



Leveraging Generative AI to Guide Ethical Behavior

Michelle Darnell, Pennsylvania State University

Tawnya Means, University of Illinois Urbana-Champaign



Takeaways

- **Takeaway 1:** A framework for categorizing ethical issues surrounding the use of generative AI in the classroom
- **Takeaway 2:** A set of guidelines for responsible integration of generative AI to enhance teaching and student engagement
- **Takeaway 3:** A model for using generative AI in crafting scenarios, activities, and ethical dilemmas for use in the classroom.



What ethical concerns do you or your peers have about the responsible use of generative AI in your discipline?

- AI and our humanity
- AI and appropriate use
- AI and responsibility for use
- AI and stakeholder impact



Takeaway 1: A framework for categorizing ethical issues surrounding the use of generative AI in the classroom



Faculty Insecurities: How do I do this?!?

- Who can help me think through this?
- What am I trying to accomplish?
- How can AI support rather than take away from learning?
- Which AI should I use?
- What if AI doesn't 'work' how I want it to?
- How do I introduce this to students?



Addressing Faculty Concerns

Focus 1: Transparency and Ethics

- Clearly communicate AI's role and limitations in education.
- Adhere to ethical standards, regularly auditing for biases and accuracy.



Addressing Faculty Concerns

Focus 2: Student-Centered Approach

- Personalize learning with AI, tailoring content and feedback to individual needs.
- Enhance engagement through interactive and immersive AI-driven experiences.



Addressing Faculty Concerns

Focus 3: Professional Development and Support

- Train and support educators in using AI effectively.
- Promote continuous improvement and sharing of best practices in AI integration.



Takeaway 2: A set of guidelines for responsible integration of generative AI to enhance teaching and student engagement



An abstract digital artwork featuring a dark background with glowing circuitry lines. A vibrant, multi-colored nebula or galaxy core is the central focus, surrounded by swirling energy and numerous small, colorful dots. A white mouse cursor points towards the center of the nebula.

INTEGRATION EXAMPLE 1

Purposeful integration of generative AI in
a business ethics course

Trust as a Cornerstone of Business

The Trust Equation

$$\begin{array}{c} \text{T} \\ \text{Trustworthiness} \end{array} = \frac{\begin{array}{c} \text{C} \\ \text{Credibility} \end{array} + \begin{array}{c} \text{R} \\ \text{Reliability} \end{array} + \begin{array}{c} \text{I} \\ \text{Intimacy} \end{array}}{\begin{array}{c} \text{S} \\ \text{Self-Orientation} \end{array}}$$

Trusted
Advisor
[ASSOCIATES LLC]

Teaching Application:

Evaluate the appropriateness and transparency of the use of AI



1. Create Content with AI

Can you create a fictional company for me in the industry that currently has the most start-ups in the United States?

... can you create a branding message ... that suggests the company will not be guilty of the most common ethical scandals in this industry?

Can you create a branding image that incorporates the ethical principles identified, namely trust, respect, and responsibility?

“ Let's create a fictional company in the fintech sector, as it's one of the most vibrant areas for startups, driven by innovation in payments, personal finance, blockchain, and more. ”



2. Students analyze AI-generated content (without knowing it was generated by AI)



What is a common way that trust could be violated in the FinTech industry?



Does FinLeap (an example FinTech company) address these potential concerns in its messaging?



2. Students analyze AI-generated content (without knowing it was generated by AI)



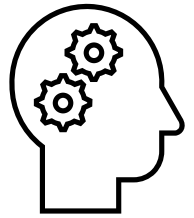
Does FinLeap educate you on potential risks (value extraction from you as a stakeholder) to which you should be paying attention?



Do you believe that FinLeap will stand by its values, based on this messaging? Why or why not?



3. The Big Reveal: How does AI impact our relationships?



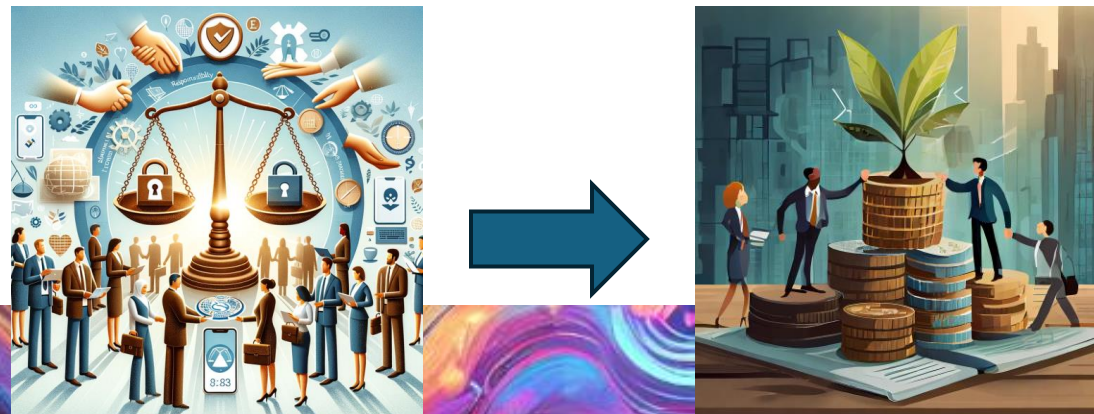
Now that you know the company and branding message was created by ChatGPT, does your confidence that FinLeap will stand by its values increase, decrease, or stay the same?



4. How effective is generative AI as a tool?

Does the provided generated image truly represent the core values that were articulated in FinLeap's branding message? Why or why not?

Try to create a branding image for FinLeap that more accurately portrays the core values that were identified.



	AI Designed Exercise/ Assignment	AI Created Content for Analysis	AI as a tool (decision making tool / deliverables)	Use of AI for developing/ enhancing interpersonal skills or habits
AI and our humanity				
AI and appropriate use				
AI and Responsibility for Use				
AI and Stakeholder Impact				



Takeaway 3: A model for using generative AI in crafting scenarios, activities, and ethical dilemmas for use in the classroom.



Contact Us



Michelle Darnell
814-867-1845
mzd560@psu.edu



Tawnya Means
217-265-5059
tawnya@illinois.edu





APPENDIX

The background of the slide is a complex, abstract digital artwork. It features a dark, almost black, space filled with intricate, glowing patterns. On the left side, there are vertical lines and circuit-like structures in shades of blue and green. The right side is dominated by a large, swirling, multi-colored vortex or nebula-like structure, with a bright yellow and orange core. This central structure is surrounded by various elements: a mouse cursor arrow pointing towards it, a series of colorful, glowing circles (red, orange, yellow, green, blue, purple) at the bottom right, and a network of thin, white, branching lines resembling a neural network or a complex data structure. The overall effect is one of high-tech, futuristic, and digital complexity.

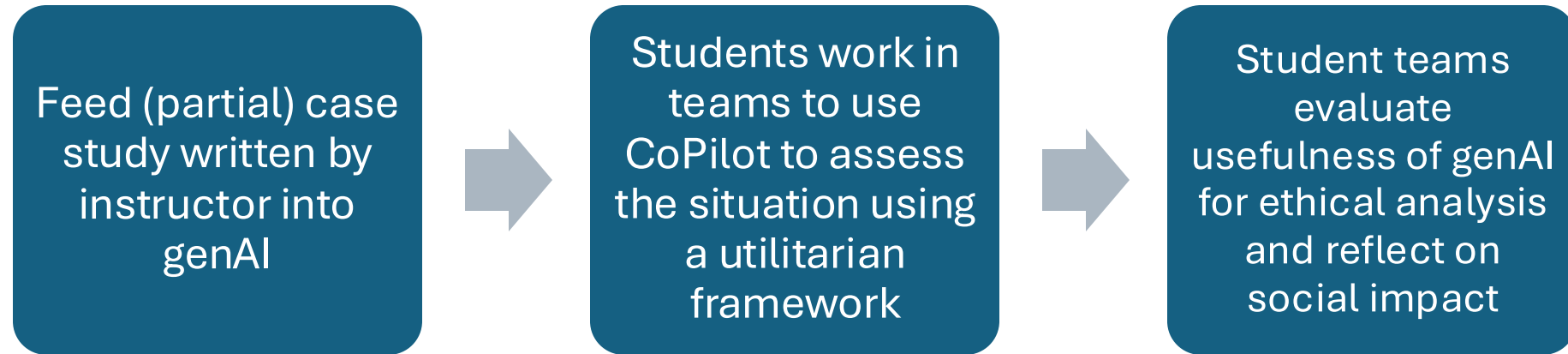
	AI Designed Exercise/ Assignment	AI Created Content for Analysis	AI as a tool (decision making tool / deliverables)	Use of AI for developing/ enhancing interpersonal skills or habits
AI and our humanity	Application 4	Application 3	Application 1 Application 2 Application 3 Application 4	Application 1 Application 4
AI and appropriate use		Application 1	Application 2 Application 3	
AI and Responsibility for Use			Application 2 Application 3	
AI and Stakeholder Impact		Application 3	Application 1 Application 2	

Integration Example 2

Can AI determine what is ethical in a given scenario by applying a Utilitarian framework?

	AI Designed Exercise/ Assignment	AI Created Content for Analysis	AI as a tool (decision making tool / deliverables)	Use of AI for developing/ enhancing interpersonal skills or habits
AI and our humanity			✓	
AI and appropriate use			✓	
AI and Responsibility for Use			✓	
AI and Stakeholder Impact			✓	

Integration Example 2: Use genAI to assess a company's decision to pilot weaponized drones powered by AI.



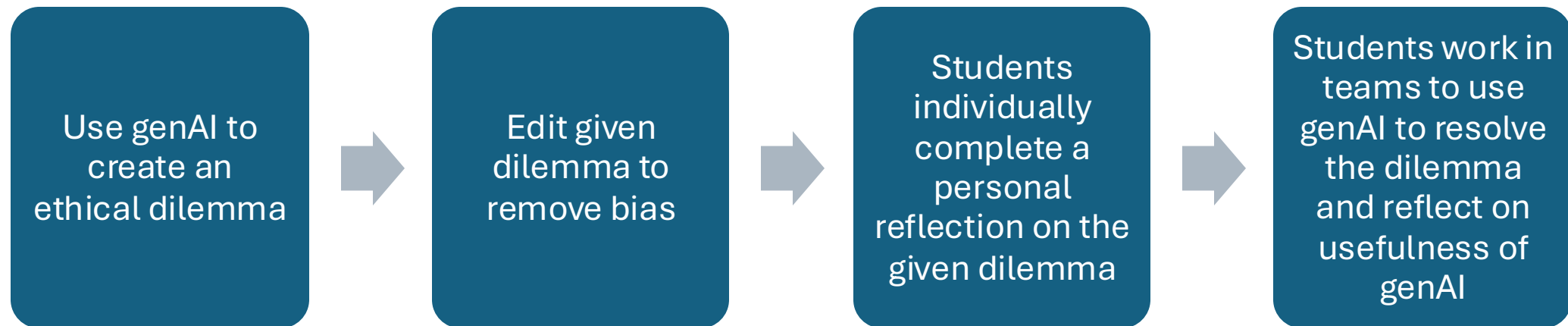
- ? Can genAI replace human ethical decision-making?
- ? Is there any conflict of interest in asking an AI to form a judgment on the creation of something that relies on AI?

Integration Example 3

Can AI help navigate scenarios when core values are in conflict?

	AI Designed Exercise/ Assignment	AI Created Content for Analysis	AI as a tool (decision making tool / deliverables)	Use of AI for developing/ enhancing interpersonal skills or habits
AI and our humanity		✓	✓	
AI and appropriate use			✓	
AI and Responsibility for Use			✓	
AI and Stakeholder Impact		✓		

Integration Example 3: Can AI help navigate scenarios when core values are in conflict?



“Ultimately, the AI pushed back on us and said that depending on our moral values and responsibilities we are the ones to decide at the end of the day” – Team 2

Integration Example 4

AI and the formation of virtuous habits

	AI Designed Exercise/ Assignment	AI Created Content for Analysis	AI as a tool (decision making tool / deliverables)	Use of AI for developing/ enhancing interpersonal skills or habits
AI and our humanity	✓		✓	✓
AI and appropriate use				
AI and Responsibility for Use				
AI and Stakeholder Impact				

Integration Example 4: Forming virtuous habits

