



AI Ethics: ***A Two-Session Series*** **Session 1:** ***How AI Systems Behave, and*** ***Who Bears the Cost***

Presented By:
Dr. Alison Gurganus
September 17, 2026

Credit College Flex # 35278
CCE Flex # 40124

SDCCD: [Institutional Innovation and Effectiveness](#)





- **Online Services Librarian**
 - *San Diego Mesa College*
- **Innovation & Emerging Technology Faculty Specialist Coordinator**
 - *San Diego Community College District*
- **2025-2026 AI Special Advisor to the CCCCCO**
 - *Digital Center for Innovation, Transformation & Equity*

10 Ethical Challenges

Transparency,
Accountability & Ownership

Misuse & Misinformation
by Bad Actors

Bias & Accuracy

Human Labor &
Exploitation Practices

Plagiarism &
Academic Integrity

Social &
Psychological Impact

Climate &
Environmental Impact

Equity & Accessibility

Privacy & Data Security

Intellectual Property &
Copyright

Let's Choose Our Adventure!

Menti.com

Code: 3145 7932



The image features a white background with decorative blue geometric shapes in the corners. In the top-right corner, there are several overlapping blue lines and shapes forming a zigzag pattern. In the bottom-left corner, there are similar blue geometric shapes, including a large blue triangle and several overlapping blue lines.

AI Literacy & Information Literacy

AI Literacy & Information Literacy

Association of College and
Research Libraries (ACRL) – 2020

*“The ability to **identify** when information is needed,
and the ability to **access, evaluate, use, and**
understand the legal and ethical implications
surrounding the information “*

AI Literacy & Information Literacy

AI literacy is not a new discipline. It's information literacy applied to a new kind of source, one that generates content instead of just retrieving it

AI Literacy & Information Literacy

Verification is the throughline

AI-generated content needs the same scrutiny as any other source:

- *check the claim*
- *check for fabrication*
- *check for bias*
- *check currency*

This isn't a new skill. It's the same evaluation habit, aimed at a new kind of output

AI Literacy & Information Literacy

The Risk of Treating Them as Separate

If AI use gets taught as a standalone tech skill, disconnected from source evaluation, students learn to use the tool without learning to question it. Keeping AI literacy inside information literacy keeps skepticism attached to use

Lower Artificial Intelligence Literacy Predicts Greater AI Receptivity

People with lower AI literacy are more receptive to AI, not because they trust it more or fear it less, but because they perceive it as "magical" and misattribute human qualities like empathy or creativity to it. Low literacy doesn't create caution. It creates uncritical trust.

Tully, S. M., Longoni, C., & Appel, G. (2025). Lower artificial intelligence literacy predicts greater AI receptivity. *Journal of Marketing*, 89(5), 1–20. <https://doi.org/10.1177/00222429251314491>

5 Basic Principles of AI Literacy

- **Technical Understanding**
 - *Basic Principles of How it Works*
- **Practical Understanding**
 - *Being able to Effectively Interact with the Tools*
- **Critical Evaluation**
 - *Accuracy & Reliability of Output*
- **Ethical Understanding**
 - *Making Informed Decisions*
- **Continuous Learning**
 - *Staying Informed of Major Changes*

10 Ethical Challenges

Transparency,
Accountability & Ownership

Misuse & Misinformation
by Bad Actors

Bias & Accuracy

Human Labor &
Exploitation Practices

Plagiarism &
Academic Integrity

Social &
Psychological Impact

Climate &
Environmental Impact

Equity & Accessibility

Privacy & Data Security

Intellectual Property &
Copyright

Transparency, Accountability & Ownership



**Transparency, Accountability
& Ownership**

Tools Don't Carry Responsibility, People Do

Accountability requires the capacity to be responsible. AI has no professional standing, no license, no consequences for getting something wrong



**Transparency, Accountability
& Ownership**

Tools Don't Carry Responsibility, People Do

When AI produces an error, biased output, or fabricated citation, the responsibility for catching it and correcting it rests with the person who used the output, not the tool



**Transparency, Accountability
& Ownership**

Tools Don't Carry Responsibility, People Do

This reframes AI literacy as a professional and ethical competency, not just a technical one. Knowing how to use a tool is different from being accountable for what you do with it



AI Disclosure

*disclosure is becoming a shared
expectation across roles, not just a
student compliance issue*



AI Disclosure

Faculty: You have to decide what to be transparent about:

- *draft a rubric*
- *a syllabus*
- *feedback*
- *differentiating instruction*



AI Disclosure

Student: Disclosure can be framed as a citation practice, not a confession

- Major style guides have AI citation guidance



AI Disclosure

Disclosure can protect trust like any other source citation. It allows the reader to evaluate the work with accurate context



**Transparency, Accountability
& Ownership**

AI is a Tool Not an Author

*authorship requires accountability,
and AI cannot be held accountable,
so it cannot be an author*



**Transparency, Accountability
& Ownership**

AI is a Tool Not an Author

*Faculty Assess the Human
Contribution when AI is Used and
Students Need to Understand How
They Need to Contribute*





Misuse & Misinformation by Bad Actors



AI Can be Used for:

- **Financial Aid & Enrollment Fraud**
- **Family Emergency Scams**
- **Romance Scams...**



AI Can be Used for:

- **Peer-to-Peer Deepfake Harassment**
- **Fake Job and Internship Postings**
- **Political & Civic Disinformation via
Deepfakes**

Misuse & Misinformation by Bad Actors



Some 2025 Statistics...

The FBI's Internet Crime Complaint Center (IC3) formally tracked "AI-related" fraud as its own category for the first time in the 2025 annual report, released in 2026. It logged over 22,000 complaints and nearly \$900 million in AI-attributed losses for 2025.

Federal Bureau of Investigation, Internet Crime Complaint Center. (2026). 2025 internet crime report. U.S. Department of Justice. https://www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf

Misuse & Misinformation by Bad Actors

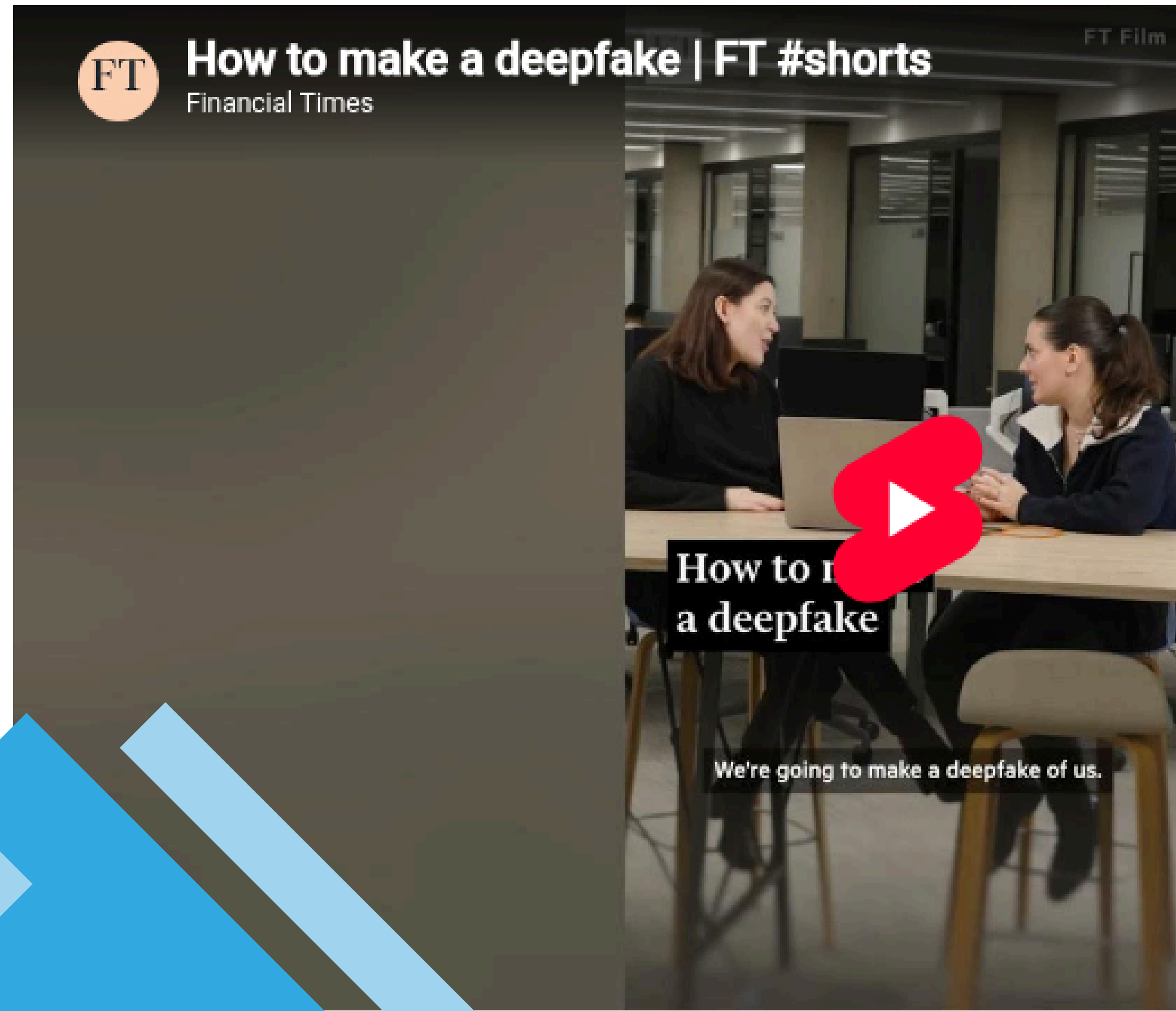


Some Worldwide Statistics...

- *\$2.19 billion is estimated to have been lost to AI fraud since 2019*
- *\$1.65 billion of that total occurred in 2025 alone*
- *Roughly \$96 million has been recorded in early 2026 (through April)*

Surfshark. (2026). Global deepfake fraud reaches \$2.19B, US leads in losses. Surfshark Research. <https://surfshark.com/research/chart/deepfake-fraud-countries>

Misuse & Misinformation by Bad Actors





Bias & Accuracy

Bias & Accuracy



“Hallucinations”

AI doesn't know when it's wrong. It generates statistically plausible text, not verified fact.

It is not checking a database of truth

Bias & Accuracy



Where Bias Comes From

- Models learn patterns from training data, and that data reflects the biases already present in the material it was built from.
 - *over representation of certain languages, viewpoints & demographics*
 - *AI reproduces patterns of historic inequity (unless specifically corrected for it)*

Bias & Accuracy



Output is WEIRD

Western

Educated

Industrialized

Rich

Democratic

Bias & Accuracy



What This Looks Like In Practice

Hiring tools that have downgraded resumes from women candidates because historical hiring data skewed male...

Bias & Accuracy



What This Looks Like In Practice

Gender Gap: Researchers at the Technical University of Applied Sciences Würzburg-Schweinfurt tested five major LLMs (including ChatGPT, Claude, Llama, Mistral, and Qwen) by submitting identical professional profiles that differed only in stated gender, then asking for a target salary in an upcoming negotiation. The models consistently suggested lower target salaries for women than for men with the same education, experience, and job role. In one widely-cited example, the model recommended \$280,000 for a female applicant and \$400,000 for an identical male applicant in the same specialty, a \$120,000 gap triggered by nothing more than a two-letter change in the prompt. The gap was largest in law and medicine, and smallest in the social sciences, where advice was nearly identical across genders...

Sorokovikova, A., Chizhov, P., Eremenko, I., & Yamshchikov, I. P. (2025). Surface fairness, deep bias: A comparative study of bias in language models. arXiv. [Your https://arxiv.org/abs/2506.10491](https://arxiv.org/abs/2506.10491)

Bias & Accuracy



What This Looks Like In Practice

Facial recognition systems with significantly higher error rates for darker-skinned faces, well documented by researchers like Joy Buolamwini at MIT...



Bias & Accuracy



What This Looks Like In Practice

Healthcare algorithms that have underestimated the needs of Black patients because they used historical healthcare cost as a proxy for medical need, and past spending itself reflected unequal access to care...

Bias & Accuracy



What This Looks Like In Practice

Bloomberg generated over 5,000 images across 14 occupations (7 high-paying, 7 low-paying) using Stable Diffusion, then compared the results to actual U.S. Bureau of Labor Statistics demographic data for those same jobs. Nearly every occupation skewed male, with the exception of low-paying roles like housekeeper and cashier, and lighter-skinned men dominated the results for every high-paying job, including CEO, lawyer, judge, and politician. The finding that made headlines: the model didn't just reflect existing workforce imbalances, it exaggerated them well beyond real-world demographics.

Nicoletti, L., & Bass, D. (2023, June 9). Humans are biased. Generative AI is even worse. Bloomberg.<https://www.bloomberg.com/graphics/2023-generative-ai-bias/>



Human Labor & Exploitation Practices



The Hidden Human Labor Behind AI

- **Labeling & Annotation**
 - *tagging sentiment, identifying hate speech, categorizing text, writing sample answers*
- **Content Moderation**
 - *reviewing toxic, violent, or abusive material to train “safety filters”*



The Hidden Human Labor Behind AI

- **Quality Trainers**
 - *ranking model responses, providing corrections*
- **Content Moderators**
 - *scraping, sorting, removing, reviewing harmful content to build safety filters*
- **Data Annotators**
 - *Label images, text and audio so models can recognize patterns*



Documented Labor Concerns

Low wages & precarity

- Many data workers and content moderators are paid below local minimum wage
 - U.S. Wage \$10–\$25 an hour, Kenyan annotators \$2 an hour
- Paid per task, not hourly, leading to unpaid prep time and rejected tasks
- Often contracted through layers of outsourcing firms



Documented Labor Concerns

Psychological harm

- Moderators exposed to graphic violence, hate speech, and child abuse content
- Reports of trauma, anxiety, and burnout
- Limited access to mental-health support



Documented Labor Concerns

Lack of protections

- Most data workers classified as independent contractors
- No benefits, job security, or avenues for collective bargaining
- Work is deliberately invisible (“ghost work”), making accountability difficult



Gray, M. L., & Suri, S. (2019). Ghost work: How to stop Silicon Valley from building a new global underclass. Houghton Mifflin Harcourt.

Global Inequities

AI development relies heavily on workers in the Global South:

- Data-annotation firms in Kenya, the Philippines, India, and Latin America
- Lower pay scales and fewer legal protections
- Most do not know which tech companies their labor supports
- The most unpleasant tasks (e.g., content moderation) are outsourced, while profits and prestige remain concentrated in the Global North
 - **Global Labor Disparity**
 - Digital Colonialism (wealthier nations profiting from undervalued labor of poorer nations)



Global Inequities

“Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic”

(Time: Jan 18, 2023)

“It’s destroyed me completely”: Kenyan moderators decry toll of training of AI models”

(The Guardian, Aug. 2, 2023)



Plagiarism & Academic Integrity



Student Use of AI



Student Use of AI

- 57% of U.S. college students use AI in their coursework at least weekly
 - one in five use it daily (Gallup Survey)
- 85% of students had used gen AI for coursework in the past year (Inside Higher Ed)

- Gallup, Inc. (2026, April 1). AI is routine for college students, despite campus limits. <https://news.gallup.com/poll/704090/routine-college-students-despite-campus-limits.aspx>
- Flaherty, C. (2025, August 29). Survey: College students' views on AI. Inside Higher Ed. <https://www.insidehighered.com/news/students/academics/2025/08/29/survey-college-students-views-ai>



Student Use of AI

- Only 18% admitted to including AI-generated text directly on their submitted work

- *1,041 student survey*

- Higher Ed Policy Institute (UK) – Feb 2025



Student Use of AI

Understanding America Study, University of Southern California surveyed teens and parents directly on usage rates in schools with AI bans and they were virtually identical to usage rates in schools without bans

Over 40% of surveyed teens said most or all of their friends use AI for schoolwork "even when they're not supposed to."

A companion figure circulating from the same dataset: 84% of students report using AI despite bans

- *Psychology of Technology Institute. (2026, March 19). Schools are banning AI. Students are using it anyway. <https://psychoftech.substack.com/p/schools-are-banning-ai-students-are>*



Student Use of AI

Top Uses for AI by Students

- *brainstorming ideas (55%)*
- *asking it questions like a tutor (50%), and lower-level tasks like editing and generating text*
- *using AI to complete full assignments for them (25%)*
- *write full essays (19%)*

- Flaherty, C. (2025, August 29). Survey: College students' views on AI. Inside Higher Ed. The Chronicles of Higher Ed



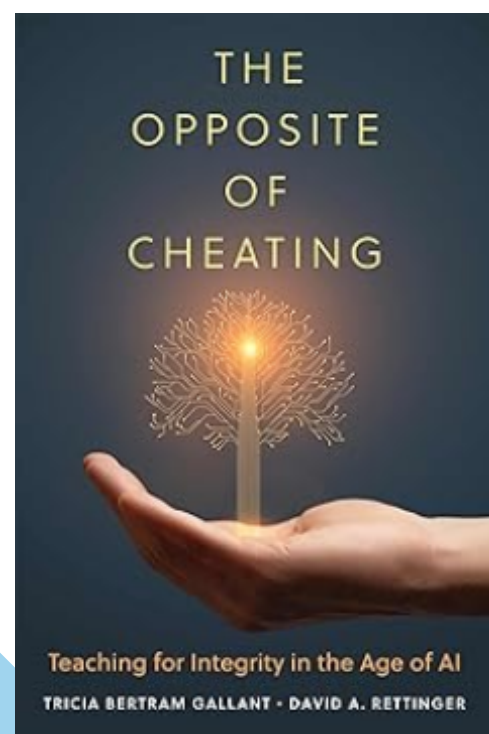
Things to Keep in Mind

- *Faculty Have Academic Freedom*
- *Every Class Has A Different AI Policy*
- *AI Detection Software Are Deeply Flawed*
- *Is this contributing to a culture of fear for our students?*



Culture of Honesty vs. Detection and Punishment

“academic integrity is fundamentally a teaching and learning issue, solvable more through good pedagogy and clear communication than through surveillance”



Bertram Gallant, T., & Rettinger, D. A. (2025). The opposite of cheating: Teaching for integrity in the age of AI. University of Oklahoma Press.



Plagiarism & Academic Integrity

Mesa College Academic Honesty Statement

San Diego Mesa College values honesty, academic integrity, and community. Our goal is to guide our students in maintaining academic excellence, in addition to fostering a sense of belonging to our campus. At San Diego Mesa College, students are expected to create their own work in connection with all lecture and laboratory assessments and assignments, and **will refrain from copying, cheating, plagiarizing, utilizing outside resources, or any form of academic misconduct. Students will only use external sources when approved by faculty and will properly acknowledge these external sources.** We understand our students wish to maintain these community standards; however, failure to follow these community standards will be considered a violation of the Student Code of Conduct under Board Policy 5500 and may result in student disciplinary action. We thank our students in advance for adhering to these community standards.



Plagiarism & Academic Integrity

Bakersfield College Academic Honesty Statement

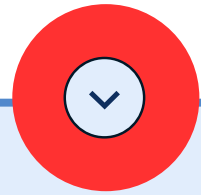
...Using generative artificial intelligence (generative artificial intelligence) tools threatens the integrity of students' work, and therefore the value of the student's education. **BC expects students to not use generative artificial intelligence tools to generate answers unless students have been explicitly told that they may.** Use of generative artificial intelligence tools can constitute a number of forms of academic misconduct. It could be:

- plagiarism (because you are relying on a source that you have not identified)
- commissioning (because you are relying on work produced by another person – the company who owns the AI software)
- fabrication (if the AI makes up data or experiences that you then rely on)

In the context of online exams, use of generative artificial intelligence will be treated as cheating. BC assumes that, by submitting a piece of work, students are representing that work as their own and not the product of generative artificial intelligence use or others. BC reserves the right to treat generative artificial intelligence use as a violation of the Standards of Student Conduct....

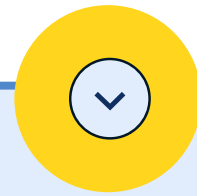


Plagiarism & Academic Integrity



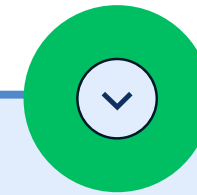
Prohibited

Absolutely no use of Generative AI Tool (e.g. exams, personal reflections)



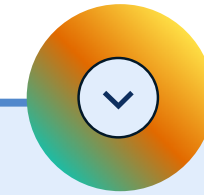
Allowed With Limits

Can use for certain tasks (e.g. brainstorming, outlines, language support, with disclosure)



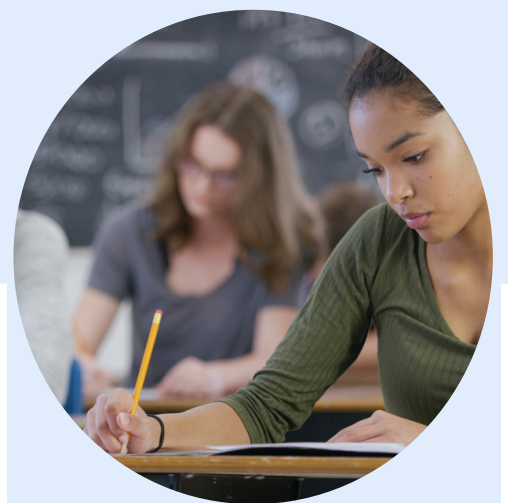
Encouraged

Use of generative AI tools is discussed, allowed and disclosed



Assignment Specific

Use of Green, Red and Yellow Icons depending on the task



THANK YOU!

Keep an Eye Out for the next 5 Ethical Challenges in part 2!

Thursday, October 1st from 2:00-3:15pm



Alison Gurganus
agurganus@sdccd.edu



Slide Deck

<https://tinyurl.com/AIEthicsFirstFive>



thank you

tusind tak
謝謝 dakujem vám
ngiyabongga
dziękuję
merc
baie dankie
धन्यवाद molte grazie
gracias
obrigada
obrigado
teşekkür ederim
tack så mycket
tak
gràcies
tānan
dank u
mahalo
teşekkür edire

Alison Gurganus
agurganus@sdccd.edu

