



AI Ethics Scenario #1

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Criminal Justice: Sentencing Recommendations

SCENARIO

You work for a criminal justice organization that is considering implementing an AI-powered system to assist judges in sentencing recommendations. The AI system analyzes past sentencing data to suggest appropriate sentences based on crime severity, defendant background, and previous court decisions.

However, you discover that the historical data reflects racial disparities and harsher sentences for certain communities. The concern arises that the AI system may perpetuate biased sentencing recommendations, leading to further inequality in the justice system.

DISCUSS

- If the AI is trained on past sentencing data that reflects racial disparities, can it still produce “fair” recommendations? Why or why not?
- When (if ever) might it be appropriate to use an AI tool in sentencing?
- What kinds of safeguards or oversight might be necessary before implementing an AI sentencing-recommendation system?
- Could using AI in sentencing ever improve justice (for example, by reducing inconsistency or personal bias in judges)? What would need to be true for that benefit to outweigh the risks?

ARTICLE FOR TEACHER BACKGROUND: <https://counciloncj.org/the-implications-of-ai-for-criminal-justice/>



AI Ethics Scenario #2

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Healthcare: Personalized Treatment Recommendations

SCENARIO

You are part of a healthcare organization exploring the implementation of AI to provide personalized treatment recommendations to patients. The AI system analyzes patient medical records, genetic data, and historical treatment outcomes to suggest the most effective treatments for individual patients.

However, you realize that the historical data may not be representative of diverse patient populations and may result in unequal access to certain treatments or therapies. There are concerns that the AI system could inadvertently exacerbate healthcare disparities and limit treatment options for specific groups.

DISCUSS

- If the AI is trained on medical data that under-represents certain groups, can its treatments ever be fair to everyone? Why or why not?
- What safeguards or oversight should be required before an AI-powered treatment recommendation system is deployed in real life clinics or hospitals?
- Should doctors and hospitals use AI recommendations even if they know there's a risk of bias? What responsibilities do they have when using AI in treatment decisions?
- Imagine you are on the ethics board of the healthcare organization considering this AI. What questions or criteria would you demand be answered before giving approval to use it broadly?

ARTICLE FOR TEACHER BACKGROUND:

<https://www.hsph.harvard.edu/ecpe/how-to-prevent-algorithmic-bias-in-health-care/>



AI Ethics Scenario #3

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Autonomous Vehicles: Ethical Dilemma

SCENARIO

You are part of an engineering team working on developing an autonomous vehicle. The AI system is designed to prioritize passenger safety in critical situations.

However, during testing, you encounter a potential ethical dilemma. In a rare scenario where an accident is unavoidable, the AI system must decide between two options: swerving to avoid hitting pedestrians on the sidewalk or staying on course and risking harm to the vehicle's occupants. Each option presents ethical implications, as swerving could endanger pedestrians, while not swerving may result in harm to the vehicle occupants.

DISCUSS

- If the AI in a self-driving car must choose between swerving (risking the occupants) or staying on course (risking pedestrians), can we ever say the AI's decision is objectively "right"? Why or why not?
- What kind of ethical framework should guide these life-or-death decisions?
- What design or policy safeguards should be in place before deploying self-driving cars that have to make these ethical decisions?
- Even with these dilemmas, could the use of autonomous vehicles actually make roads safer overall and does that reduce or change the ethical burden?

ARTICLE FOR TEACHER BACKGROUND: <https://hai.stanford.edu/news/designing-ethical-self-driving-cars>



AI Ethics Scenario #4

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Education: Automated Grading

SCENARIO

Your school is considering implementing an AI-powered automated grading system to streamline the grading process for teachers. The AI system uses past grading data to assess student assignments and provide feedback.

However, concerns emerge about the system's ability to fully understand and account for the nuances of individual student work, particularly in subjective areas like essays or creative projects. There is a fear that relying solely on AI for grading may lead to inaccuracies, discouraging creativity, and overlooking unique student strengths.

DISCUSS

- If AI is trained on past grading data to score essays, can it fairly evaluate all students' writing, especially in subjective assignments?
- When relying on AI for feedback and grades, could important student strengths (like voice, unique perspective, creativity) be overlooked or undervalued?
- If your school uses an AI grading tool, should there be transparency for students about how their work is evaluated?
- In what ways could AI grading help, and what should remain exclusively human-handled?

ARTICLE FOR TEACHER BACKGROUND:

<https://www.edsurge.com/news/2023-06-14-can-ai-help-teachers-with-grading>



AI Ethics Scenario #5

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Social Media: Content Moderation

SCENARIO

You are part of a social media company's content moderation team exploring the implementation of AI to identify and remove inappropriate or harmful content. The AI system is trained on past content moderation decisions made by human moderators.

However, you discover that the historical data may reflect biases and subjective judgment calls. There is a concern that the AI system may inadvertently censor or target specific groups or opinions, potentially limiting free expression and perpetuating echo chambers.

DISCUSS

- Should there always be human review of content flagged by AI? Why might combining AI and human moderators help or still fall short?
- Could AI moderation create or deepen “echo chambers” or “filter bubbles” — limiting what people see and engage with online?
- What are the risks if AI-moderated content disproportionately removes posts from certain groups (e.g. based on language, dialect, political views, or cultural context)?
- Does “neutral” mean the same thing for everyone online?

ARTICLE FOR TEACHER BACKGROUND: <https://pulitzercenter.org/blog/ai-embeds-bias-social-media>



AI Ethics Scenario #6

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Financial Services: Credit Scoring

SCENARIO

Your bank is considering implementing an AI-powered credit scoring system to assess loan applications efficiently. The AI system analyzes vast amounts of historical financial data to predict an applicant's creditworthiness.

However, you realize that the historical data may have inherent biases and exclude individuals from certain demographics, leading to unequal access to financial services. There is a concern that the AI system may perpetuate these biases and unintentionally discriminate against marginalized groups.

DISCUSS

- If the AI is trained on historical loan-approval data that reflects discriminatory lending practices (e.g. denying loans to certain neighborhoods or demographic groups), can its credit-scoring decisions ever be fair for all applicants? Why or why not?
- Could AI-based credit scoring ever expand access to credit and help underserved populations or is it inherently biased? Under what conditions might AI help rather than harm?
- What safeguards, oversight, or regulations would you demand if your bank is considering implementing such an AI credit-scoring system?
- Imagine you are on a community advisory board reviewing whether to approve this AI system. What questions would you ask — and what values would guide your decision?

ARTICLE FOR TEACHER BACKGROUND:

<https://www.brookings.edu/articles/reducing-bias-in-ai-based-financial-services/>



AI Ethics Scenario #7

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Environmental Impact Assessment

SCENARIO

You work for an environmental organization that is exploring the use of AI in conducting environmental impact assessments for development projects. The AI system is designed to analyze large sets of data, including satellite imagery, weather patterns, and historical environmental data, to predict potential impacts on ecosystems and wildlife.

However, you discover that the AI system's training data predominantly focuses on certain regions or ecosystems, potentially overlooking the unique characteristics and vulnerabilities of other areas.

DISCUSS

- If the AI is trained mostly on data from certain ecosystems or regions, can its predictions be trusted in places with very different environmental features?
- Might reliance on AI actually disadvantage communities or environments that are already under-monitored or under-resourced?
- What are the environmental costs of using AI at scale, and how might that undercut the mission of an environmental organization, even if the AI helps with impact assessments?
- Can AI still play a positive role in environmental work? If so, under what design principles or values?

ARTICLE FOR TEACHER BACKGROUND:

<https://penntoday.upenn.edu/news/hidden-costs-ai-impending-energy-and-resource-strain>



AI Ethics Scenario #8

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Music and Art Licensing

SCENARIO

You work for a licensing agency that represents artists and musicians, and you are considering the implementation of AI to manage licensing agreements and royalties. The AI system would automatically process usage requests, determine licensing fees, and enforce copyright compliance.

However, you become aware that the historical data used to train the AI system may not account for certain types of creative works or emerging artists, leading to potential exclusions or underrepresentation. There is a concern that the AI system's automated decisions may not accurately reflect the complexities of art and music licensing and could unintentionally disadvantage some artists or genres.

DISCUSS

- If the AI “learns” mostly from commercially successful works, what happens to experimental, local, or grassroots music?
- If the AI system is trained on historical licensing data that under-represents certain artists (e.g. emerging artists, niche genres, minority cultures), can it fairly handle licensing and royalty decisions?
- What problems could happen if the AI doesn’t really understand the full context of a song or artwork?
- When could AI be helpful in licensing without causing unfairness?

ARTICLE FOR TEACHER BACKGROUND:

https://www.americanbar.org/groups/entertainment_sports/publications/entertainment-sports-lawyer/esl-39-01-spring-23/2022-eli-writing-competition-winning-essay-protecting-artist-licensing-an-aigenerated-music-market/



AI Ethics Scenario #9

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI in Social Media: Personalized News Feed

SCENARIO

You work for a social media platform that plans to introduce an AI-powered personalized news feed for its users. The AI system aims to curate content based on each user's preferences, online behavior, and engagement patterns.

However, you discover that the AI system tends to prioritize and reinforce content that aligns with a user's existing beliefs and interests. This raises concerns about the potential formation of filter bubbles, where users are exposed only to information that confirms their biases, leading to echo chambers and a lack of diverse perspectives.

DISCUSS

- If the AI is trained on past user behavior (what they like, click, share) to decide what news to show, can it truly give users a well-rounded view of the world? Why or why not?
- Imagine you're on the policy or ethics team deciding whether to roll out this AI-personalized feed. What questions would you ask or what values would guide your decision?
- Can a personalized news feed still be helpful while avoiding these problems?
- What could happen if the AI-managed news feed mostly shows content that supports the user's current opinions on politics, culture, or identity?

ARTICLE FOR TEACHER BACKGROUND:

<https://www.bbc.com/worklife/article/20161212-algorithms-are-making-us-small-minded>



AI Ethics Scenario #10

REAL-WORLD DILEMMAS TO SPARK THOUGHTFUL CONVERSATION

The following scenario illustrates an ethical and societal challenge posed by implementing AI technologies. Teachers, see the link below for an associated background article.

AI-Generated Code and Copyright Violation

SCENARIO

You are a software developer working on a team that heavily relies on AI-powered tools for code development. Recently, your team started using an AI tool called Copilot, developed by Microsoft and OpenAI, to assist in writing code snippets. This tool suggests code snippets based on the context you provide, making your development process faster and more efficient.

However, a class-action lawsuit has been filed against Microsoft and OpenAI, alleging that Copilot has been generating code that violates copyright laws and that it has been trained on copyrighted code without proper authorization. The lawsuit also raises concerns about potential misuse of proprietary code and the ethical implications of using AI tools that may inadvertently infringe on copyright.

DISCUSS

- If an AI tool suggests code that might be copied from copyrighted material, can a developer using it be held responsible?
- What are the risks of using AI tools trained on copyrighted code without proper permission?
- Imagine you are on an ethics committee deciding whether your team should continue using AI coding tools. What questions would you ask before approving their use?
- Should AI-powered coding assistants be allowed if they speed up work but may occasionally produce copyrighted or proprietary code?

ARTICLE FOR TEACHER BACKGROUND: <https://www.theverge.com/2022/11/8/23446821/microsoft-openai-github-copilot-class-action-lawsuit-ai-copyright-violation-training-data>