

Online Workshop

Bharati College | R&D Cell × CDC (Samarth Bharat) | Under IQAC

India's AI in Research Ecosystem Workshop(Basic)

AI Research Tools: Ethics Guidelines: Laws

Relevance for Academics, Researchers & PhD Scholars

Principal: Dr Prof.Saloni Gupta

Training Co-Ordinator: Dr Anubha Bhargava

AI & Data Analytics Leader | Researcher | Educator | 20+ Years Experience

Session 1 of 5

Module: Introduction AI lanscape

Every Monday | 11:00 AM – 12:00 Noon | Online (Google Meet) | Starts 6th July 2026



ETHICAL USE OF AI

IN RESEARCH & TEACHING ENGAGEMENT

Empowering Educators with AI for Research, Teaching and Academic Excellence

Organized by

Research and Development Cell

in Collaboration with CDC (Samarth Bharat)

Under the aegis of IQAC

ONLINE MODE

Online
(Google Meet)

STARTS

17 6th July
2026
(Monday)

DURATION

17 Every
Monday for 5
Weeks

TIME

11:00 AM
– 12:00 Noon

RESOURCE PERSON

Mr. Anurag Mittal

- **AI & Data Analytics Leader / Researcher / Educator**
- **20+ Years of Experience in Research, Teaching, Analytics and AI-driven Solutions.**

Register here:

Payment details are mentioned in the form



Last date to register:
30th June, 2026.

PROGRAMME HIGHLIGHTS

- Literature Review & Academic Writing
 - Plagiarism Detection & Academic Integrity
- Research Data Collection & Analysis
 - AI-enabled Teaching Material Development
- Multimedia Content Creation
- Session Recordings will be Provided
- Notes, Practice Resources & Useful Links will be Shared
- Certificate of Participation
- Google Meet Link will be Shared with Registered Participants

FOR ANY QUERIES

Contact: 93195 73696

The Researcher's Blueprint: The 16-Step Master Checklist

1. AI is Assistant,
Not Researcher.

2. Understand Ethical
AI Meaning.

3. Know Safe Uses
(Summaries).

4. Avoid Misuse
(Fake citations).

5. Never List AI
as Author.

6. Verify Every AI
Output.

7. Language Support,
Not Original Thinking.

8. Cite Original
Sources.

9. Disclose Use
Transparently.

10. Protect Data
Privacy.

11. Avoid Plagiarism
& Patchwriting.

12. Use Detectors
Carefully.

13. Maintain AI-Use
Record.

14. Follow Journal
Policies.

15. Use Ethical
Workflow.

16. Apply via Real
Examples.

**AI strengthens research only when the researcher remains honest,
responsible, transparent, and accountable.**

AI strengthens research only when the researcher remains honest, responsible, transparent, and accountable.

THE 7 AI SUTRAS

(MeitY Governance Guidelines 2025) — the seven principles behind every rule



1

Safety & Reliability

Ensuring AI-assisted research outputs are high-quality, reproducible, and free from dangerous hallucinations.



2

Fairness & Non-Discrimination

Mitigating algorithmic bias in student evaluations, research scoring, and grant allocations.



3

Transparency & Explainability

Openly disclosing the use of AI tools in methodology sections and identifying the logic behind AI-driven decisions.



4

Privacy & Data Protection

Safeguarding personal information of students and research participants within AI training datasets.



5

Accountability & Oversight

Maintaining “human-in-the-loop” control; researchers bear non-delegable responsibility for all AI-generated findings.



6

Ethics & Human Values

Aligning AI use with academic integrity and the moral rights of original human authors.



7

Innovation & Inclusivity

Leveraging AI to create adaptive learning pathways and support minority language contexts through tools like Bhashini.

A model AI-use statement you can adapt

Adjust to your institution's exact wording (Module 6), but this structure covers what UCT, Wits, UP, UKZN, SU and the ICMJE all ask for.

Sample declaration

"In preparing this thesis I used generative AI tools as follows. I used ChatGPT (GPT-4o, OpenAI) to improve the grammar and clarity of text I had written myself in Chapters 2 and 5, and to explain statistical concepts, which I then verified against [named sources]. I used [reference/discovery tool] to identify candidate literature, all of which I located and read in the original before citing. No AI tool was used to generate data, results, or interpretations, and no identifiable participant data was uploaded to any AI tool. I have verified all AI-assisted content and I take full responsibility for the entire work, which represents my own analysis and writing. No AI system is credited as an author."

AI use declaration

Table 3: AI use declaration

AI system used	What was it used for?	Where in the work was it used/what was it used for?
To what extent did you use AI and why do you consider the work as your own?		

AI Tool and Version	Task	Researcher	Researcher Contribution	Evaluation Notes	Dates, Prompts, and Output
<i>Example: Asta (no version listed)</i>	<i>Finding papers for lit review</i>	<i>Laila</i>	<i>Vetted results by reading abstracts at the publishers' websites; used Google Scholar to find additional papers.</i>	<i>Asta identified several very helpful papers that I didn't find in Google Scholar, but some it marked as "perfectly relevant" were not.</i>	<i><link></i>
<i>Example: Gemini (version 2.5)</i>	<i>Summarizing responded to open-ended questions</i>	<i>Armando</i>	<i>Read raw responses first, noted themes, then read the AI summary, checking for additional themes.</i>	<i>Summaries were not helpful; they were too general and vague.</i>	<i><link></i>

The AI-use log — fill it as you go

A running log turns disclosure from a stressful reconstruction at submission into a two-minute copy-paste. It is also your evidence trail from Module 7.

DATE	TOOL & VERSION	WHERE (CHAPTER/TASK)	WHAT YOU ASKED	HOW YOU VERIFIED & OWNED IT
12 Mar	ChatGPT (GPT-4o)	Ch2 · language	Improve clarity of my paragraph on sampling	Compared to original; kept my meaning; rewrote two sentences
18 Mar	Elicit	Ch2 · discovery	Find studies on CDF governance in SADC	Located and read 6 originals; discarded 2 AI summaries as inaccurate
02 Apr	ChatGPT ADA	Ch4 · code help	Explain error in my R regression script	Re-ran corrected code myself in RStudio; results reproduced

The main detectors, factually

1 Turnitin

2 GPTZero

3 Copyleaks

4 Originality.ai

5 DrillBit

Turnitin AI writing detection

Integrated into the similarity report many universities already use. Turnitin publicly targets a "less than 1% false positive rate" and stresses the tool "does not make a determination of misconduct" but only "provides data for educators". Yet its own Chief Product Officer acknowledged a ~4% false-positive rate at the sentence level and higher false positives on scores below 20% AI, prompting a qualifying asterisk on low scores (Turnitin, 2023; Inside Higher Ed, 2023).

The main detectors, factually

1 Turnitin

2 GPTZero

3 Copyleaks

4 Originality.ai

5 DrillBit

GPTZero

Launched January 2023 by Princeton undergraduate Edward Tian. Uses perplexity and burstiness; widely used in education. Independent reporting has documented meaningful false-positive rates that can wrongly accuse innocent students.

The main detectors, factually

1 Turnitin

2 GPTZero

3 Copyleaks

4 Originality.ai

5 DrillBit

Copyleaks

A plagiarism and AI-content platform founded in 2015, combining database comparison with a model that assesses semantic style. Like all detectors, its AI judgments are probabilistic, not proof.

The main detectors, factually

1 Turnitin

2 GPTZero

3 Copyleaks

4 Originality.ai

5 DrillBit

Originality.ai

A commercial detector aimed at web publishers and marketers, founded 2022. In one study it detected GPT-4 text at 91.3% mean accuracy — which collapsed to 27.8% after the text was passed through adversarial "humanizer" software, showing how easily detectors are defeated (and therefore how weak they are as evidence).

The main detectors, factually

1 Turnitin

2 GPTZero

3 Copyleaks

4 Originality.ai

5 DrillBit

DrillBit

A plagiarism and AI-detection platform widely used in India, where it powers the national ShodhShuddhi initiative. Multilingual support and institutional scale, but subject to the same fundamental accuracy limits as its peers.

Detectors are biased against non-native English writers

- Across seven GPT detectors, a set of 91 TOEFL essays by non-native English speakers was misclassified as AI at high rates — more than half were flagged, and one detector flagged around 98%
- The same detectors correctly classified over 90% of essays by native US eighth-graders as human
- The cause is structural: additional-language writing tends to have lower perplexity, which detectors read as "AI-like"
- For South African postgraduates — most of whom write in English as an additional language — this is not a footnote; it is a direct threat of false accusation

The Big Picture (The Headline)

AI detectors falsely accuse non-native English speakers of cheating far more often than native speakers.

The software meant to catch AI-written text has a heavy built-in bias.

Breaking Down the Key Findings

1. High Rate of False Accusations

- **What the slide tested:** Researchers ran 91 essays written by non-native English speakers (taking the TOEFL exam) through seven different popular AI detectors.
- **The result:** Over **50% of these human-written essays** were wrongly flagged as AI. One detector even falsely accused **98%** of them.

2. Native Speakers Don't Have This Problem

- When the same detectors were given essays written by native US 8th graders, they correctly recognized that a human wrote them over **90% of the time**.
- **The contrast:** The tools work well for native speakers, but fail miserably for non-native speakers.

Why Is This Happening? (The "Perplexity" Issue)

The slide mentions a technical term called "**perplexity**":

- **What is perplexity?** Think of perplexity as how unpredictable, creative, or complex a person's word choice is.
- **Why AI detectors get confused:** AI tools look for simple, predictable, formulaic sentence structures. Non-native writers naturally tend to use simpler, more straightforward vocabulary and repetitive grammar rules because they are writing in a secondary language.
- **The flaw:** Because non-native writing is more predictable (**lower perplexity**), AI detectors mistake simple English for AI-generated text.

The Real-World Consequence

The final point highlights that this isn't just a fun math stat—it causes real harm:

- For students or researchers using English as an additional language (like South African postgraduates mentioned on the slide), this flaw poses a constant, unfair risk of being **wrongfully accused of academic dishonesty or cheating**.

Native speakers were rarely flagged because their writing style features **high perplexity**—a technical metric AI detectors rely on.

1. High "Perplexity" (Unpredictable Word Choices)

- **How native speakers write:** Native writers naturally mix uncommon idioms, varied sentence lengths, regional slang, complex vocabulary, and unexpected stylistic choices.
- **How detectors see it:** This unpredictability triggers high "perplexity" scores. To an AI detector algorithm, **unpredictable text = human text**.

2. Low "Burstiness" in AI vs. Native Writing

- **"Burstiness"** measures sentence variety. Native human writers tend to alternate between very short, punchy sentences and long, complex ones.
- AI detectors associate this natural "bursting" rhythm with human authorship, which native speakers naturally display far more than non-native speakers or standard AI models.

3. How the Researchers Proved This Bias

To prove that sentence structure—and not actual AI usage—was the cause of the false flags, the Stanford researchers did a reverse experiment:

- **Simplifying Native Writing:** They took the native 8th-grade essays and simplified the vocabulary to match non-native writing patterns. **Result:** The detectors immediately started flagging the native essays as AI-generated.
- **Sophisticating Non-Native Writing:** They took the non-native TOEFL essays and asked ChatGPT to rewrite them using advanced, complex native vocabulary. **Result:** The detectors suddenly marked the AI-edited essays as 100% human!

The Core Flaw

AI models (like ChatGPT) are trained to pick the **most probable next word**. Non-native speakers naturally rely on standard, high-probability textbook words. Because both write using simple, expected patterns, detectors lump non-native humans and AI into the exact same category.

Why serious universities are switching detectors off

Internationally

Vanderbilt University disabled Turnitin's AI detector in August 2023, citing opacity, false-positive risk, bias against non-native speakers and privacy, concluding: "We do not believe that AI detection software is an effective tool that should be used" (*Vanderbilt University, 2023*). Cambridge and UT Austin are among others reported to have opted out.

ZA In South Africa

UCT is discontinuing AI detectors from October 2025, with a director stating "AI detectors are simply not reliable". Stellenbosch calls detection software unreliable and warns against a "culture of suspicion". Rhodes flags detectors as biased against non-native English writers and says they must not be the sole basis for discipline. UP and NWU, by contrast, retain detection as one signal among many.

How to protect and evidence your own authorship

If your work is genuinely yours, the strongest protection is a visible trail of how you produced it. This is legitimate, encouraged, and far more robust than any attempt to game a score.

Build an authorship trail as you write

These habits both improve your scholarship and defend it if a detector misfires.



Version history. Draft in tools that keep revision history (e.g. Google Docs, tracked changes); your evolving drafts are evidence of authorship.



Notes and outlines. Keep your reading notes, mind-maps and outlines — the scaffolding only a real author accumulates.



An AI-use log. Record where and how you used AI (Module 8), so you can account for it precisely rather than vaguely.



Be able to explain it. If you can discuss your argument and defend your choices, you demonstrate the understanding no model can fake.



Verified references. A bibliography of sources you have actually read is itself strong counter-evidence to a "machine-written" claim.



Know the due process. A detector score is an *indicator*, not proof; you are entitled to fair process, to see the basis of any allegation, and to respond.

✓ KNOWLEDGE CHECK

A student who wrote her chapter herself is flagged at "62% AI" by a detector. What is the best, most ethical response?

- A** Frantically rewrite the chapter until the detector score drops.
- B** Present her drafts and notes, explain her work, and cite the documented unreliability and bias of detectors through due process.
- C** Run it through an AI "humanizer" so it stops being flagged.

Correct. Correct. Contemporaneous logging makes your disclosure accurate (you won't forget or misstate uses), effortless at submission, and doubles as an authorship evidence trail if you are ever wrongly flagged.

CARRY THIS FORWARD

Make transparency a workflow, not an afterthought

Set up your AI-use log today, adapt the declaration to your institution, and run the 30-second pre-flight before each AI action. Do this and disclosure stops being a worry — it becomes proof of how carefully you worked.

The five commitments

If you remember nothing else, remember these. Every module feeds one of them.

- 1 I verify.** Every AI-provided fact, number, quotation and reference is checked against the original before it enters my work. *(Modules 1-3)*
- 2 I own the thinking.** The argument, analysis and interpretation are mine; AI never originates data or findings. *(Modules 4-5)*
- 3 I disclose.** Meaningful AI use is declared, honestly and specifically, per my institution's rules. *(Modules 5-6, 8)*
- 4 I protect.** Participant data stays lawful, minimised and off unvetted tools; my drafts and log evidence my authorship. *(Modules 4, 7)*
- 5 I stay accountable.** I can explain and defend everything in my thesis without hiding a thing. *(All modules)*

Build your personal AI-use policy

Turn the course into a one-page standard you actually follow. Complete each line for your own project.

My PhD AI-use policy (complete these)

Write your answers, keep it with your thesis file, and revisit it each term.



My institution's rule is... (link the exact document and its disclosure requirement)



I will use AI for... (your green and conditional list)



I will never use AI for... (your red list)



My data rule is... (what may/may not be uploaded, and to which tools)



I log AI use in... (where your log lives)



My disclosure statement reads... (your adapted declaration)

Would you know the line? Decide, then reveal

Read each scenario, decide green / amber / red and what you would do, then reveal the model answer.

1 A student asks ChatGPT to write her entire discussion chapter, then edits a few words and submits it.

HIDE

Red. The discussion is where original interpretation lives; generating it wholesale and lightly editing is passing AI text off as her authorship — a misrepresentation regardless of disclosure, and treated by several SA institutions as contract cheating. The fix: she writes the interpretation herself and may use AI only to polish her own sentences, with disclosure.

2

A researcher uploads anonymised, aggregate survey counts to an AI tool to help draft a plain-language summary.

[HIDE](#)

Amber — likely acceptable. The data are anonymised and aggregate, so POPIA risk is low, and drafting a summary is legitimate. Conditions: verify the summary does not overstate the findings, rewrite it in her own voice, and disclose the assistance. She should still confirm no re-identification is possible from the aggregates.

4 A student uses Elicit to find 20 papers and cites all 20 from the AI summaries without opening them.

HIDE

Red. Citing sources from summaries you have not read risks hallucinated or misrepresented references and misstates your scholarship. Discovery tools are green for *finding* literature; the moment you cite, you must have located and read the original.

3 An additional-language writer is flagged at "40% AI" though he wrote the chapter himself.

HIDE

A false positive to defend, not evade. Detectors are documented to over-flag non-native English writers (Liang et al., 2023). He presents his drafts, notes and version history, explains his argument, and invokes fair process and the evidence of detector unreliability. He does *not* run it through a "humanizer" — that would misrepresent honest work and is itself a breach.

Which single question best captures the ethical test taught throughout this course?

A "Will an AI detector be able to catch this?"

B "Can I stand behind this at my viva, explain how I produced it, and disclose it without embarrassment?"

C "Did this save me the most time?"

Correct. Correct. If you can stand behind it at your viva, explain how you produced it, and disclose it without embarrassment, you are almost certainly acting ethically. Concealment is the warning sign.

Your university has no standalone AI policy yet. What binds you?

A Existing integrity rules, the shared disclose–verify–account expectations, and any guidance in guides or frameworks.

B Nothing — you can do anything until a formal policy exists.

C Only international journal rules, which don't apply to a thesis.

Correct. Correct. Existing academic-integrity, plagiarism and research-misconduct rules apply, along with the shared expectations to disclose, verify and stay accountable, and any guidance in library guides or frameworks.

What is the strongest protection against a false AI-detection accusation?

- A Never using any AI tool at all.
- B Running your honest writing through a humanizer pre-emptively.
- C A visible authorship trail — drafts, notes, verified references, a log — and knowing your right to due process.

Correct. Correct. A visible authorship trail — version history, notes, verified references, an AI-use log, and the ability to explain your work — plus fair process, is the strongest and most ethical defence.

What each tool is for, and its main academic risk

CATEGORY	REPRESENTATIVE TOOLS	WHAT THEY DO	MAIN ACADEMIC RISK
General LLM assistants	ChatGPT, Claude, Gemini, Copilot	Draft, explain, summarise, reformulate text	Hallucinated facts, quotations and references; variable data handling
Literature discovery	Elicit, Consensus, Scite, Research Rabbit, Semantic Scholar, SciSpace	Semantic search, citation mapping, smart citations	Summaries misread findings; you must still read the originals
Writing & editing	Grammarly, Writefull, Paperpal, Trinko, QuillBot	Grammar, clarity, academic-style feedback	Edits language not substance; paraphrase/humanize features can disguise text
Reference managers	Zotero, Mendeley, EndNote	Collect, organise, cite and format sources	AI summaries misrepresent papers; metadata needs manual checking
Transcription	Otter.ai, Whisper (local), Word Transcribe	Convert interview/focus-group audio to text	Accuracy drops on accents/jargon; cloud upload risks participant confidentiality
Qualitative analysis	ATLAS.ti, NVivo, Taguette (manual)	Organise and code qualitative data; AI code suggestions	Inconsistent, decontextualised codes; transcript confidentiality
Quantitative & code	ChatGPT ADA, GitHub Copilot, R/Python assistants, ChatPDF	Run analysis, write code, chat with documents	Wrong methods stated confidently; must be reproduced and verified

Where nine institutions stand (July 2026)

INSTITUTION	INSTRUMENT (YEAR)	AI AS AUTHOR	DISCLOSURE	AI DETECTORS
UCT	EiRC research guideline (2023); AI-in-education framework (2025)	Prohibited	Required (substantial use)	Discontinuing (Oct 2025)
Wits	Academic Integrity-Based AI Framework v2.2 (2025)	"Cannot be an author"	Required (template)	Not addressed
Stellenbosch	GenAI TLA resources; interim guidelines (2025)	Not explicit	Required (postgrad form)	Sceptical
Pretoria	Library AI guidelines (2026); Academic Integrity Policy	Rejected	Required (tool + prompt)	Uses Turnitin AI indicator
UKZN	GenAI Guidelines SE/01/0312/24 (eff. 2025)	Not co-authors	Required (declaration)	Not addressed
UJ	Student GenAI Quick Guide (2023)	Implied prohibited	Required	Not mentioned
Rhodes	Guidelines for Assessment in the Time of AI (c. 2023–24)	Implied prohibited	Required (tools + prompts)	Cautious; not sole basis
NWU	Policy on AI 5P_5.10 (2025)	Attribution required	Required	Recognised as legitimate
UNISA	Draft AI Policy + Academic Integrity Policy	Not explicit (draft)	Required (cite & disclose)	Not addressed

Before you upload participant data

Four checks

- **Lawful basis (s.11):** usually informed consent
- **Minimise:** only what is "adequate, relevant and not excessive"; de-identify
- **Cross-border (s.72):** foreign-hosted tools = restricted transfer
- **Secure & account:** your responsibility; fines up to R10m

RULE OF THUMB

When in doubt, keep it offline

If data can identify a participant and you have not confirmed the tool's data terms and your ethics approval, do not upload it. Use local tools and speak to your ethics committee first.

A student who wrote her chapter herself is flagged at "62% AI" by a detector. What is the best, most ethical response?

- A** Frantically rewrite the chapter until the detector score drops.
- B** Present her drafts and notes, explain her work, and cite the documented unreliability and bias of detectors through due process.
- C** Run it through an AI "humanizer" so it stops being flagged.

Correct. Correct. She calmly presents her version history, notes and drafts, explains her argument, and points to the well-documented unreliability and bias of detectors — invoking fair process. Genuine authorship, evidenced, is the strongest defence.

Prompt & verification hygiene

- Never paste confidential or participant data into a public tool
- Ask the model to show reasoning and sources, then check them
- Assume every reference is fake until you find the original
- Cross-check facts against at least one authoritative source
- Keep your own voice — rewrite, don't paste

A 30-second pre-flight

- Is this task green, amber or red? (Module 2)
- If amber: have I verified and logged it?
- Am I about to upload data I should not? (Module 4)
- Could I explain this at my viva without embarrassment?
- Does my institution require me to disclose it? (Module 6)

Why is keeping an AI-use log as you work better than writing your disclosure at the end?

- A Because a longer disclosure looks more thorough to examiners.
- B** It keeps the disclosure accurate, makes submission effortless, and doubles as an authorship evidence trail.
- C Because logging replaces the need to verify AI outputs.

Correct. Correct. Contemporaneous logging makes your disclosure accurate (you won't forget or misstate uses), effortless at submission, and doubles as an authorship evidence trail if you are ever wrongly flagged.

Type of AI Use	Example	Permission Level	Additional Caveats
Ideation	"Give me five different titles for a research project about . . ."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Ideation	"I want to write a proposal for a grant that has these specifications. . . What would be a strong proposal?"	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Productivity	Automatic meeting transcripts or summaries for online meeting platforms.	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	

Type of AI Use	Example	Permission Level	Additional Caveats
Grant Writing	"Write a ten-page grant application based on these bullet points."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: ☐ Required ☐ Not required	
Non-academic Writing	"Write three paragraphs for a website summarizing the research study in the uploaded file. The target audience is parents of high school students."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: ☐ Required ☐ Not required	
Agentic AI for Project Management	"Send emails to all participants in phase one (as indicated in the spreadsheet) asking them to complete the survey. Send follow up reminders at 5 and 10 days to those who don't complete the survey (as indicated in the survey system account)."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: ☐ Required ☐ Not required	

Type of AI Use	Example	Permission Level	Additional Caveats
Generation of Ancillary Materials	"Create four images. Each one should show a diverse group of college students conducting an experiment in a chemistry lab."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Generation of Research Materials	"Create ten questions designed to assess the sense of belonging of high school students in STEM courses."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Idea Refinement	"Refine this research question: ..."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	

Type of AI Use	Example	Permission Level	Additional Caveats
Feedback	"List ways that this research question might be improved:"	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Feedback	"Write a peer review of this paper, identifying at least five areas for improvement."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Generation of Synthetic Data	"Provide 20 responses to this interview question as if you were an engineering major."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	

Type of AI Use	Example	Permission Level	Additional Caveats
Data Analysis	“In this set of interviews, identify the five main themes related to influences that encourage students to study computer science.”	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Editing	Use of tools that check grammar and spelling	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Academic Writing	“Write an abstract of 220-250 words based on this research study.”	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	

Type of AI Use	Example	Permission Level	Additional Caveats
Content for Alternative Dissemination	"Create a 20-minute podcast for high school physics teachers based on the research study described in the attached file."	☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Other:		☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	
Other:		☐ Not permitted ☐ Permitted with closed AI system only ☐ Permitted with open AI system, but no model training ☐ Permitted with open AI system and model training Disclosure: __ Required __ Not required	

4.2 AI use checklist

- 4.2.1 I am convinced and can support my claim that my assessment product is an indication of my own learning, knowledge, skills, and understanding.
- 4.2.2 Where I have used AI tools for enhancing my own creation of ideas and words, I acknowledge that I have to declare it.
- 4.2.3 Where I have used AI tools for generating new ideas, words, code, image-prompts for other AI Image-generating tools, or structure and even presentations (or other AI tools that can be used as assistants to the knowledge building and representing process), I have declared and documented the use of such tools and I am prepared to talk about the process I used and what it contributed to my learning and insights.
- 4.2.4 I am aware that the lecturer can ask me to demonstrate my learning, for example through explaining the choices I made in terms of approach, content used, literature selected, conclusions drawn, etc. through an additional assessment like an oral (for example).
- 4.2.5 Where the use of AI tools was explicitly not allowed, I can say with integrity and honesty that I did not use any such tools.

- 4.2.6 I understand that if I am not able to agree to the above points, there is a chance that my academic behaviour will be deemed unethical and might lead to a disciplinary case being brought against me on the grounds of cheating or plagiarism and that the standard procedures for such behaviour will be followed.
- 4.2.7 As per the Disciplinary Code of SU (par. 10.2.1 and 10.2.2) I understand that I take responsibility for the integrity of my work, which includes the obligation to ask for clarification from an academic member of staff if I am unsure of anything, and that I strictly adhered to all instructions received in the course of the academic assessment by relevant and authorised staff (whether the instruction is in oral or written format).
- 4.2.8 I understand that when AI-tools were not allowed in the assessment, or where I am not able to document and declare my use of such tools, this behaviour will be deemed as cheating in examinations and assessments (Disciplinary Code 1.1 c.) as I referred to "unauthorized notes, books, electronic devices or other reference material".

5.1.4 Where used in assessment of student work, the following conditions hold:

- 5.1.4.I SU allows the use of AI tools such as ChatGPT and Quillbot to assist the lecturer with reviewing the assessment, e.g., finding themes or weaknesses in student work, but on condition that the assessment is not fully outsourced to machines. The tools can assist, to minimize potential bias (e.g. related to language skills), but their output should be verified to ensure that you agree with it. No confidential (e.g. personal reflections) or personal student information should be shared with these kinds of AI systems. Final accountability, as with writing, lies with the lecturer (AI systems are not fool proof and cannot be held accountable).
- 5.1.4.II The lecturer, just as with students using these tools, should check the results for accuracy.
- 5.1.4.III Declare which tools will be used as well as how they will be used and how student queries and AI output will be verified and/or moderated, both for AI-generated text detection, and for assessing student work, in the module framework (open to internal and external moderation).
- 5.1.4.IV No personal student information or sensitive information, i.e., reflections sharing personal experiences should be uploaded to platforms outside the SU digital firewall.
- 5.1.4.V Be aware that there are no detection systems that are 100% reliable, they can be easily circumvented, and some systems can create different levels of false-positives (saying text is AI-generated or plagiarized when it isn't).

5.2 AI Lecturer AI use Checklist

- 5.2.1 I am familiar with the implications that the existence of generative AI technologies can have on how my planned assessment will be done.
- 5.2.2 I have given thought to how to redesign my assessment accordingly.
- 5.2.3 I have decided to either allow the responsible use of AI tools, but with the clear understanding that students should declare their use or prohibit the use of AI tools for a given assessment altogether.
- 5.2.4 I have clearly communicated my decision in terms of SU's overarching approach and guidelines to my students.
- 5.2.5 Where I plan to use AI tools to assist me in assessing students' work, I have also communicated that to my students.
- 5.2.6 When students submit long-form English language text through Turnitin, I understand that the AI-generated percentage score is not immediate proof of plagiarism (or "aigiarism"), but an important a caution as to possible use of generative AI tools to create text that is not the student's own words, or where the student did not acknowledge, declare, cite or explain their process in using the AI-generated outputs. I understand that, as indicated in point 5.3.4 below, students currently cannot see the % of AI-generated text and that penalising them for it, can be deemed unfair.ⁱ

5.3 Turnitin guidelines¹⁹

There are several concerns with the use of AI text detectors, including the functionality in Turnitin, which needs to be considered.

- 5.3.1 Although Turnitin states that it identifies long-form English language AI-generated text with 98% certainty it is not determinative²⁰. Whilst it can be used as a form of evidence, it cannot be used exclusively and determinatively to prove academic misconduct.
- 5.3.2 According to Turnitin's website, they state that "[they] will likely miss up to 15% of text written by AI". That is fair scope for wrongdoing. This means in a text suggested to be 20% written by AI, the real AI contribution could be 35% and higher, which is substantial, given that only longform text is considered.
- 5.3.3 Turnitin states that they currently have a 1% false positive rate (i.e., incorrectly identifying fully human-written text as AI-generated). The challenge with this is that there is no way of verifying the result, unlike with the Turnitin similarity report.
- 5.3.4 The AI text % is not currently shared with students, so there is no formative value in it. This also means that we are now using the tool differently than we have for the past years, and if we choose to do this, it should be communicated with our students clearly.
- 5.3.5 If this approach is chosen, how will an acceptable AI text % be determined?

5.3.6 What will be indicated as AI-generated text: can it indicate where AI has been used to improve or edit text as opposed to AI-generated content. What about tools such as Grammarly and Quillbot?

The SU procedure for the investigation and management of allegations of plagiarism or of cheating according to the Disciplinary code is followed. As with plagiarism allegations, it is suggested that:

6.1 The student be contacted via email and asked to a discussion of the concern

6.2 The student should also be supplied with

6.2.1 The allegation/s

6.2.2 Details of the work that will be discussed

6.2.3 A mark-up of the work

6.3 If this is the first time that the student has been suspected of this kind of irregularity, the discussion should be seen as mostly formative. The student can be advised to bring

6.3.1 Their declaration of AI use

6.3.2 Evidence such as relevant AI chat history

6.4 During the discussion, the student will be given the opportunity to

6.4.1 Explain how they used AI, why they used it as well as what they used

6.4.2 Answer questions related to the content of the submitted work

CARRY THIS FORWARD

Detectors are weak evidence; your trail is strong evidence

AI detectors are unreliable and biased, which is exactly why the ethical strategy is not to game them but to render them irrelevant: write honestly, keep your drafts and notes, log your AI use, and be ready to explain your work. That protects you far better than any score-lowering trick — and it keeps your integrity intact.

Permitted, conditional, prohibited

Permitted

Understanding concepts · outlining your own ideas · proofreading your own prose · brainstorming · learning a tool or method · finding candidate literature to read yourself.

Conditional (verify + disclose)

Drafting sections you rewrite · AI-assisted coding you test · summarising sources you read · paraphrasing your own text · qualitative code suggestions you audit.

Prohibited

Fabricating data or results · submitting unverified AI references · passing AI text off as your authorship · uploading protected personal data · evading detection to misrepresent authorship.

PRACTICAL COMPLIANCE CHECKLIST FOR ACADEMICS

Research & Papers

- Disclose use of AI tools in methodology / acknowledgements
- Do not submit AI-only content as your original work
- Verify copyright status of training data used
- Ensure research participant data is DPDP-compliant

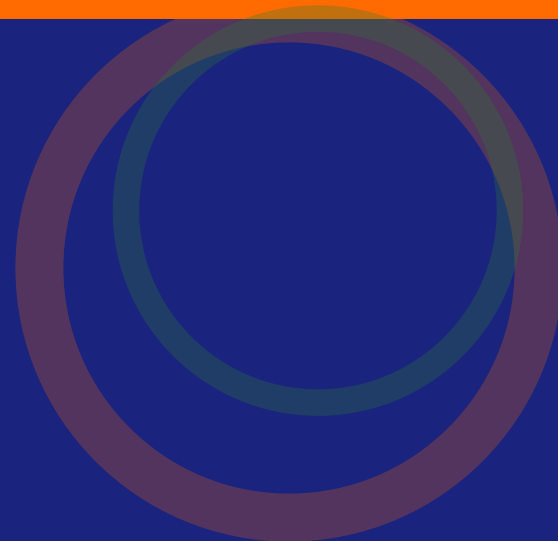
Teaching & Content

- Label AI-generated teaching content clearly
- Obtain student consent before using AI with their data
- Preserve human oversight in AI-based grading
- Follow your institution's GenAI use policy

Data & Privacy

- Do not feed personal student data into public AI tools
- Ensure AI vendors are DPDP-compliant data fiduciaries
- Anonymise data before AI-assisted research analysis
- Keep records of consent for AI-processed personal data

Thank You



Mr. Anurag Mittal

AI & Data Analytics Leader | Researcher | Educator |

Email id: anurag7@illinois.edu

Online Workshop — Ethical Use of AI in Research & Teaching

Bharati College | R&D Cell × CDC (Samarth Bharat) | Under IQAC

#IndiaAI

#DPDP2023

#ResponsibleAI

#AILaw

#AcademicIntegrity

Questions & Discussion

Session 4 of 5 | 11:00 AM – 12:00 Noon | Online (Google Meet)

WHAT IS PROMPT ENGINEERING?

Writing instructions that get you an answer worth using

Prompt engineering is the skill of giving an AI tool clear context, a precise task, and the format you want — so it produces a focused, usable answer instead of a generic one.

WHY IT MATTERS FOR YOUR RESEARCH



Literature Reviews

Get a focused synthesis of themes, gaps and methods — not a generic summary.



Academic Writing

Draft abstracts, emails and proposals that already match your tone and audience.



Data & Coding

Get more accurate analysis scripts and outputs in fewer back-and-forth attempts.

Figure 3 The Way Teachers Seek Generative AI Support



Seeking Input:

Through reflection, the teacher identifies a potential that there may be a better way---that they do not know about---to achieve the target learning goals.

Teachers prompt generative AI to iterate on and explore new ideas about how to teach and get students to understand target learning goals.

Teachers either:

1. Feel more confident that what they are doing is an effective way forward.
2. Have ideas for a new workflow plan that they can be more confident about, which they then create.

Seeking Outputs:

The teacher identifies materials that they need but do not have even a draft of, and so they would like help creating it.

Teachers prompt generative AI to make the materials they need.

Teachers either:

1. Have a draft of the needed materials, which they then modify and improve upon.
2. Have a useable, final version of the materials they need.