University’s annual ‘Sydney Teaching Symposium’, held in July each year. These symposia feature educators’ best practice approaches to applying generative AI.

Students

- The Education Portfolio and Library worked with a team of student partners to develop the [AI in Education](#) site [OP, LRS]. This team included students from equity backgrounds to ensure that diverse perspectives were accommodated. The aim of this site is to build familiarity around generative AI, its applications, and the rules around its use in learning and assessment. A link to this site is provided to all students as part of their unit outline documents.
- The Education Portfolio has developed a 20-minute activity that will be embedded in key first-year ‘transition units’, introducing commencing students to generative AI [OP] in all degrees. This activity is run by tutors in the context of a particular discipline and introduces students to key generative AI concepts such as its impact, use, ethics, academic integrity, and individual perspectives.
- The Library, in collaboration with its peer learning advisors, has run regular introductory generative AI sessions for students [LRS]. These cover information and digital literacies around generative AI, how they work, key applications for students, and guardrails for use.
- The Education Portfolio’s Learning Hub has run regular workshops for students on key generative AI-enabled tools for literature searching, structuring writing, and proofing work [LRS]. These are run to support

students with responsible use of AI in assessments, in consultation with unit coordinators.

- The Library is adapting its information and digital literacy support model to incorporate the University's two-lane approach to assessment. This will include an increased focus on lane 2 assessments and learning activities supporting those assessments [SD].

Future action plan

- The Education Portfolio, Library, and Research Portfolio will work to consolidate the suite of student-facing training and resources by the beginning of 2025. There will be regular reviews and continuous improvement activity to ensure training and resources remain relevant and responsive to generative AI's changes.
- In partnership with students from diverse backgrounds, the University updates the [AI in Education](#) website [OP, LRS] on an annual basis to ensure that information on the site is responsive to policy updates, technology improvements, and contemporary practices around AI [MRI]. The University will explore additional locations from which to link this website, such as from Canvas.
- The Education Portfolio will extend the 20-minute transition activity for first year units, building an extended drop-in module that can be adapted to disciplinary contexts. This module will cover what generative AI is and how it works, the ethical and integrity considerations when using it in education, how to use generative AI effectively, practical applications of generative AI tools for learning and assessment, and impacts of generative AI on the discipline, society, and the economy. This will build student familiarity with generative AI, especially in units with lane 2 assessments.
- The Education Portfolio will continue curating exemplars of good practice around practical generative AI application in teaching and assessment [SD]. A significant update of the existing resources is due by the end of 2024 to assist staff in preparing for assessment redesign, ahead of the timelines for moving towards lane 2 assessments. These resources will continue to be refreshed on a bi-annual basis [MRI].
- In collaboration with faculties, the Education Portfolio will continue to run professional development, assessment redesign consultation and faculty sessions [SD, CD] over the next 18 months to transition the University's assessments towards the two-lane approach. The significant majority of the University's courses and assessments are blended or on campus, with only a small selection available fully online. These online courses will receive extra support and oversight to ensure appropriately secured lane 1 assessments are designed to assure attainment of learning outcomes [LOA].
- Workshops for Honours/HDR students and research supervisors are currently under development in collaboration with the Research Portfolio and Library. These will cover academic integrity issues related specifically to theses and protection of research data, as well as applications for information literacy [ARI, RT, LRS].
- The University will continue engaging with students regularly to improve familiarity with generative AI [MRI, CGA]. Developments will be coordinated via the AI in Education Working Group and include initiatives such as hackathons and other events/channels where we can platform the student voice. Students will continue to form the fundamental role as partners in the development of our policies for assessment, learning and teaching.

- The Library and Learning Hub are collaborating on aligning co-curricular AI support offerings to students to ensure they receive coherent and consistent messaging about AI use. These groups are also developing referral pathways for unit coordinators to ensure they maximise use of curricular and co-curricular AI skills support and are directed to relevant University policy where needed.

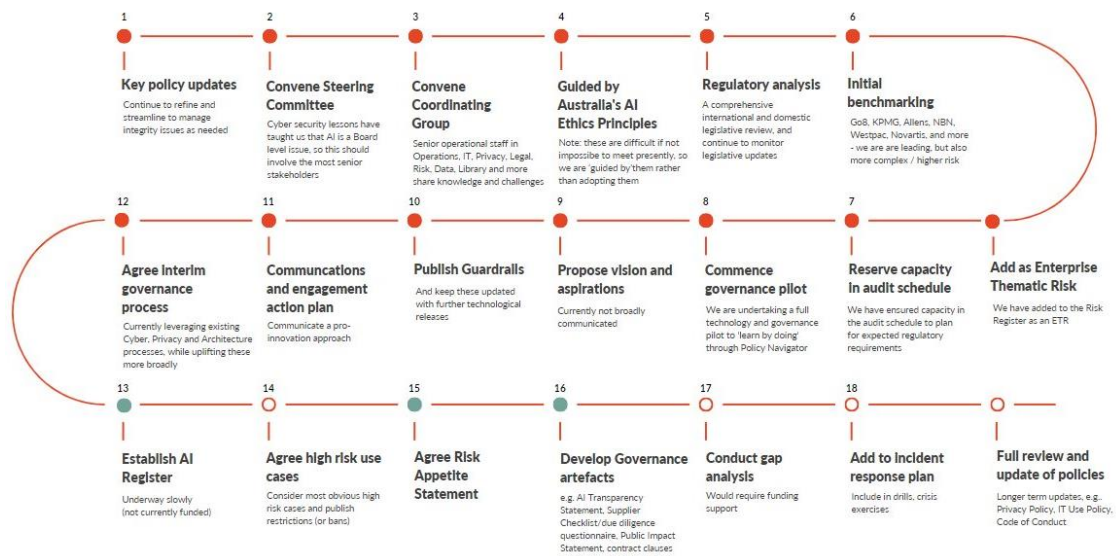
Principle 4: Fostering trust

The three action areas of rules, access, and familiarity are built on a foundation of trust. This trust is a triumvirate between students, teachers, and AI, and is based on integrity, honesty, security, transparency, openness, and human agency.

- **Radical transparency.** Our AI guardrails ([Appendix 3](#)) and aspirations ([Appendix 5](#)) highlight the need for open acknowledgement of AI use, by staff and students. As generative AI is still a relatively new concept, this overt transparency will help build trust and normalise the use of these tools in day-to-day work and study.
- **Secure computing environments.** The University's ICT team works with key vendors and internal stakeholders such as the Education Portfolio, Research Portfolio, and operations group to establish robust and secure generative AI infrastructure. This elevates trust by providing environments within which data is protected and experimentation can occur safely.
- **Control by, and visibility for, teachers.** The University's Cogniti AI platform allows teaching staff to steer and control their own 'AI agents' and see how students are interacting with them. This enhances the trust that students have in AI tools that are specifically designed to support their learning, and the trust that teachers have in being able to monitor and interpret student-AI interactions.

Appendices

Appendix 1 – Generative AI Governance Roadmap



Appendix 2 – Generative AI Steering Committee Terms of Reference

COMMITTEE	Generative AI Steering Committee
PURPOSE	The purpose of the Generative AI Steering Committee is to develop, and make recommendations about, a unified and coordinated leadership approach to generative artificial intelligence across the University.
TERMS OF REFERENCE	<p>The Generative AI Steering Committee will:</p> <ol style="list-style-type: none"> 1. (AI Strategy) Develop a coordinated institutional approach to generative AI across the full range of University activity, including teaching, research, and administration. 2. (Capacity building and expertise) Support and encourage staff, affiliates, and students to engage productively, effectively, and ethically with generative AI. 3. (Human centred culture) Support the University in taking a human-centred approach to developing and using generative AI. This approach will be grounded in our institutional values. It will serve and support the University community and our organisational culture. 4. (Governance) Encourage a cohesive and consistent approach to generative AI governance, risk, and compliance. To do this the Committee will oversee, and make recommendations about: <ol style="list-style-type: none"> a. identifying and assessing opportunities and risks presented by generative AI; b. developing a single, transparent, risk-based approach to generative AI; c. developing and communicating guidance and strategies to support staff, affiliated academics and students in understanding engaging with AI safely, ethically and responsibly; d. advising about cross-institutional roles, responsibilities and resources, including models for distribution of and equitable access to the University’s AI resources; and e. reviewing, and making recommendations about updates to, delegations, policies, and procedures as required. 5. (Reporting) Providing reports to the University Executive (UE) and the Senate Performance and Risk Committee (PARC) on its activities, recommendations, and outcomes.

COMMITTEE	Generative AI Steering Committee
CHAIR	Deputy Vice Chancellor (Education) Deputy Vice Chancellor (Research)
MEMBERSHIP	General Counsel Vice President, Operations Deputy Vice Chancellor (Education) Deputy Vice Chancellor (Research) Chief Information Officer Chief Risk Officer Chief of Staff Others as required
SECRETARIAT / COORDINATION	Director, Legal Operations
MEETINGS	Monthly
REPORTING	University Executive (UE) and the Senate Performance and Risk Committee (PARC)

Appendix 3 – Guardrails for Generative AI use at the University of Sydney

This is an extract from the Intranet page available at

<https://intranet.sydney.edu.au/strategy-governance/gen-ai/our-guardrails.html>

How you can act responsibly

You are a central component of our safe and responsible approach to generative AI. Please stay informed, ask questions and help us lead the way.

Whilst we encourage you to learn and experiment with generative AI in your work, you must follow our guardrails below.

The risks of not following these guardrails are serious. They include loss of confidentiality, privacy and intellectual property rights, and non-compliance with laws, research codes, funding rules, ethics approvals, University policies and journal policies.

Our guardrails

Be very careful what information you enter

It is important that you carefully manage information held by the University.

Public tools

You must never enter confidential, personal, proprietary or otherwise sensitive information into ChatGPT or other public cloud-based generative AI platforms:

- Once information is entered into a public cloud-based generative AI tool, it can be incorporated into the datasets that the tool uses to generate content for all users and could therefore be presented to another user at a later date (including those maliciously seeking to collect and aggregate it). This can compromise the security of that information – which can impact the University and others and breach the law.
- In public cloud-based services like ChatGPT, we also recommend that you [opt-out of having your data used for training future models](#).

Examples of confidential, personal, proprietary and sensitive information include:

- a person's name, birth date, address, identifying numbers or passwords
- other personal or health information (which are [broadly defined](#) under law)
- information that is intended for internal use only (including confidential or commercially sensitive information of the University or our partners)
- unpublished research data manuscripts or grant proposals of your team or [others' grant proposals](#)
- copyright materials, such as the Library's eResources, the University's proprietary teaching resources and student assessments.

University-endorsed tools

A broader range of data and information can often be entered into University-endorsed tools.

In these cases:

- you must manage information in accordance with the classifications set out in the University’s [Data Handling Standard](#); and
- for cloud-based tools, always check that you are using the ‘SECURED’ versions of these tools by logging in via OKTA.

See our list of [endorsed cloud-based tools](#).

Do not rely on the accuracy of outputs

Be aware that generative AI tools often produce incorrect information.

You remain responsible for your work. This means you must independently verify and edit AI-generated content to ensure the integrity, accuracy, and suitability of the output.

Generative AI tools use probabilistic models that are trained on enormous datasets to generate plausible new content; they use existing, uncorroborated content from many sources and can produce inaccurate, biased or creatively fictitious content. Repeated use of a generative AI tool can create different content each time it is run.

Using clear and specific prompts can improve the quality and relevance of AI-generated output. View some [practical examples of effective prompts here](#) and [here](#).

Openly acknowledge your use of AI

Where you create content using generative AI, model best practice by acknowledging this, e.g. “An AI generated image of DNA”.

A higher bar also applies to academic research and teaching work:

- Researchers using generative AI need to understand the proper way to attribute such work based upon the policy of their institution, grant funding body, or journal and relevant government agencies. Bear in mind these may change over time.
- Educators using generative AI at Sydney must model best practice for students by being transparent and clear with students about how they use the tools, acknowledging the author’s responsibility for the generated content.

Check for restrictions and conditions

Before using a generative AI tool, check that all relevant contracts (eg. funding agreements, data sharing agreements, commercial contracts), approvals (eg. your ethics protocols) or third-party stakeholders (eg. [journal conditions](#)) permit that use.

There are an increasing number of restrictions being placed on the use of generative AI tools, for example in the terms and conditions of open access materials, and by academic journals.

Report security incidents

It is important to report any cyber security and privacy security incidents as they happen, so that ICT and our Privacy team can quickly address any issues and mitigate risk exposure. You can [use this Service Now form to confidentially report issues](#).

Follow any further guidance specific to your work

- [Detailed guide for researchers: Generative AI and research at Sydney](#)
- [Guide to responding to generative AI in assessments](#)

Appendix 4 – Generative-AI: guidelines for researchers

This is an extract from the Intranet page available at https://sydneyuni.service-now.com/sm?id=kb_article_view&sysparm_article=KB0031813

The recent and rapid emergence of Generative-AI (Gen-AI) tools presents many opportunities for research. The benefit these tools may bring to research and contribute to the public good is being explored by researchers across all disciplines. While 'the sky is the limit' in what these new tools may offer, there are identified risks that must be managed so that Gen-AI tools can be used safely in research.

These [guidelines](#) provide a simple set of considerations that all researchers need to take into account before using Gen-AI tools in research.

What is Generative-AI?

Gen-AI refers to a set of algorithms, data and applications that analyse and generate text, images, code and other outputs based on models created from immense amounts of publicly available content on the internet (e.g. Wikipedia, open datasets, websites, image banks). Well-known Gen-AI tools include ChatGPT, Bing, Midjourney, Dall-E and Bard, but there are thousands more available and under development.

What are the top risks to consider when using Generative-AI in research?

1. Handling Sensitive Information

Where we choose to store and process our data – whether it be [on institutionally approved platforms](#), or via cloud-based solutions – is especially important when dealing with sensitive data. This could include personally identifiable or private information, [protected or highly protected research data](#), research IP such as unpublished paper drafts or grant applications, and confidential or contractually secret data. Sending these types of data to servers in other jurisdictions or analysing them using unapproved online services may violate agreements, contracts, university policy and state and federal law, so it is important to consider what you are sharing with these services.

Some Gen-AI model providers will retrain new versions of their models using our private data – so someone else may get a result with our sensitive info in it. Make sure to read the terms and conditions of any cloud-hosted service (especially Gen-AI) you are using with your data to ensure that your use of the service complies with university policy and state & federal privacy legislation. In online services like ChatGPT, make sure to [opt-out of having your data used for training future models](#). Note that even with this opt-out, sending some sensitive or personally identifiable information to a service like ChatGPT could constitute a privacy breach (which may carry heavy legal penalties), as people such as OpenAI employees may now have unauthorized access to it, and/or the information has left NSW jurisdiction to be processed by a server in the US.

Cloud services in other jurisdictions (including Gen-AI tools) are ok for some use cases such as looking at publicly available data, or where you are ok with the information becoming publicly available.

See also:

[University of Sydney Privacy Policy 2017](#),

[Research Data Management Procedures 2015](#)

[Cybersecurity Standard – Data Classification](#)

2. Research Integrity

Accuracy: The content created by a Large Language Model (LLM) can be creative and therefore sometimes inaccurate. It is important to independently verify the information gained from a large language model, just as it is important to check the source's trustworthiness when you do an internet search. The example of [a lawyer citing non-existent case law](#) shows what can happen if you don't independently verify your sources. If you are building a tool based upon. If you are building a research tool or workflow involving LLMs there are prompting techniques and architectures which you can use to increase accuracy beyond what you will get directly out of something like ChatGPT.

Accountability + Responsibility for accuracy and truthfulness in research – be that writing, data analysis or decision making – lies with the researcher using the AI and not with the AI itself. The researcher must ensure they are not lying, fabricating, misleading or misrepresenting, and they cannot wash their hands of this responsibility by blaming the AI.

Attribution: A researcher using Gen-AI needs to understand

- How to attribute this properly? E.g. “Alice and the AI worked together on this”
- How to cite the works/training data this is most derivative from? “The photos/articles/poems/proteins this draws upon the most are X, Y and Z” Is it possible to tell which pieces of training data contributed the most to a given answer, and therefore cite in your derivative work?
- Does this use align with the policy of your institution / funding body / journal / government agency? E.g.:
 - [ARC Discovery Project grant proposal reviews were done by ChatGPT](#) – a breach of confidentiality agreements that the assessor has signed on to, and [prompting a response from the ARC](#)
 - Nature Publishing Group allows use of Large Language models in published work but [disallows use of generative images or video](#)

Reproducibility: Gen-AI models such as ChatGPT are stochastic and as such their output is not always exactly the same for the same input. Sometimes we are ok with it being different – no two authors will write the same novel, and no two software engineers will write the same computer program. However, once written, the novel is the same for everyone who reads it, and the program will run (hopefully) the same every time. We probably don't want legal or medical advice, or a scientific analysis to come to a different conclusion based upon the same input but we probably do want creative output such as writing a poem, an essay or creating visual art to be different

most of the time. We can lower the temperature of the model, use truth-grounded architectures and track model versions to create reproducibility in use cases in which that is desirable.

3. Research Ethics

Bias and fairness

All models, research and work is biased in some way or another. It is important to consider what those biases are and be intentional about which of these are undesirable and we would like to change. For example, GPT-4 has been found to give different medical diagnoses and recommend different procedures based upon race and gender. AI models will make the human biases they were trained on faster, not different, unless we are careful about how we structure them, and what qualities we reward and penalize them for.

Models are heavily weighted towards the dominant cultural bias and language of the training data as well as the perspectives of the model developers. For example, 93% of the training data is in English for GPT3, the original version behind ChatGPT, this can bias it towards American cultural views. It is many times more expensive to use ChatGPT directly in non-English and especially minority languages as the tokenisation stage is less efficient.

Source of training data

The large foundational models that are used for Gen-AI applications require immense amount of data for training. These datasets are very large, and this scale makes ensuring that data is licensed properly and / or ethically sourced difficult if not impossible. Some models including ChatGPT and Google's Bard do not reveal much about their training data at all and so we cannot know what was in there. If a GenAI tool unintentionally plagiarizes improperly licenced data on your behalf this may expose you to liability. Additionally, many components of the Gen-AI product lifecycle (e.g. data labelling, content filtering) may involve ethically questionable, sanctioned, or illegal labour practices, including ones which qualify as modern slavery.

How can I use Generative-AI safely?

- **Ensure that the only content you are uploading to a cloud-based Gen-AI tool is material that can be shared safely with anyone and anywhere.**
 - Do not upload unpublished research findings into a Gen-AI tool. For example, using Chat-GPT to assist in writing a literature review is probably low risk, but writing a discussion, synthesis of research findings or entire thesis is high risk.
 - Do not upload research data that contains confidential or sensitive information into a Gen-AI tool (this includes personally identifying information of study participants, commercially sensitive information, and culturally or environmentally sensitive information – these would be classed as protected or highly protected data).
 - Do not upload your personal data into a Gen-AI tool.
 - Do not upload data to or make use of Gen-AI tools if your data has geographic restrictions in place. Gen-AI tools, including specifically

OpenAI's GPT products, are usually running on servers housed outside of Australia. Be aware of the geographic restrictions around your data before using Gen-AI.

- Copyright material owned and/or controlled by the University, including published research available in [library databases](#), must not be used as an input to Gen-AI unless permitted to do so.
- **Validate that the output from Gen-AI tools is accurate and attributable.**
 - Gen-AI tools use probabilistic models that are trained on enormous datasets to generate plausible new content; they use existing, uncorroborated content from many sources and can produce inaccurate, biased or creatively fictitious content. Repeated use of a Gen-AI tool can create different content each time it is run. If using Gen-AI tools to generate research content, researchers need to ensure that outputs are accurate, factual, able to be attributed properly, and that research results are replicable.
- **Ensure that you are allowed to use Gen-AI in your research.**
 - Before using a Gen-AI tool for research, check that your ethics protocols, funding agreements, data sharing agreements or commercial contracts permit use of Gen-AI, and/or cloud-based services in other jurisdictions in general.
- **Check the Gen-AI policy of the journal that you plan to publish in.**
 - Some conferences, journals, and societies limit and/or prohibit use of Gen-AI. Carefully consider if the use of Gen-AI will have sufficient positive impact on your research to outweigh any potential limitation in publication.
- **Observe confidentiality when using Gen-AI**
 - Protect the IP of researchers and maintain professional standards: **do not use Gen-AI when [peer-reviewing grant applications](#), manuscripts and publications**
- **Ensure your students are aware.**
 - If you are supervising a research student (HDR or Honours), make sure that they are aware of the risks of using Gen-AI in their research, including proposal and thesis writing.

See also:

[**Generative-AI: guidelines for researchers**](#)

[**Australian AI Ethics Principles**](#)

[**AI in Education Canvas Site**](#)

[**HTI report on Governance and Legal implications of AI adoption in Australia.**](#)

Queries? Submit a [**General Research Enquiry form**](#).

Appendix 5 – Dynamic Generative AI Roadmap



Generative Artificial Intelligence *Dynamic Roadmap*

Vision

To be a leader in the safe and responsible use of generative AI in the international higher education sector.

To augment and redefine our educational and research experiences, and enhance the productivity of our operations, allowing Sydney students and staff to focus on advancing **human endeavours**.

Guiding principles

We are guided by [Australia's AI Ethics Principles](#):

Human, societal and environmental wellbeing	AI systems should benefit individuals, society and the environment
Human-centred values	AI systems should respect human rights, diversity, and the autonomy of individuals
Fairness	AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups
Privacy protection and security	AI systems should respect and uphold privacy rights and data protection, and ensure the security of data
Reliability and safety	AI systems should reliably operate in accordance with their intended purpose
Transparency and explainability	There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI, and can find out when an AI system is engaging with them
Contestability	When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system
Accountability	People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled

Aspirations

Sydney takes the following aspirational positions on generative AI in the higher education sector:

Cross functional leadership	AI has applications in all facets of University work	<ul style="list-style-type: none"> We recognise that AI has many applications throughout the range of work performed by individuals and groups at the University. AI will have uses in teaching, learning, assessment, research, operations, and productivity. In these contexts, we will be guided by our principles, developing the comfort and caution of accepting work that blends input from human and machine.
	Human agency, expertise, and accountability are central	<ul style="list-style-type: none"> We recognise that human agency and expertise are paramount. Students must develop and demonstrate disciplinary skills and competencies, and staff expertise must guide this. We recognise that the role of AI will be different at varying levels of skill development, and we will use AI appropriately to this end. We will ensure that humans are appropriately accountable for AI-assisted decisions.
Human-centred	AI must benefit the University and its community	<ul style="list-style-type: none"> Development and application of AI at the University will respect diversity and humanity, be inclusive and accessible, and should seek to close equity gaps. AI-enabled systems and approaches will seek to benefit students, staff, the University community, and the work we do together.
	We engage productively and responsibly with AI	<ul style="list-style-type: none"> We will ensure that we engage appropriately with AI instead of ignoring it. We will do this in ways that deeply consider privacy, intellectual property, legal, copyright, data security, data integrity, and ethical issues around its development and use in all facets of University work.
	Where AI is used, it is transparent and documented	<ul style="list-style-type: none"> We are an institution committed to integrity, information and digital literacy. When we use AI we want to make sure humans are in control for high stakes decisions, and we want to make sure that AI contributions are transparent in all situations. We are committed to developing appropriate processes to allow all members of the university to contest the use and outputs of generative AI where possible
Safe and responsible	Our staff and students will model the use of AI	<ul style="list-style-type: none"> As leaders in education and research, we have a responsibility to demonstrate to students and the wider community how to engage appropriately with AI. Our staff will model the productive and responsible use of AI, and our students will become experts in its use.
	AI-human collaborations are normalised	<ul style="list-style-type: none"> As AI becomes increasingly embedded into the software and processes that we use day-to-day, as well as those in industry and the community, we will gradually normalise the nature of AI-human collaborations. This means together developing the comfort and caution of accepting work that blends input from human and machine.
Building capacity		

We will evaluate and periodically update our AI positions to keep pace with rapidly evolving AI technologies and trends.

Appendix 6 – The AI x Assessment menu

This menu approach recognises that there are many (and constantly expanding) ways in which students might use AI in the process of completing an assessment. The menu approach is preferred because it emphasises choice and suitability, as opposed to a traffic light or assessment scale approach which suggests that one can restrict or control AI use in unsecured assessments (one cannot). The menu analogy also emphasises the role of the educator in guiding students' choice of productive and responsible engagement with AI, much like a maître d' would guide diners' choice of culinary experiences.

As a critical friend – Soups

- Suggest analyses
- Provoke reflection
- Provide study/organisation tips
- Practicing

Getting started – Entrees

- Suggesting structure
- Brainstorming ideas

Engaging with literature – Bread service

- Suggesting search terms
- Performing searches
- Summarising literature
- Identifying methodologies
- Explaining jargon
- Fixing reference list

Generating content – Mains

- Writing some text
- Making images, video, audio
- Making slidedecks

Analyses – Lighter mains

- Performing analyses of data, text
- Suggesting counterarguments

Editing – Coffees

- Editing tone
- Improving clarity and readability
- Fixing grammar
- Shortening

Feedback – Desserts

- On all of the above elements
- Specifically on rubric criteria

Appendix 7 – Guidelines for determining academic integrity breaches involving AI



Guidelines for determining breaches related to the use of generative AI tools

When determining academic integrity breaches involving AI, it is crucial to consider the severity of the breach and the specific circumstances surrounding each case. The severity of the findings and penalties will vary depending on the specific details of the AI-related integrity breach, including the affected proportion of the student's work and their own contribution and the student's academic record and previous breach history. Penalties should be applied consistently and fairly, with discretion for individual circumstances and the opportunity for students to present their responses. Feedback, informal warnings and penalties should educate on appropriate use of AI.

The Office of Educational Integrity will conduct a preliminary assessment to gather any supporting information and determine the appropriate decision maker to refer the case. In these circumstances, the interview will form a key component of the investigation to determine the use of AI tools on the balance of probabilities.

Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
Inappropriate use of digital technology to complete an assessment task (incl. use of generative artificial intelligence tools, paraphrasing or translation software)	<ul style="list-style-type: none"> Use of generative AI tools Student acknowledged the tools but unpermitted by the unit coordinator 	<ul style="list-style-type: none"> Turnitin AI detection software confirms use of AI Minimal content generated by AI with substantial contribution from student 	<ul style="list-style-type: none"> Refer to faculty. 	Minor breach	<ul style="list-style-type: none"> Feedback or informal warning on permitted use within unit. Assessment to be marked on merit – unit coordinator to resolve within mark of assessment.
	<ul style="list-style-type: none"> Suspected use of AI tools Lack of proper citation or attribution of AI-generated content. 	<ul style="list-style-type: none"> No content identified by the Turnitin AI detection software Authorship report finds no evidence of a discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student provides a clear explanation of the process for completing their assignment and demonstrates an understanding of the content. Unable to substantiate allegation of AI use. 	No breach	<ul style="list-style-type: none"> No further action.
	<ul style="list-style-type: none"> Instances of unintentional or minor misuse of AI tools Lack of proper citation or attribution of AI-generated content. First-time breach with no significant impact on academic work or others. 	<ul style="list-style-type: none"> No or limited content identified by Turnitin AI detection software Unit coordinator's assessment of the submission suggest content is indicative of AI usage Authorship report finds limited evidence of discrepancy in the submission under investigation 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or cannot adequately explain their process of completing their assignment. 	Minor breach	<ul style="list-style-type: none"> Undertake development activity (AHEM) Submit corrected submission to rectify the breach, or Minor mark reduction for the assessment and mark on merit for the genuine component of the assessment.

Guidelines for managing AI-related academic integrity breaches – 7 July 2023



Allegation Type	Description of conduct	OEI review	Investigation process	Finding	Outcome
	<ul style="list-style-type: none"> Deliberate misuse of AI tools to gain unfair advantage Submission of AI-generated content with limited evidence of the student's original contribution and without proper acknowledgement or permission Substantial similarity to AI-generated content without adequate citation 	<ul style="list-style-type: none"> Moderate content identified by Turnitin AI detection software Authorship report finds no evidence of discrepancy in student submission 	<ul style="list-style-type: none"> Refer to faculty – allegation issued with a request for a meeting. Determination based on the evidence produced by the student as part of their response. The student admits to using AI tools or is unable to adequately explain their process for completing their assignment or the content of the submitted work. If the student admits to engaging a third party or there is evidence to suggest contract cheating, these should be referred as potential misconduct. 	Major breach	<ul style="list-style-type: none"> Mark reduction for the assessment and mark on merit for the genuine component of the assessment Undertake development activity (AHEM)
	<ul style="list-style-type: none"> Submission of content generated solely or primarily by software, which is presented as the student's own original content. Systematic or repeated use of AI tools to cheat or deceive Creation or distribution of AI models designed to facilitate academic dishonesty Significant impact on the academic work of others or the institution's integrity 	<ul style="list-style-type: none"> High volume of content identified by the Turnitin AI detection software; and/or Authorship report finds evidence of a discrepancy in the submission and/or potential indicators of contract cheating 	<ul style="list-style-type: none"> Refer to Registrar 	Potential Misconduct	Refer to Registrar

Guidelines for managing AI-related academic integrity breaches – 7 July 2023

Appendix 8 – Generative AI essentials for educators workshop slides

The following table provides a detailed overview of the 24 slides in the workshop, organized by their content focus:

Slide #	Topic	Key Content
1	Generative AI essentials for educators	Welcome. We will start at 5 past.
2	Demystifying large language models	Visual explanation of LLMs.
3	Demystifying large language models	Visual explanation of LLMs.
4	Demystifying large language models	Visual explanation of LLMs.
5	Neural networks	Diagram of a neural network.
6	Neural networks	Diagram of a neural network.
7	Neural networks	Diagram of a neural network.
8	Neural networks	Diagram of a neural network.
9	Implications of this understanding	LLMs are not a database, hallucinations are expected, and the way you 'converse' matters.
10	Implications of this understanding	LLMs are not a database, hallucinations are expected, and the way you 'converse' matters.
11	Generative AI is like...	Comparison of generative AI to a search engine.
12	AI and assessment	Introduction to AI in assessment.
13	The reality of generative AI and assessment	Real-world examples of AI in assessment.
14	Two-lane approach	Lane 1: Assessment of skills; Lane 2: Human-AI collaboration.
15	Example of two-lane approach	Example of a two-lane approach in a course.
16	Another example of two-lane approach	Another example of a two-lane approach in a course.
17	Example of a lane 1 assessment	Interactive oral assessment example.
18	AI x assessment menu	Menu of AI-related assessment options.
19	Of lanes and menus	Discussion of lanes and menus in assessment.
20	AI in assessment	AI in assessment overview.
21	Using available generative AI tools at Sydney	Approved tools: Microsoft Copilot, Copilot, etc.
22	AI tools available to staff and students	Approved tools and other AI tools.
23	Quick demos: Copilot for web and in Edge browser	Demos of Copilot for web and in Edge browser.
24	Prompting AI: key terms	Key terms for prompting AI: Prompt, Model, Completion.

Prompting AI: key terms

Context windows: how much text the model can take into account

- GPT-3.5: ~3000 words
- GPT-4: ~8000 words
- GPT-4o: ~96,000 words
- Claude 3: ~150,000 words
- Gemini 1.5: ~1,000,000 words

Prompting AI: Structuring prompts through 'RTRI'

- Role (act as...)
- Task
- Requirements (include, contain, be...)
- Instructions

Example: Generate practice Q&A

Use a 3-5 word leading and 5-7 word question format.

Make 2-3 more extra questions that test secondary or primary student understanding of a component of the following topic:

- Chemistry
- Geography
- Government

The AI will generate 10-15 questions. They will be based on the content and format you provide. Use the topic together with your feedback to help students connect ideas together in your feedback, and to ensure the AI generates the correct type of questions. Feedback of good questions will be provided.

That's it! Enjoy your questions!

How can AI help teachers and students?

Student perspectives haven't changed much

- Awareness and acceptance
- Careless usage
- Supplement to teachers
- Saving time
- Variable skill
- Wanting guidance from the university

ChatGPT and other AI learning

Good learning, teaching, assessment hasn't changed

Learning & teaching

- Practice
- Socratic questioning
- Analogies
- Concrete examples
- Building relationships
- etc.

Assessment

- Authentic assessment motivates learning
 - Realistic tasks and contexts
 - Cognitive challenge
 - Helping students judge quality
- Feedback

How can generative AI help you with these?

Example: Suggesting analogies

You are an expert tutor in university-level microeconomics.

Come up with three creative analogies to explain the complexity of market structures, especially focusing on the concepts of market power, barriers to entry, and product differentiation.

Example: Drafting case studies

Act as an expert toxicologist and pharmacologist with extensive experience in working with clinical cases.

Write a case study of an adverse drug reaction to methotrexate. Include detailed information about the patient, symptoms, the drug, timing and onset, and clinical presentation.

Also write a suggested analysis of the case study that students should provide, including management and interventions.

Example: Perspectives generator

Give 5 different perspectives on the Great Resignation from employees and employers who have very different cultural backgrounds. Write these in the form of short interview quotes from the perspective of these people.

...

Explain why these perspectives stem from their cultural backgrounds.

Example: Discussion topic suggestions

Come up with 5 examples of discussion prompts to get university students thinking about key concerns that might affect a new humanitarian engineering project. Base these prompts on the following topics which they are interested in:

- Social media
- Climate change
- Part-time work

Example: Ideation assistant

Look up modern perspectives on commercial architecture. Predict what architectural designs may be popular. Describe these succinctly. Use Image Creator to render a few of these descriptions.

Example: Assessment ideas generator

Use AI to generate 5-10 creative assessment ideas. Support 1-2 ideas, including a question and answer, to help you develop a question and answer. Support 1-2 ideas, including a question and answer, to help you develop a question and answer. Support 1-2 ideas, including a question and answer, to help you develop a question and answer.

Example: Rubric drafting assistant (link)

When Your Technical Skills for Edged, You Really Will Matter

What are our core capabilities as humans? If we remove that question from a piece of paper about what's left for people in the age of AI, we can end up conceiving a diminished view of human capability. Instead, it's critical for us all to start from a place that recognises what's possible for humans in the age of AI.

Thank you

Q&A

Q&A