De La Salle University

Policies on Generative Artificial Intelligence in Higher Education

I. Purpose

The rise of generative artificial intelligence (AI) is transforming education in various ways. Generative AI is a type of artificial intelligence that can create human-like output such as text, images, audio, and video after being trained on massive amounts of data. Despite exhibiting many signs of intelligence, generative AI models are probabilistic, prone to hallucinations, and run the risk of weakening human cognitive skills. Therefore, the use of generative AI tools necessitates careful oversight and responsible implementation.

De La Salle University's advisory on the use of generative AI in academia in September 2023[1] highlights the need to be receptive towards generative AI's potential benefits while also being mindful of its risks. It also encourages the responsible use of generative AI in teaching, learning, and research through augmentation, rather than replacement, of human output. This policy aims to support the position of the university towards generative AI by (1) providing a unified but flexible structure to the integration of generative AI in teaching, learning, and research, (2) ensuring that the use of generative AI is responsible and aligned with Lasallian values, and (3) continuing a critical engagement with generative AI and its effects on the various aspects of life to maintain learner agency.

II. Scope

This policy applies to the use of generative AI in De La Salle University for teaching, learning, and research. It applies only to generative AI, not AI in general. This policy document is based on various international references, inputs from focus group discussions and workshops held with the Lasallian faculty members, and inputs from the University Student Government Legislative Assembly and the Student Discipline Formation Office.

III. Definitions

Key Terms

- **University** refers to De La Salle University.
- **Student** refers to those enrolled in any subject or course in the University, regardless of the number of units enrolled; those included in the official student list of the Office of the University Registrar (OUR), whether enrolled, on official leave of absence (LOA), on residency, or absence without leave (AWOL); and those allowed to graduate from the University but has not yet secured the required clearances from the University.
- **Faculty** refers to all teaching personnel employed by the University, whether full-time or part-time, who are responsible for delivering academic instruction and assessing student performance.
- **Policy** refers to a formal set of principles and regulations established by the University to guide decision-making and ensure consistency in academic and institutional practices.

[1] Office of the Provost (2023), Advisory on the Use of AI in Academia. Released via Help Desk Announcement

- **Guidelines** refer to recommendations or best practices provided by the University to assist students, faculty, and staff in interpreting and applying policies effectively.

Course-Related Terms

- **Syllabus** refers to the official document that outlines the structure, objectives, policies, and grading system of a course. It serves as a guide for both students and faculty, detailing the course components, assessment tasks, learning plan, and expectations for academic conduct.

- **Grading System** refers to the structured method outlined in the course syllabus used to evaluate student performance. It is composed of several components, which refer to specific graded elements of a course that contribute to a student's final grade. These components, as defined in the course grading system of the syllabus, include but are not limited to quizzes, projects, exams, and other assessment tasks, each assigned a specific weight or point value.

- **Course Assessments** refers to any output submitted for academic purposes, including but not limited to essays, research papers, coding assignments, presentations, and projects.

- **Course Activities** refer to structured learning experiences designed to engage students in the course material. These activities are primarily formative in nature, aiming to reinforce understanding and skill development. They may include lectures, discussions, group work, hands-on exercises, problem-solving sessions, and other instructional methods.

AI-Related Terms

- **Artificial Intelligence (AI)** refers to computer systems that can perform tasks that require intelligence. There are many different types of AI. Some types of AI make use of predefined rules to accomplish their tasks, while some types of AI rely on learning patterns from data. Some AI systems are designed to specialize in a singular, specific task, while some are designed to handle a wider range of tasks.

- **Generative AI** refers to a specific type of artificial intelligence that can produce human-like content such as text, images, audio, and videos after being trained on large amounts of data through a process called machine learning. It is important to make a distinction between generative AI and AI in general. The latter is a more general term used to refer to any computer system that can perform intelligent tasks, which may or may not follow a probabilistic approach.

- **Generative AI models** are the underlying computer models that generate content. These models, based on neural network architectures such as transformers, are trained on large amounts of data until they can mimic human output. Examples of generative AI models are GPT-4o (a large language model that powers ChatGPT), DeepSeek (another large language model), and DALL-E (a text-to-image model that OpenAI uses in their image generation tools).

- **Generative AI tools** refer to software or services that make use of generative AI models. Examples of generative AI tools include general-purpose chatbots (ChatGPT, Gemini, Claude, DeepSeek), image generation services (Midjourney), video generation services (Sora, Goku), and research assistants (Consensus, Perplexity)

- **Hallucinations** refer to misleading, false, or nonsensical responses by generative AI models presented as facts. They occur as a side-effect of the probabilistic approach used by generative AI models in their training.

IV. Policy Details

a. General Conduct in Using Generative AI

Rationale. Adoption of generative AI tools is already high among both students and faculty members of the University[2]. As a Catholic University that prepares its students to become achievers for God and country, it is part of our mission to ensure that the usage of such disruptive technologies aligns with the Lasallian core values of faith, service, and communion, while at the same time advance the attainment of the expected Lasallian graduate attributes (ELGA)[3] among students.

- The following principles shall govern the conduct of the members of the academic community in using generative AI tools:

Principle	Expected Lasallian Graduate Attributes
Transparency must be observed when using generative AI. Disclosure is required whenever generative AI is used to assist in the preparation of any material.	Effective Communicator Engaged and Service-Driven Citizen Virtue-Guided Decision-Maker
Human agency must be upheld when generative AI is used for any purpose. No decisions should be made based solely on generative AI output, and there should be an awareness that outputs of generative AI may bias or skew human evaluation or interpretation.	Culturally Sensitive Critical and Creative Thinker Technically Proficient Professional Engaged and Service-Driven Citizen Virtue-Guided Decision-Maker
Accountability should accompany every use of generative AI. Outputs of generative AI must undergo critical human validation and synthesis before they are incorporated into any work or decision, and the author or creator shall ultimately be solely accountable for the final output.	Culturally Sensitive Critical and Creative Thinker Technically Proficient Professional Engaged and Service-Driven Citizen Virtue-Guided Decision-Maker
The use of generative AI should be **beneficial**. Generative AI should not be used for malicious or harmful purposes, including but not limited to: disinformation, deception, cheating, fraud, bullying, harassment, and the erosion of the quality of teaching, learning, and research.	Virtue-Guided Decision-Maker Engaged and Service-Driven Citizen
The use of generative AI should be **aligned** with the University's educational goals, especially the practice of critical thinking, active engagement with subject matter for understanding, academic integrity, and the acquisition of lifelong learning habits.	Culturally Sensitive Critical and Creative Thinker Reflective Life-Long Learner and

[2] Based on the results of the university-wide survey on AI perspectives and adoption, conducted in De La Salle University in August to October 2024.
[3] https://www.dlsu.edu.ph/colleges/gcoe/academic-departments/elga/

	Competent Self-Nurturer Virtue-Guided Decision-Maker
Compliance with existing regulations and policies, including those that pertain to academic honesty, research ethics, and data privacy, must be observed when using generative AI.	Engaged and Service-Driven Citizen Virtue-Guided Decision-Maker
The use of generative AI shall be accompanied by personal acknowledgment and **reflection** of its limitations, including the ethical issues that surround it. Amidst the presence of challenging issues surrounding generative AI, some of which are still uncharted (e.g., environmental impact, data bias), the community must place importance on education and awareness to promote active participation and citizenship in future conversations or efforts on its development and governance.	Service-Driven Citizen Reflective Life-Long Learner and Competent Self-Nurturer Virtue-Guided Decision-Maker

b. Disclosure Requirements

Rationale. Transparency is one of the key principles of responsible generative AI use. Although current policies on academic honesty cover transparency, there is a need to clarify transparency in the context of generative AI tools. There is currently no universal standard on when and how generative AI use should be disclosed. Therefore, there is a need to set a minimum disclosure standard for the University with the goal of promoting a culture of honesty and transparency in generative AI use.

- For both faculty and students, the use of generative AI in the production of any material, presentation, or submission, must be accompanied by a written disclosure statement.

- The disclosure statement must contain:

    1. The specific generative AI tool/s or services that were used
    2. A clear, concise description stating how the tool was used and the extent of influence it had on the final output.

    Here are some examples of disclosure statements:

ChatGPT was used to improve the writing style of this paper. I manually checked and validated the resulting improvements.
DeepSeek was used to generate an initial outline for this presentation. The contents, however, were entirely created by us.
The core idea presented in this paper was heavily inspired by a response from Gemini during brainstorming sessions, which was [describe generative AI response]. We took this idea and [explain how the idea was tweaked and incorporated to the work].

- A separate disclosure statement is not required anymore if the disclosure is already embedded within the output itself (e.g., if the generative AI was used as part of a research methodology, and this was already explained in the Methodology section of the manuscript). Likewise, a separate disclosure statement is not required anymore if the type

of generative AI use is already obvious from the specifications of the activity (e.g., the activity requires students to use generative AI for a specific purpose).

- Instructors may include additional disclosure requirements for students in specific course activities or assessments. These must be properly communicated to the students upon the release of the activity or assessment specifications. Examples of additional disclosure requirements are: (1) disclosure of relevant interaction logs of the student with the generative AI (e.g., prompts used, response of the AI), and (2) compliance with a specific format for citation (e.g., APA[4], MLE[5], or a custom format for the course).

- In research, students and faculty must comply with the generative AI disclosure requirements of the target venues for publication.

c. Declaration of Generative AI Use Policy in Course Syllabi

Rationale. The introduction of generative AI in education necessitates a thoughtful consideration of how such technologies impact the attainment of target learning outcomes. The declaration of generative AI policies in course syllabi aims to (1) clearly communicate to students the expectations pertaining to the use of generative AI in each course in relation to the course objectives, and (2) encourage faculty members to think more deeply about the potential benefits and risks of generative AI with respect to the knowledge, skills, and values that students are expected to develop.

- Starting Term 1, 2025-2026, every course must include an explicit declaration of AI usage policies in a dedicated section of the syllabus with the heading *Generative AI Use Policy*. This declaration must include the following:

  1. The **usage policy level** for **each component in the grading system** of the course
  2. A clear and concise **Generative AI Use statement** explaining the usage policies for the course, including the rationale for the said policies

- The table below shows the different usage policy levels.

Usage Policy Level	Description
Free to Use	Generative AI can be used for any purpose, provided that transparency and accountability are observed.
Allowed in Specific Contexts	Generative AI can be used only for specific purposes defined in the syllabus, provided that transparency and accountability are observed.
Banned	Generative AI is banned for use in any submission.

- A usage policy must be defined for each component in the grading system. Components that do not involve submissions (e.g., attendance or recitation) are not required to have a usage policy. For example, if a course's grading system is broken down into 20% class activities, 30% final exam, and 50% project, a possible declaration would be to say that the generative AI usage policy is "free to use" for the class activities and the project, but "banned" for the exams.

[4] https://apastyle.apa.org/blog/how-to-cite-chatgpt
[5] https://style.mla.org/citing-generative-ai/

- In cases where the usage policy is "allowed in specific contexts", it must include a clear explanation on which purposes or contexts generative AI is allowed. The explanation can be embedded within the Generative AI Use Statement, or alongside each grading component. If there are ambiguities about a particular use, students must first obtain permission from the instructor.

- Individual assessments can override the usage policy for a grading system component. These must be properly communicated in the syllabus or by the instructor upon the release of the assessment specifications.

- Appendix A contains guidelines in writing the *Generative AI Use Policy* section in the syllabus (see Appendix A). Templates are also provided, which can be modified freely to suit the specific needs of each course.

- Violations of the AI usage policy shall be treated as academic dishonesty (see Section IV-d for more details).

d. Handling of Generative AI-Related Academic Dishonesty

Rationale. Cheating is one of the most worrying concerns of both students and faculty members in De La Salle University when integrating generative AI in education[6]. As a research university, it is important to uphold academic integrity even in the face of disruptive technologies. Policies and guidelines are needed to ensure that the university's approach in handling academic dishonesty linked to generative AI use is consistent, fair to both faculty and students, and grounded on facts and evidence.

- For students, the following shall be considered as generative AI-related academic dishonesty:

    1. Using generative AI in producing a submission in a way that is prohibited by the usage policy defined for that assessment, even if it was disclosed (see section IV-c for more details on usage policy)
    2. Failing to properly disclose the use of generative AI in a submission (see section IV-b for more details on proper disclosure)

- Cases of generative AI-related academic dishonesty shall be treated the same way as academic dishonesty in general. All provisions in the Student Handbook[7] regarding academic dishonesty shall also be applied in these cases. Section 4.15 of the Student Handbook contains the University's policy statement on academic honesty. It contains the following statements:

    "As a Catholic University that prepares its students to become professionals and be the resource for Church and nation, De La Salle University puts premium on academic honesty. Students' academic requirements, such as assignments, term papers, computer programs, projects, and thesis papers, should be their own work. They must distinguish between their own ideas and those of other authors. Students must cite references, direct quotes, and other sources (including data obtained from tables, illustrations, figures, pictures, images, and video) following the prescribed format of the discipline."

[6] Based on the results of the university-wide survey on AI perspectives and adoption, conducted in De La Salle University in August to October 2024.
[7] De La Salle University Student Handbook 2021-2025 (https://www.dlsu.edu.ph/wp-content/uploads/pdf/osa/student-handbook.pdf)

In this case, any generative AI tool shall be treated as an external source of ideas, and the above principles shall also apply to such tools.

Additionally, the following provisions in the Student Handbook shall also apply to generative AI-related academic dishonesty:

1. Academic dishonesty is considered as cheating, and is a major offense (Section 5.3.1.1)
2. Faculty members have the right to demand the presentation of a students' ID, to give a grade of 0.0, and to deny admission to class of any student caught cheating. Students should immediately be informed of their grades and barred from further attending their classes (Section 5.2.4.1)
3. In cases of alleged cheating, the faculty member should report the incident to the Student Discipline Formation Office (Section 9.3)
4. Any form of academic dishonesty is considered a non-mediatable offense (Section 5.2.3.4.9)
5. To qualify for the Dean's Honors List, a student must have not been found guilty of cheating or academic dishonesty within the term (Section 11.1.4)
6. To qualify for the University Honors List, a student must have not ever been found guilty of cheating or academic dishonesty within the term (Section 11.2.4)
7. A student found guilty of cheating or academic dishonesty is disqualified from graduating with honors (Section 12.4.4)

- Faculty can use AI detector tools (Turnitin, GPTZero, etc.) for flagging suspicious submissions. However, results of such tools cannot be used as the sole basis for establishing generative AI-related academic dishonesty, as these tools can produce both false positives and false negatives.

- Faculty cannot refuse submissions or impose grade penalties solely because of AI detector results. For AI detector tools that produce percentage values such as Turnitin, this policy shall apply regardless of any threshold value, since even high percentage outputs are not a definitive indication of cheating.

- Faculty can impose penalties and file academic dishonesty cases if suspicious submissions are accompanied by additional pieces of evidence that indicate students' non-ownership of the submitted work. This may be one or more of the following:
    - Failure to answer questions in a way that is consistent with the submitted work
    - Failure to explain the thought processes used to produce the submitted work
    - Failure to demonstrate the same skills in the submitted work in a controlled environment
    - Traces of hallucinations in the submitted work that are unlikely to be due to human error (e.g., citing a non-existent source)
    - Drastic difference in the writing style of the student compared to previous submissions (must be used with caution, as students' writing styles can slowly change over time)
    - Large similarity with the submissions of other students despite the wide range of possible ideas for addressing the assessment

  Such grade penalties must be clearly communicated to the student. The communication must include the specific indicators of non-ownership to justify the penalty.

- Appendix B contains more information and guidelines on AI detector tools, their strengths, and limitations.

- The above policies on AI detector tools only apply to AI detection, not plagiarism detection. Some tools like Turnitin offer both services, so it is important to distinguish between the two. Plagiarism detection is more reliable because it gives you the sources where the text was copied from, making it simpler to validate.

- The lack of a guaranteed way to distinguish human output from AI output remains a big challenge. In navigating this landscape, faculty are encouraged to shift their focus from strictly policing generative AI use to a more holistic view that considers why the use of generative AI is harmful in a particular context. Cheating involving generative AI shall be viewed through the broader frameworks of academic honesty[8] and shall be addressed with a combination of both reactive and preventive measures.

- Faculty members are encouraged to implement the following preventive measures to help deter generative AI-related academic dishonesty:

    1. Communicate clearly the expectations on AI use to students
    2. Design AI-resistant assessments
    3. Redesign rubrics by putting more emphasis or weight on items that would be difficult to achieve if there was an over-reliance on generative AI
    4. Rethink learning objectives in the age of generative AI
    5. Combine take-home assessments with face-to-face assessments

    Appendix C contains more detailed guidelines and recommendations for preventing generative AI-related academic dishonesty.

e. Data Privacy and Intellectual Property

Rationale. As generative AI services process large amounts of data, there is a risk of exposing sensitive information, including personal data, confidential research, and institutional records. To mitigate these risks, the University must establish guidelines to ensure responsible handling of data in academic and research contexts. Furthermore, the University must ensure that faculty and students use AI tools ethically and in compliance with existing frameworks.

- The use of generative AI must comply with existing Philippine data privacy laws[9] and the data privacy policies of the University.

- The use of generative AI must not infringe on the data privacy rights and the intellectual property rights of other individuals.

- Any data or information given to generative AI services through prompts and other channels is considered "shared" to that service. In doing so, faculty and students must be mindful of the terms and conditions of that service. Most free services include options to turn off the utilization of the user's data for training their models. It is ideal to turn these options off whenever appropriate.

---

[8] DiPietro, M. (2010). 14: theoretical frameworks for academic dishonesty: a comparative review. *To improve the academy*, *28*(1), 250-262.

[9] Republic Act No. 10173, *Data Privacy Act of 2012* (Philippines)

- Personal and highly sensitive information should never be shared with generative AI services.

- Internal University documents, proprietary research, and confidential information should never be shared with generative AI services.

- In research, if any participant or subject data is to be processed by generative AI services, this must be indicated in the informed consent form in the data collection process. Furthermore, such data must always be anonymized, and ethical approval must be secured when necessary.

- Consistent with guidelines published by most academic publishers, AI cannot be treated as an author or creator of any work, as AI does not have accountability.

Appendix A: Guidelines on Defining AI Usage Policies

In choosing the **usage policy level** for each grade component, we make the following recommendations:

Usage Policy	Recommendation
Free to use	This category is recommended for assessments that aim to push the boundaries of knowledge, for assessments that emphasize exploration (e.g., formative assessments), for assessments where generative AI use is a central component, and for situations where the students already have the requisite knowledge and skills to be able to validate and synthesize AI output. It should be noted that this usage policy does not exempt the students from accountability, so the students must still maintain human agency in the final output regardless of how AI was used.
Allowed in certain contexts	This category is recommended for assessments where AI can be used to augment parts of the work but must not undermine core learning outcomes of that assessment. For example, consider a course where generative AI can greatly improve the organization, structure, and clarity of student submissions, but the ideas and content must be developed solely by the students as part of the course's learning objectives. When choosing this category, it is important to articulate clearly which contexts generative AI would be allowed or not. The context can be the purpose of generative AI usage (e.g., you can only use it for writing improvements) and/or the types of tools that can be used (e.g., you can use Grammarly but not GrammarlyGo).
Banned	This category is recommended for foundational courses where students are not yet expected to have the knowledge and skills needed to sufficiently validate and synthesize generative AI output. This category is also recommended for assessments that aim to measure students' acquisition of skills in a controlled environment (e.g., summative assessments).

Here are four points to consider when defining the AI usage policy level:

- **Be open to the potential benefits of generative AI in learning.** While it may seem simpler to just ban generative AI in every course, this will not allow us to prepare our students for a future where AI is strongly integrated into society. Therefore, we should give a thoughtful assessment of generative AI's benefits and risks and consider allowing its responsible use when it is appropriate.

- **Make sure that generative AI use is aligned with learning outcomes.** The use of generative AI should not undermine any skill or knowledge that the students are expected to master by the end of the course. Even in competencies where generative AI tools perform exceptionally well, students will still need to develop sound foundational knowledge and skills to be able to validate and synthesize AI output[10]. Therefore, we must ensure that generative AI is not being used to bypass the attainment of a course's core learning objectives.

- **Consider the common uses of generative AI among students.** There are a variety of ways that students can make use of generative AI. One way to categorize generative AI usage is by its purpose:

Type of Use	Description
Brainstorming	Generative AI is used to generate outlines, ideas, and potential solutions. AI response is not directly used in the submission itself.
Writing improvement	Generative AI is used to improve writing, which may range from simple grammar checking to rewriting entire sections to fit the desired tone.
Summarization	Generative AI is used to summarize existing content for easier understanding, or for picking out the main points.
Presentation Help	Generative AI is used to generate a script for an oral presentation, or presentation slides or outlines.
Problem-Solving	Generative AI is used to solve structured, possibly complex problems such as math problems or programming problems[11].

[10] Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO Publishing.
[11] Modern generative AI tools called reasoning models are now capable of complex problem solving to a certain degree of accuracy (https://platform.openai.com/docs/guides/reasoning)

Note that the list above is non-exhaustive. Furthermore, there may be field-specific use cases of generative AI. For example, in the computer science and information technology field, a common use case of generative AI is to generate computer codes or scripts for a specified purpose. When stating the AI usage policies for a given course, think about the possible ways that students might use generative AI in that course, and consider which types of use are acceptable and which types of use undermine the course's learning objectives.

Another way to categorize generative AI use is by the amount of contribution the AI plays in the final output, as seen in North Carolina Generative AI Implementation Recommendations and Considerations for PK-13 Public Schools[12] and more recently also adopted by the Department of Management Organization in De La Salle University[13]. In this categorization, the different levels indicate the extent of collaboration or power-sharing between students and AI. For instance, level 0 can be considered "banned", while level 4 can be considered "free to use". Levels 1 to 3 fall under "allowed in specific contexts".

No.	Level of AI Use	Full Description
0	No AI use	The assessment is completed entirely without AI assistance. AI must not be used at any point during the assessment. This level ensures that the students rely solely on their own knowledge, understanding and skills.
1	AI-Assisted Idea Generation and Structuring	No AI content is allowed in the final submission. AI can be used in the assessment for brainstorming, creating structures, and generating ideas for improving work.
2	AI-Assisted Editing	No new content can be created using AI. AI can be used to make improvements to the clarity or quality of student crafted work to improve the final output.
3	AI for Specified Task Completion	AI is used to complete certain elements of the task, as specified by the teacher. This level requires critical engagement with AI-generated content and evaluating its output. You are responsible for providing human oversight and evaluation of the AI-generated content.
4	Full AI Use with Human Oversight	You may use AI throughout your assessment to support your own work in any way you deem necessary. AI should be a "co-pilot" to enhance human creativity. You are responsible for providing human oversight and evaluation of the AI-generated content.

Note that when defining the specific contexts where generative AI is allowed, it's not necessary to map an existing categorization such as the ones shown above to a binary "allowed" or "not allowed" tag. Departments can be flexible in declaring the allowable parameters of generative AI use according to the needs of the course. The categorizations above only serve as guides, but departments may ultimately want to create their own categorizations that are more suitable for their specific courses.

- **Be mindful of data privacy.** Ensure that the parameters of generative AI use in any course do not violate data privacy, especially in activities that deal with highly sensitive data and information.

[12] North Carolina Department of Public Instruction. (2023). *NCDPI generative AI implementation recommendations and considerations for PK-13 public schools* [Press release]. Retrieved from https://www.dpi.nc.gov/press-release/ncdpi-generative-ai-implementation-recommendations-and-considerations-pk-13-public-schools/download?attachment

[13] https://docs.google.com/document/d/1JDetgP3TDlQWpmSVmPkZCIDYdsqOng1wGXRWaFbGsJk/edit?tab=t.0

In writing the **Generative AI Use statement**, the following components must be present:

- A clear, concise summary of the usage policies in the course
- The justification or rationale on why such usage policies are implemented, with respect to the learning objectives of the course

Here are some templates that you can use for a course with predominantly "free to use", "allowed in specific contexts", and "banned" usage policies. These templates can be modified in any way to suit the needs of the course. Alternatively, the statement can be crafted entirely from scratch.

Dominant Usage Policy of the Course	**Sample Generative AI Use Statement**
Free to use	This course [describe the nature of the course and / or its core learning objectives (e.g., "focuses on exploration and pushing the boundaries of knowledge", "is a specialized course that assumes the student already has mastery of core skills", "is a course about exploring the strengths and limitations of AI")]. Therefore, students are free to use generative AI for any purpose in their submissions for this course, as long as they exercise transparency and accountability with any AI-assisted output. Exceptions to this are [state the relevant assessments], where the use of generative AI is [banned or allowed in some contexts]. Students are reminded that they are still accountable for any output that they submit, and that inaccuracies, lack of quality, or failure to show ownership over the submitted works will still incur penalties.
Allowed in specific contexts	This course [describe the nature of the course and / or its core learning objectives]. While generative AI can augment [explain parts of the course that can benefit from generative AI], there is also a risk of [explain the threats of AI mis-use towards the learning objectives of the course]. Therefore, in this course, students are allowed to use generative AI only for the purposes of [state contexts in which generative AI is allowed, OR which AI tools are allowed] in their submissions. Students are **not** allowed to use generative AI for [state contexts in which generative AI is not allowed, OR which tools are not allowed]. Exceptions to this are [state the relevant assessments], where the use of generative AI is [free to use or completely banned]. Violations of this usage policy will be treated as academic dishonesty.
Banned	This course [describe the nature of the course and / or its core learning objectives (e.g., "aims to develop foundational skills", "places a high emphasis on human creativity and originality", "focuses on interpersonal skills")]. The use of generative AI can [explain how generative AI can undermine the nature or learning objectives of the course]. Therefore, in this course, students are **not** allowed to use generative AI, either partially or fully, in their submissions. Exceptions to this are [state the relevant assessments], where the use of generative AI is [free to use or allowed in some contexts]. Violations of this usage policy will be treated as academic dishonesty.

For additional discipline-specific inspiration, you may refer to examples from courses in other universities compiled by Dr. Lance Eaton[14]. Please be reminded that the Generative AI Use statement must include a short justification on how the usage policies align with the learning objectives of the course.

[14] https://docs.google.com/spreadsheets/d/1lM6g4yveQMyWeUbEwBM6FZVxEWCLfvWDh1aWUErWWbQ

Appendix B: Guidelines on the Use of AI Detector Tools

AI detector tools are services that try to detect whether a textual output is AI-generated or human-written. Popular examples of such tools include Turnitin, GPT Zero, and Grammarly AI Detector. These tools are helpful in flagging suspicious submissions. However, since AI detector tools are also probabilistic AI models, it is important to understand that such tools are not perfect. They can produce both false positives (human content mistakenly flagged as AI-generated) and false negatives (AI-generated content mistakenly flagged as human). Independent empirical studies on common AI detector tools show varying, sometimes contradictory performance in flagging AI-generated content[15].

Turnitin states on its own website[16]: "While Turnitin has confidence in its model, Turnitin **does not make a determination of misconduct**, rather it provides data for the educators to make an informed decision based on their academic and institutional policies. Hence, we must emphasize that **the percentage on the AI writing indicator should not be used as the sole basis for action or a definitive grading measure by instructors**."

In addition to the false positives and false negatives of AI detector tools, they also present several limitations:

1. **AI detector tools only detect "visible" uses of AI.** AI detector tools are designed to check if the submission itself was generated by AI. However, students may use AI in other ways, such as generating ideas, asking questions, solving problems, or summarizing readings. These "hidden" uses cannot be detected by AI detector tools if the AI responses were not directly used in the submission.
2. **AI detector tools can be bypassed by simple evasion techniques.** Empirical studies have shown that evasion techniques such as paraphrasing and clever prompt engineering can bypass detection[17] [18]. It is very probable that students who intend to cheat are smart enough to apply some of these evasion techniques in preparing their submissions, rendering AI detector tools ineffective.
3. **Over-reliance on AI detector output may encourage undesirable behaviors.** If there is a fixation on passing AI detector models, it may encourage undesirable behaviors among students such as repeatedly editing or paraphrasing a piece of work (AI-generated or otherwise) until it passes an arbitrary threshold of a particular detector tool. Ultimately, it is against the interest of educators to promote this kind of behavior as it does not contribute to critical thinking and learning.
4. **AI detector tools are playing catch-up with developments in generative AI.** AI detector tools can struggle to keep up with the rapid advancements in generative AI models. Current AI detector models work by learning patterns of known generative AI models and checking whether those patterns align with student submissions. With the release of new generative AI models, it is possible that those patterns may change, necessitating AI detector tools to be updated.
5. **AI detector tools are mostly limited to English text.** Other languages, particularly Filipino, as well as other forms of outputs (e.g., images, video), are currently not fully supported by AI detector tools.

Therefore, the following recommendations are given in using AI detector tools:

- Recognize the AI detector tools are not perfect and may produce false positives and false negatives.
- Use AI detector tools to flag suspicious submissions, but human judgment and verification must be done before any penalty is applied.
- Recognize that some uses of generative AI (e.g., those that do not directly appear in the output, or those that used evasion techniques) are not caught by AI detector tools, and thus it is still important to do a critical assessment of students' submission.

---

[15] As summarized in the AI Detector Tools Fact Sheet, released along with the HDA on "Reminder on the Use of AI Detector Tools" (https://drive.google.com/file/d/1RQpksfJkZFpdTXtqAX-LT3X4EZ1nJUGg/view)

[16] Turnitin. (n.d.). *Turnitin's AI writing detection capabilities FAQs*. Retrieved from https://guides.turnitin.com/hc/en-us/articles/28477544839821-Turnitin-s-AI-writing-detection-capabilities-FAQs

[17] Pudasaini, S., Miralles, L., Lillis, D., & Salvador, M. L. (2025, January). Benchmarking AI Text Detection: Assessing Detectors Against New Datasets, Evasion Tactics, and Enhanced LLMs. In *Proceedings of the 1stWorkshop on GenAI Content Detection (GenAIDetect)* (pp. 68-77).

[18] Zhang, Y., Ma, Y., Liu, J., Liu, X., Wang, X., & Lu, W. (2024, April). Detection Vs. Anti-detection: Is Text Generated by AI Detectable?. In *International Conference on Information* (pp. 209-222). Cham: Springer Nature Switzerland.

Appendix C: Guidelines on Preventive Measures Against Academic Dishonesty

As generative AI tools become more complex, it also becomes more and more difficult to distinguish between human and AI-generated content. Reactive measures, i.e., checking for evidence of cheating after submission requires additional effort on the part of educators due to the absence of reliable tools that can perfectly distinguish human from AI content. Therefore, now more than ever, educators are highly encouraged to combine reactive measures with preventive ones to help maintain a culture of academic integrity. This set of guidelines, compiled from various sources, offer some strategies that educators can use to discourage generative AI cheating.

- **Clearly communicate parameters of generative AI use to students.** This is often overlooked by educators who assume that parameters of use are always obvious to students. However, by clearly articulating the policies for generative AI use, students are more likely to adhere to them[19]. Discussion of the policies should also be accompanied by an explanation of why such policies are being implemented in relation to the learning objectives of the course.

- **Evaluate not just the output, but also the process.** Students can be required to submit artifacts of their ideation, thinking, and design processes. Also, consider emphasizing the process more in the evaluation rubrics of assessments, rather than just focusing on grading the final output[20].

- **Redesign traditional assessments.** Generative AI tools perform better in certain types of assessments than others. Educators must now reflect on how traditional assessments can be redesigned to ones that are more AI-resistant. According to the UNESCO Guidance for Generative AI in Education and Research, we must contend with the fact that generative AI can already do a lot of tasks that we traditionally assess well, and we should therefore rethink how learning is to be assessed or validated[21]. Here are some strategies that can be used:

    - **Consider using authentic assessments.** Authentic assessment are assessments wherein students must apply their skills and knowledge in real-world situations[22]. Due to the complexity and unpredictability of real-world situations, reliance on generative AI will be much less effective in these types of assessments.

    - **Focus on higher-order skills.** Generative AI can perform lower-order skills exceptionally well. However, while generative AI can also perform higher-order skills such as application, synthesis, evaluation, their responses will usually be more lacking in depth. Thus, designing assessments that focus on higher-order skills can help make them more AI-resistant[23].

    - **Incorporate localized or personalized contexts.** Assessments that require students to connect knowledge and skills to local or personalized situations make it harder for generative AI to give responses with depth[24].

- **Consider using face-to-face modality.** For assessments that measure lower-order skills, and for assessments that cannot be redesigned and where generative AI cheating cannot also be easily distinguished, consider using face-to-face modality, as it is the only true, guaranteed way to ensure full human contribution.

---

[19] Northern Michigan University Center for Teaching and Learning. (n.d.). *Academic dishonesty using generative AI*. https://nmu.edu/ctl/academic-dishonesty-using-generative-ai
[20] https://www.timeshighereducation.com/campus/designing-assessments-generative-ai-mind
[21] Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO Publishing.
[22] https://citl.indiana.edu/teaching-resources/assessing-student-learning/authentic-assessment/index.html
[23] Lye, C., & Lim, L. (2024). Generative Artificial Intelligence in Tertiary Education: Assessment Redesign Principles and Considerations. *Education Sciences*. https://doi.org/10.3390/educsci14060569.
[24] Dwyer-Kuntz, Tricia & Kuntz, S. Guardians of Integrity: Navigating Cheating in the Age of AI. *Shaping Future Classrooms*, 21.