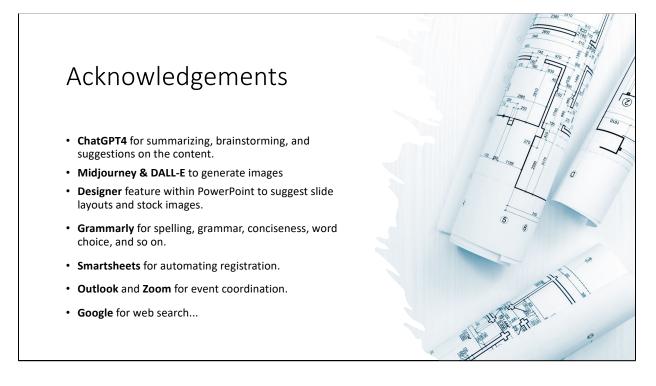


Today's primary learning objectives

- Describe and discuss risks, policy guidance, and best practices related to AI and:
 - Academic integrity
 - Privacy and security
 - Citation and attribution
 - Copyright and intellectual property
 - AI hallucinations and their causes
 - Bias
 - Inequity
 - Environmental impact
 - Misinformation
- Self-evaluate how compatible with AI your existing course is.

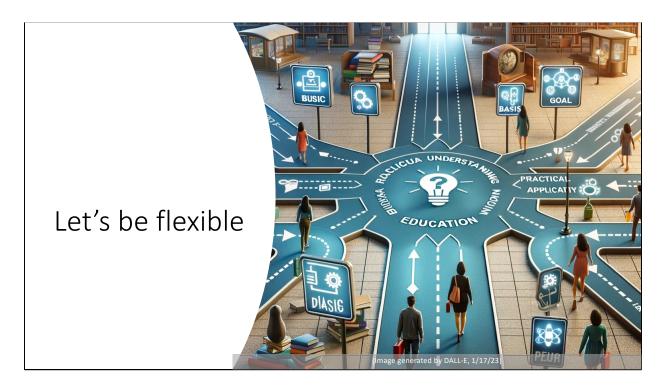




For the sake of transparency and to model best practices, here are some AI tools that I used when developing this workshop.

I used ChatGPT, and Midjourney and DALL-E, Designer in PPT. Oh yeah, I also have Grammarly installed. And if we mean algorithms and machine learning, not strictly LLM. Then I technically should say I used Google a lot too. I also used Outlook for event planning, it has AI features too. And Smartsheet for the registration forms. Zoom! Zoom has AI features too.

Yes, I am being melodramatic! But my intent is to suggest, hopefully in a humorous way, that there are fuzzy and ambiguous lines of distinction here. And warm us up to this topic. This is an ambiguous, and often contradictory topic. We are entering a twilight zone.



This is a big topic so we can't cover it all. Also, because it is so emergent, nobody has all the answers. Those answers are emerging in places like this. So, thank you for joining the discourse!

Everyone is at a different starting point. Some may be learning the very basics, some may be already thinking about bigger issues, or have begun adapting your course. Some of this you may already have heard, other parts might be new to you. Wherever you are, let's be mindful that different people are at different points and be flexible about where we might end up today.

The benefit of being together here is that we can share our thoughts with each other. This workshop will generate more questions than answers. It is a starting point for you to engage further with each other, CTL, and your own learning.

My expectation is the we will improvise a bit and hop around different topics that interest you. So let me know as we go, what you are thinking, if we want to skip ahead or circle back. Maybe even things that aren't on here yet!

Why should you care about AI?

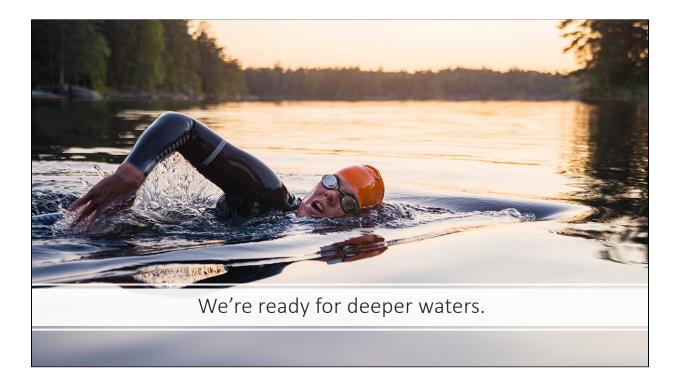
- Our perspective and expertise is important to the discourse
- The landscape is evolving quickly
- Our students and instructors want guidance
- Using AI could improve how we do our work



- Other groups on campus are, as expected bringing their perspective to it, whether it's computer science, business, medicine, and so on. Also, in the bigger discourse, Higher Education is being scrutinized. Think Claudine Gay or Liz Magill or even Michael Tessier-lavigne. People are asking, are we preparing students for success? Is Stanford leading the way responsibly? Is a college education valuable? What role does education play in our society?
- I'm of an age where WWW and web search came to prominence when I was in high school and college. I'm a 90's kid. Some of us might remember dozens of different search engines, Lycos? Alta Vista, Ask Jeeves? Dogpile? It seemed to be evolving so rapidly. I remember teachers saying "Don't cite anything from the web! It's not a valid source" Then the next semester saying, "You can cite web sources from journals or news articles, but Wikipedia is not a valid source!", then later "Wikipedia is okay, but be sure to check the references!" It was constantly evolving.
- Students want guidance. For many of us, this is not our first rodeo. But for some of our students, this might be. And for some of us, this might feel like a first rodeo. If we abdicate then for-profit will fill that void, students will go and do it anyway. We

owe it to our students to give them guidance. .

• Lastly, AI can help us improve how we do our work. Web search, googling it, is now an indispensable part of how we do our work. We couldn't do our work without web search. People smarter than me are predicting that AI ubiquity, AI everywhere is coming soon. And it's going to become a routine part of how we do work like web search has become. And it will have a big impact on how we work, so it makes sense for us to get ahead of it.





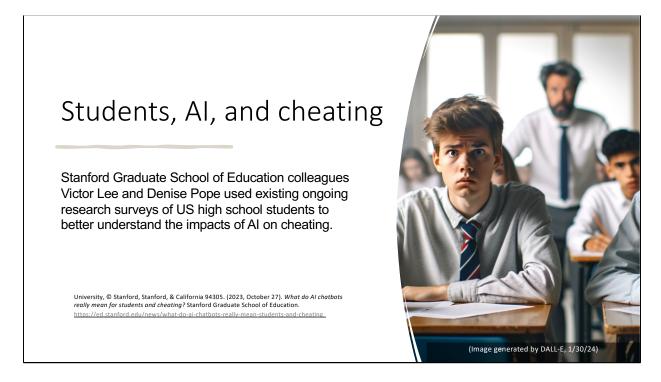
Scenario

Your current course includes several written essay assignments. Currently your syllabus does not include any policy about students using generative AI tools. In a recent meeting, your TA has told you that several students have asked them whether they can use AI tools in the course or not. They suspect that some students are already using AI tools for class assignments.

How might you approach this conversations with your TA?

What factors might you consider when thinking through this issue?





GSE colleagues Victor Lee and Denise Pope used existing ongoing research surveys of US high school students to better understanding the impacts of AI on cheating.

Gen-Al writing detection tools

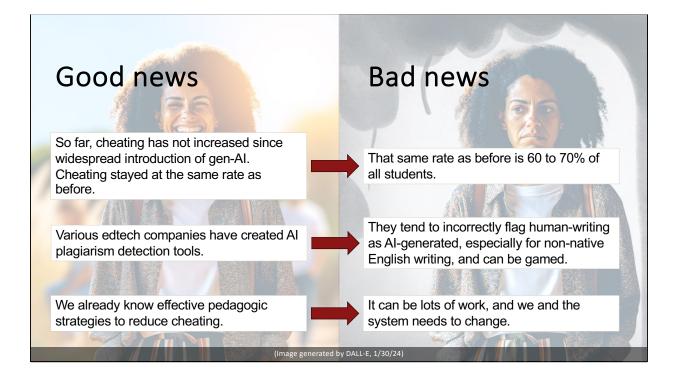
Education researchers evaluated AI content detection tools.

Stanford HAI colleagues Weixin Liang, et. al tested AI plagiarism detectors with non-native English writing.

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of Al content detection tools in differentiating between human and Al-generated text. International Journal for Educational Integrity, 19(1), Article 1. https://doi.org/10.1007/s40979-023-00140-5

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers (arXiv:2304.02819). arXiv. <u>http://arxiv.org/abs/2304.02819</u>





Regarding AI detection tools, keep in mind that these tools don't work the same way as plagiarism detectors. They don't compare human-written works to other human-written works. They work like LLMs themselves, to statistically analyze the probability that text was generated by an LLM. They are inherently guessing based on probability, so too are subject to the biases of LLMs themselves.

- In a recent study, incorrectly flags ~20% of human-writing as AI-generated
- Impacts non-native English writers more
- Less reliable at detecting more sophisticated LLMs
- Savvy users can deceive the detectors

Strategies to reduce cheating

- Build trust and community
- Meaningful experiences
- More in-class and/or low-stakes assessments
- Support robust citation practices
- Focus on the process
- Leverage formative feedback
- Emphasize metacognitive reflection

- Prioritize higher-order thinking
- Leverage multiple modalities
- Personalize assessments
- Use authentic tasks
- Clarify and adapt your assessment criteria

Systemic changes

- Purpose of grading and evaluation
- Structure of curriculum degree programs, prereqs, and so on
- Standardized curriculum and evaluation for every student
- Elitism and exclusionary nature of college entrance
- Age-based cohorts and waterfall paradigm of the school system
- Financial incentive of tuition to pass classes, graduate more students, more quickly.

Stanford's academic integrity policy guidance



OCS—AI guidance

Instructors can decide their own AI course policy and should communicate it in their syllabi.

Absent any course policy, AI use is considered same as assistance from another person.

Students, when in doubt ask the instructor and disclose use.



OCS—Academic integrity

Promote honorable behavior and support learning. Let OCS handle enforcement.

Consult with OCS for any suspected non-compliance.

In consultation, you decide how to proceed.

Generative Al Policy Guidance | Office of Community Standards. (n.d.). Retrieved February 12, 2024, from https://communitystandards.stanford.edu/generative-ai-policy-guidance

Stanford's academic integrity policy guidance



OCS—AI guidance

Instructors can decide their own AI course policy and should communicate it in their syllabi.

Absent any course policy, <mark>Al use is considered</mark> same as assistance from another person.

Students, when in doubt ask the instructor and disclose use.

- What specific behaviors do we mean by "assistance"?
- When do you already allow or prohibit assistance from another person?
- What kind of assistance from another person is beneficial? What kind is detrimental?
- How might that inform how we think about AI?

Generative AI Policy Guidance | Office of Community Standards. (n.d.). Retrieved February 12, 2024, from https://communitystandards.stanford.edu/generative-ai-policy-guidance

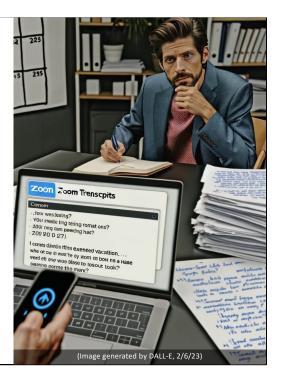
What might you tell a colleague who is concerned about AI and academic integrity are the most important factors to consider?

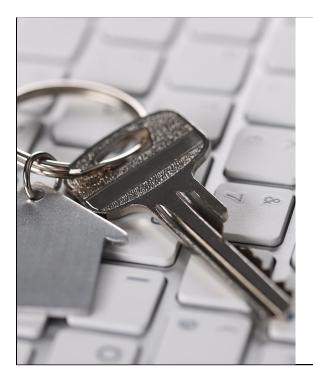


Scenario

You have recently returned from an extended vacation (it was amazing). While you were out, there were many important meetings about an unfortunate incident involving staff and students. You need to catch up on everything that was discussed in the meetings and determine any action items for yourself. You have access to the Zoom transcripts and typed meeting notes. A colleague offers to use their ChatGPT+ account to summarize the transcripts and identify any action items.

What might you consider when using AI for this task? How might you respond to that colleague?

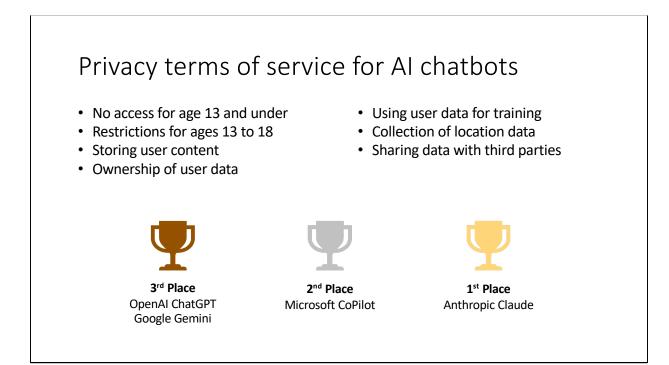




Privacy and security guidance

Responsible AI at Stanford webpage from UIT

- Inform yourself and students about AI
- Don't put sensitive or private info into chatbots
- Use private mode when appropriate
- Get informed consent
- Use third-party tools with care
- Be transparent about use
- Promote dialogue and discourse

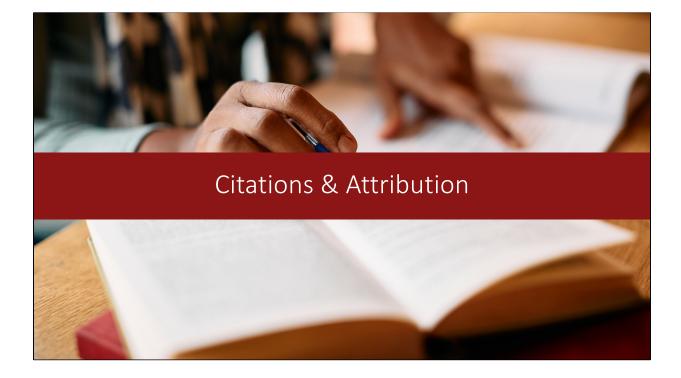


In general:

- No access for users under age 13
 - None allow under 13
- Restrictions for users age 13 to 18,
 - ChatGPt requires parental permission
 - Google doesn't allow Google Education accounts
 - CoPilot defaults to SafeSearch Strict mode
 - Claude is 18 and over only.
- Stores user content
 - They all do.
 - Gemini and CoPilot have some options to delete.
 - Only Claude follows European data protection laws and automatically deletes data after 90 days.
- Owns your data for business purposes.
 - They all do, but Claude is more limited, allows you to opt out.
- Using data for training
 - ChatGPT lets you opt out if you have a paid account
 - Gemini has no opt out
 - CoPilot varies on which product, Office365 accounts have more control

- Claude by default is opted-out, you must opt-in
- Collect location data
 - All track IP address location info
 - Only Gemini also track geolocation
- May share data with third parties
 - They all do when required by law
 - CoPilot and Claude don't "sell" your data.

What are other situations where you might be extra mindful of how you use AI tools?



Scenario

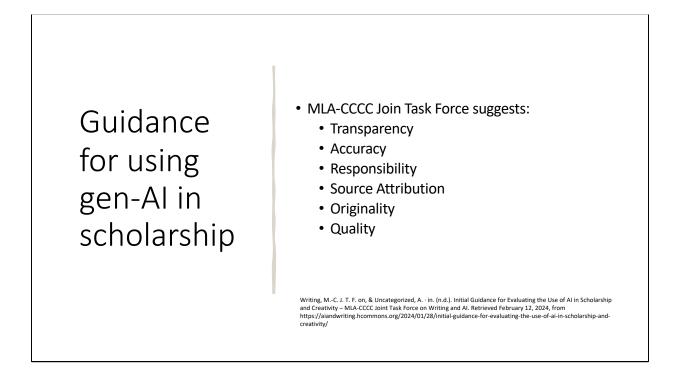
A graduate student that you mentor has been working on a research article for submission to a journal for publication. They've been experimenting with chatbots recently and intend to use one to help them write the article.

They have some concerns about using AI for this. They asked you for advice before getting started and asked if you would give them feedback before they submit the article.

What advice would you give them?

What might you look for when giving them feedback?





Modern Language Association and Conference on College Composition and Communication

Transparency – meaning you disclose how you used AI

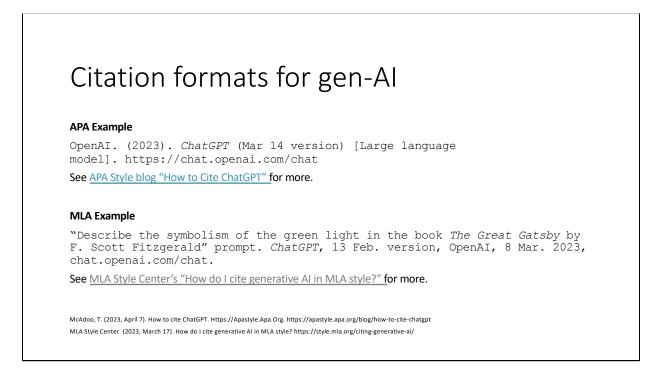
Accuracy – fact check and assure that what the AI said is accurate

Responsibility – you are responsible for vetting and checking what the AI generates.

Source attribution – you verify the sources the AI generates are true. Go to the real source

Originality – Assure that the writing reflect your ideas, with AI assistance or not, the writer must use discernment to determine what is an original contribution to the field

Quality – Writing holds to the quality standards of the discipline and institution



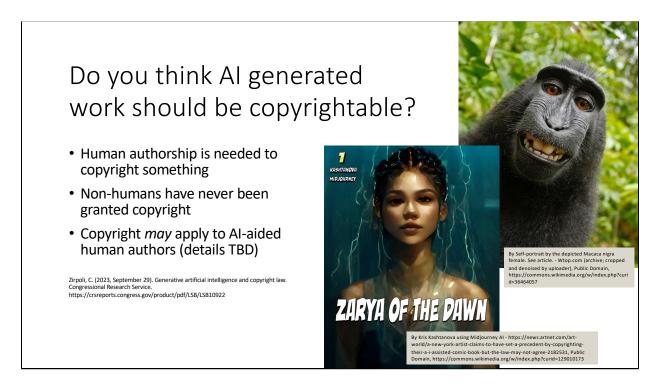
APA treats it as "personal communication". Like it were a conversation with another person.

The author is the entity that wrote the AI's alogrithm or trained it. Hence Open AI as the author here.

MLA, however treats it differently. They don't see it as having an author. It is more like an unauthored work. Perhaps they see it more as the outcome of a random roll of the dice.

Therefore the Title is the prompt, and the container is the AI, the developer is like the publisher.





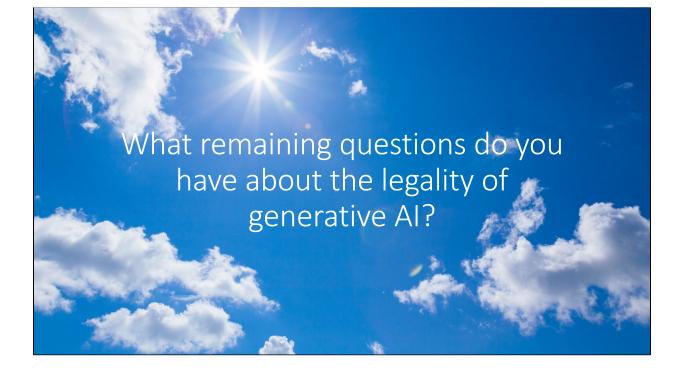
This monkey took a selfie after a photographer left his camera unattended. The photographer argued he had copyrights. Activists argued that the monkey had copyright. Wikimedia argued it was public domain. Judge ruled it was public domain.

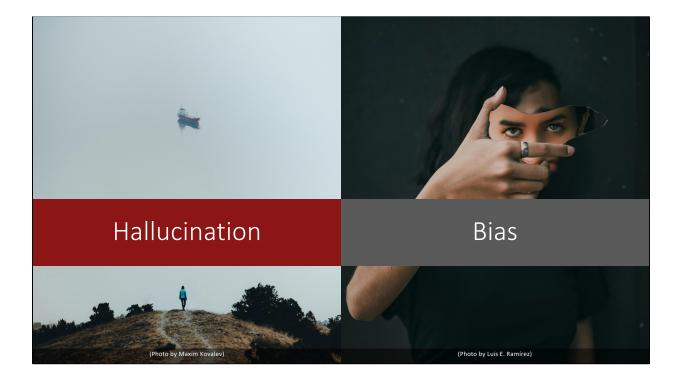
Next the graphic novel writer who used Midjourney to create the images. The judge concluded that it was public domain, but left the door open that copyright could apply in some cases where the human authors can prove that they authored the work not AI. The onus fell to the author to prove that they created the work. This is a slight inversion of previous precedent where typically the plaintiff must prove that the author copied the work.

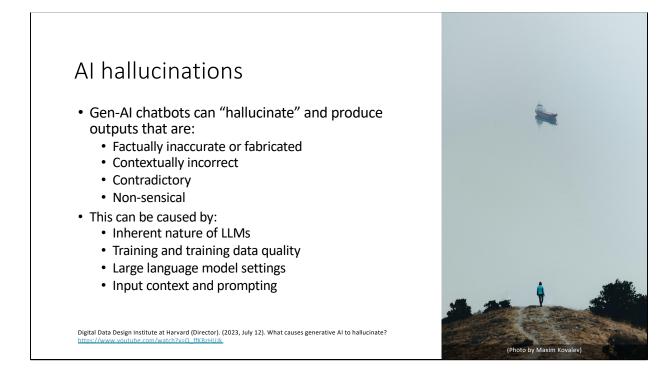
Future court cases will determine how this goes.



Al developers argue that using copyrighted material for training AI is Fair Use. The purpose is for training not replication. The amount is huge, they concede. The nature of machine learning is that AI training data is only being analyzed for the statistical relationships between the words or shapes, colors, etc. So the nature of training would be allowable under Fair Use. And they argue it won't impact the original authors because AI's aren't generating copies or claiming to be the same as the original works.. What do you think? It's up to courts and juries to decide.







There are many kinds of hallucinations, not just it fabricating made up facts. For example, if you ask it "Can cats speak English?" The answer should be "No.", but if you were talking about Garfield or Puss-in-boots, then the answer might be different, so depending on the context the responses could be incorrect. It may also produce language that is contradictory or non-sensical.

Hallucinations could be caused by a number of things.

First, the inherent nature of LLMs is that they were never designed to be accurate. They just statistically generate words that are most probable to occur in particular sequences based on enormous amounts of data. Also, because they are generative, they probabilistically create language. It is in some ways a random generator, so it is inevitable that they will sometimes produce weird outputs.

Another one is the quality of the training data itself. Most chatbots cannot search the web, and even if they could, they may not be able to differentiate between reliable or unreliable sources. The training process and training data may contain biases or inaccuracies as well.

Different LLM may also have different tolerances and "temperatures". Like CoPilot has creative mode. That is, in terms of probability of words appearing together, it may have a wider tolerance for combinations of words that are less likely to appear together. You might actually want this, if you're doing a creative task.

Then the input context is what you put into the chatbot.

The the most sophisticated models are becoming more reliable, have guardrails in place to avoid certain topics, or provide more nuanced responses.



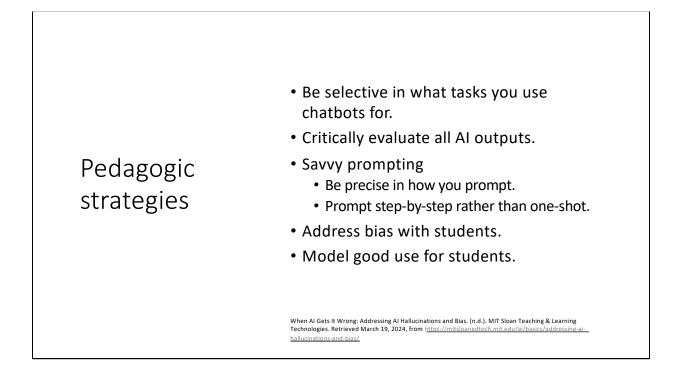
- · Research on bias in AI predates chatbots and has shown them to be racist, sexist, and ableist.

 - Buolamwini, J. (2017). Gender shades: Intersectional phenotypic and demographic evaluation of face datasets and gender classifiers. DSpace@MIT. https://dspace.mit.edu/handle/1721.1/114068
 Safiya Umoja Noble. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press; eBook Collection (EBSCOhost). https://stanford.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db= nlebk&AN=1497317&site=ehost-lve&scope=site
 Our Bodies Encoded: Algorithmic Test Proctoring in Higher Education. (2020, April 2). Hybrid Pedagogy. https://hybridpedagogy.org/our-bodies-encoded-algorithmic-test-proctoring-in-higher-education/ education/
- Many types of biases can manifest or be exacerbated by chatbots.
 - A nuanced view of bias in language models. (2023, September 27). Viden.Al. https://viden.ai/en/a-nuanced-view-of-bias-in-language-models/
- Non-English cultures and languages are excluded in many ways.

 Tidy, J. (2023, November 3). ChatGPT bias: 3 ways non-English speakers are being left behind. Medium. https://medium.com/@ioetidv/chatzol-bias-3-ways-non-english-speakers-are-being-left-behind-299h4898eee6. .

This slide was adapted from Torrey Trust. (2024). GenAl & Ethics: Investigating ChatGPT, Gemini, & Copilot.





Understand the strengths and limits of AI chatbots and be selective. Consider what tasks AI really excel at and what tasks they do not. Chatbots are generally not as good at things like:

- Deep expertise in specialized fields
- Citing specific and credible sources
- Understanding context and nuance
- Complex problem-solving
- Critical thinking and ethical judgement
- Empathy and emotional intelligence

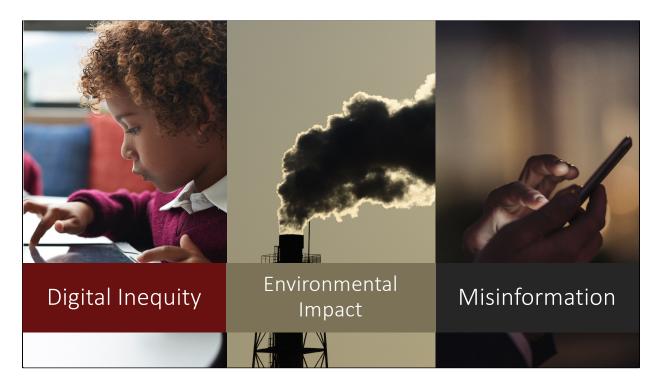
But you might instead, breakdown that big task into smaller tasks that AI can help with such as:

- Acting as a brainstorm partner
- Generating analogies for a concept
- Breaking down a process into steps

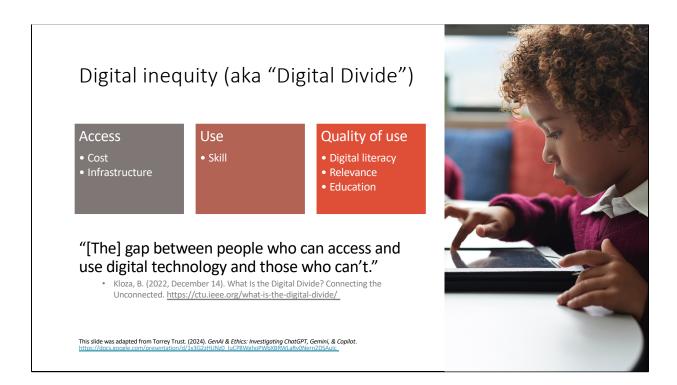
Chatbots are getting better and better with each subsequent version, reducing the likelihood of hallucinations. So consider which chatbots are best suited to your

purpose and try to use the most advanced models. There are also more guardrails in place now. For example, if you prompt for something dangerous or illegal, it will surely refuse to comply and eventually deny access. Or if you tell a chatbot "I love you." It will likely break your heart; refusing to reciprocate.

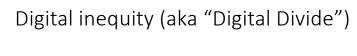
Lastly, always be critical of AI outputs. You are in the driver's seat. Always fact check and use multiple sources.



While not directly impacting



Access Cost Infrastructure Use Skill Quality of use Digital literacy Relevance Education



• LLMs not equitable for those who don't speak English or other major languages

 How language gaps constrain generative AI development. (n.d.). Brookings. Retrieved March 7, 2024, from <u>https://www.brookings.edu/articles/how-language-gaps-constrain-generative-ai-development/</u>

- UK-based survey finds emerging gaps in Al-usage among students by income, gender, ethnicity
 - Freeman, J. (2024, February 1). Provide or punish? Students' views on generative AI in higher education. HEPI. https://www.hepi.ac.uk/2024/02/01/provide-or-punishstudents-views-on-generative-ai-in-higher-education/



This slide was adapted from Torrey Trust. (2024). GenAl & Ethics: Investigating ChatGPT, Gemini, & Copilot:

Environmental impact

- Al uses a lot of energy and resources; better policy is needed
 - OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint", OECD Digital Economy Papers, No. 341, OECD Publishing, Paris, <u>https://doi.org/10.1787/7babf571-en.</u>
- Al energy footprint likely to increase dramatically
 - de Vries, A. (2023). The growing energy footprint of artificial intelligence. Joule, 7(10), 2191–2194. <u>https://doi.org/10.1016/j.joule.2023.09.004</u>
- AI has data centers using more water for cooling
 - Artificial intelligence technology behind ChatGPT was built in Iowa—With a lot of water. (2023, September 9). AP News. https://apnews.com/article/chatgpt-gpt4-iowa-ai-waterconsumption-microsoft-f551fde98083d17a7e8d904f8be822c4
- How impact is measured is not good enough
 - OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The Al footprint", OECD Digital Economy Papers, No. 341, OECD Publishing, Paris, <u>https://doi.org/10.1787/7babf571-en.</u>

This slide was adapted from Torrey Trust. (2024). GenAl & Ethics: Investigating ChatGPT, Gemini, & Copilot.

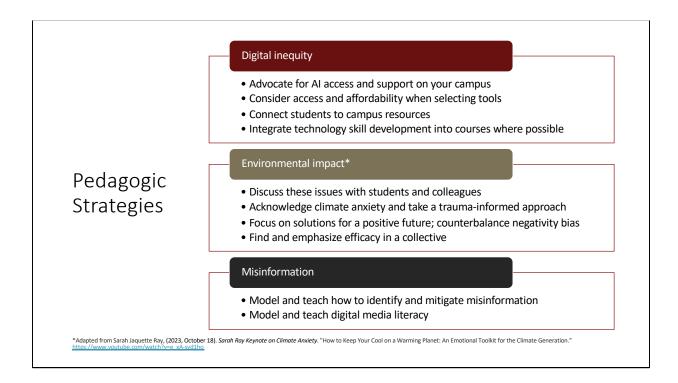


Misinformation

- Media watchdog tracks hundreds of unreliable Algenerated news sources and trends
 - Tracking AI-enabled Misinformation: Over 700 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools. (n.d.). NewsGuard. Retrieved March 7, 2024, from <u>https://www.newsguardtech.com/special-reports/ai-tracking-center</u>
- Al-driven misinformation is biggest short-term threat says World Economic Forum
 - Elliott, L., & editor, L. E. E. (2024, January 10). Al-driven misinformation 'biggest short-term threat to global economy.' The Guardian. https://www.theguardian.com/business/2024/jan/10/ai-driven-misinformation-biggestshort-term-threat-to-global-economy.
- Health disinformation study calls for AI vigilance
 - Menz, B. D., Modi, N. D., Sorich, M. J., & Hopkins, A. M. (2024). Health Disinformation Use Case Highlighting the Urgent Need for Artificial Intelligence Vigilance: Weapons of Mass Disinformation. JAMA Internal Medicine, 184(1), 92–96. https://doi.org/10.1001/jamainternmed.2023.5947

This slide was adapted from Torrey Trust. (2024). GenAl & Ethics: Investigating ChatGPT, Gemini, & Copilot.







Positive examples of these strategies in action at Stanford are all around us if you look:

- There are more AI events happening across campus than you could possibly attend
- Academic technology and Teaching and Learning staff are developing and improving resources all the time
- IT service providers, research labs, and other campus units are developing AI infrastructure and operations as we speak
- More Stanford instructors are trying different ways of addressing AI in their courses every quarter
- Students of all kinds are organizing, inventing, and obsessing over AI with their usual passion and intensity
- World-class Stanford experts are conducting research on AI and developing new inventions
- Heavy hitters with significant budgets like Learning Accelerator, HAI, and so on are offering funding
- You are attending this workshop right now and are super engaged!

I challenge you to name a single better place to be than Stanford University in Silicon

Valley for navigating AI technology. So stay connected. Let's go Cardinal!

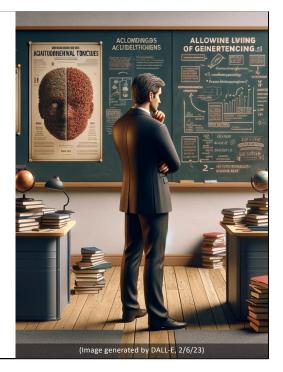


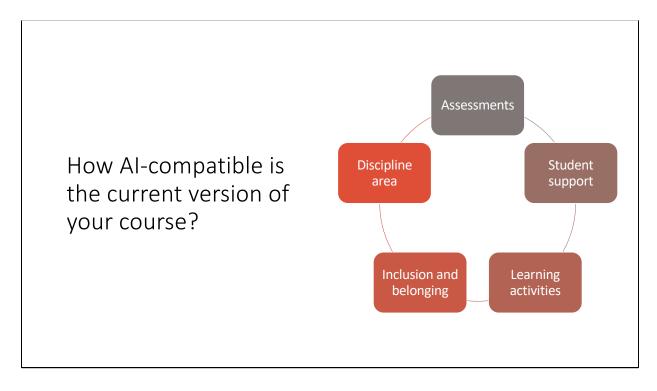
Scenario

You are preparing for a new quarter of teaching. In addition to a mid-term and final exam, you typically have a group assignment where students complete a project based on a real-world issue in your discipline.

Students then submit a written report to describe their project goals, background research, proposed solution, reflection, and so on. Last quarter, many students requested that they be allowed to use generative AI tools for this project and report.

What factors might you consider when thinking about students using AI in your course?





There is a lot to think about. Hopefully this framework can help you think though the academic integrity issues. If you're answering, "Yes" "A lot" or "Very much", then maybe consider more AI. If not, then maybe work on these things first.

Assessments

Measuring how well students learn what you intended and how you assign grades to students

- How clearly have you identified the underlying objectives of your course assignments?
- To what degree do your learning objectives align with higher-order thinking skills, such as creating original work, proposing solutions to complex problems, and internalizing values?*
- How effectively do your current assessments, rubrics, and so on measure your learning objectives?*
- How difficult would it be for an AI chatbot to successfully complete your current assessments?
- How clearly and consistently does your method help you in fairly grading student work?
- To what degree does your course provide multiple opportunities and forms

of assessment and avoid single high-stakes assessments?

- To what degree could integrating AI chatbots make your assessments more compelling or effective?
- How well do current assessments align with students' goals and needs?

Student support

Providing students with what they need to succeed

- How well do you communicate to students what integrity means in your course?
- How clearly do you communicate to students any course or campus policies about academic integrity and AI chatbot use?
- How well might your students know how to use AI chatbots in responsible and honorable ways?
- How well can or do you provide support for students on how to use AI chatbots effectively?
- To what degree are your students independent, experienced, and skilled in self-directed learning with technology tools?
- To what degree does your course promote and do your students leverage relevant support services, such as academic coaches, writing tutors, language partners, and so on?
- To what degree do you and your students understand and consent to the inherent privacy and data security risks that come with using AI tools?

Learning activities

Activities that students do to reinforce learning

- To what degree do ungraded (and therefore not subject to OCS policy) learning activities factor into your course?
- To what degree have students already mastered foundational skills that Al chatbots might augment?
- To what degree could AI chatbots make learning activities more compelling or effective?
- To what degree do you value experience using AI chatbots for students in your course?

Inclusion and belonging

Supporting a wide range of diverse students

- To what degree do you understand the different issues, challenges, and preferences of students typically enrolled in your course?
- To what degree would using AI chatbots benefit students, particularly firstgeneration/low-income, under-represented minority, or students with less academic preparation?
- How flexible do you consider yourself and your course in adapting to the

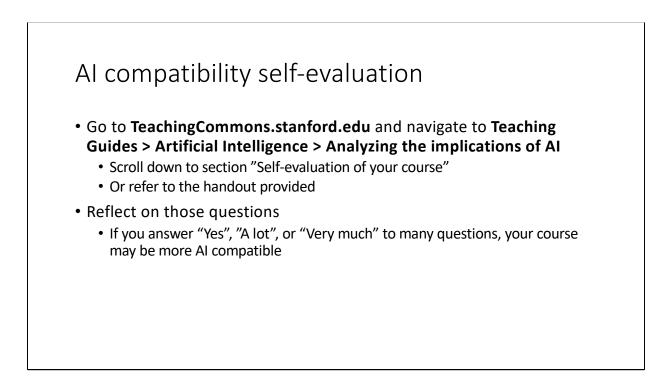
needs of diverse students?

- To what degree can you support equal access for students to AI tools in terms of affordability and accessibility?
- To what degree can you give students and your teaching team informed choices and alternatives in how or if they use AI tools?

Discipline area

Unique characteristics of your discipline area

• How important is it for students in your discipline area to have experience with AI tools or understand AI-related issues?

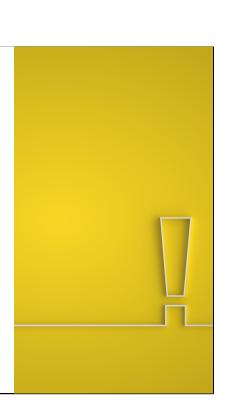


https://teachingcommons.stanford.edu/teaching-guides/artificial-intelligence-teaching-guide/analyzing-implications-ai-your-course



Summarizing major implications of AI in education

- Academic integrity
- Data privacy
- Citation & Attribution
- Copyright & Intellectual Property
- Environmental impact
- Digital inequity
- Misinformation



Truthfulness & Accuracy Privacy & Security Bias and Stereotypes Access Attribution Transparency Support Academic integrity

While there is a lot to unpack here, not to mention broader concerns about AI and its impact on society, economies, politics, ethics, social justice, environmental impacts and so on. That is out of the scope of this workshop.

For our purposes today, we just want to frame this as food for thought, as you begin unpacking some of the implications of AI. When you get further along and are integrating AI into your work in a significant way, there are support services you can connect to

Continue to engage

- Use AI chatbots for your work tasks
- Read the AI Teaching Guide on the Teaching Commons website
- Attend upcoming CTL workshops on AI
- Talk about AI with colleagues and students

