

AI Deepfakes: The Impact on Truth in Media

Why?	
<p>This lesson raises awareness about the ethical implications and societal impact of deepfake technology, a rapidly evolving field in artificial intelligence. It teaches critical thinking and media literacy skills, enabling students to consider how they would tell the difference between real and manipulated content in an era where digital misinformation is prevalent. By exploring the challenges of deception and authenticity in digital media, the lesson underscores the importance of ethical standards and personal rights in the digital landscape. Additionally, it fosters a deeper understanding of the balance needed between technological innovation and maintaining public trust, preparing students to navigate and contribute responsibly to our increasingly digital world.</p>	
Materials Needed	Time needed
<ul style="list-style-type: none"> ● Case study handout ● Projector for video ● Computers or tablets for simulation 	<ul style="list-style-type: none"> ● Approximately 60 - 90 minutes
Objectives	
<ul style="list-style-type: none"> ● Students will be able to explain the concept of deepfakes and how they are created. ● Students will be able to critically analyze and articulate the ethical implications of deepfakes in media. ● Students will be able to evaluate and respond to a real-world scenario involving deepfakes. 	
Key Concepts & Vocabulary	
<ul style="list-style-type: none"> ● Deepfake: A highly realistic video made using artificial intelligence techniques to superimpose one picture over another, typically to make it appear that a person is saying or doing something they did not. 	
Lesson Components	
<ol style="list-style-type: none"> 1. Before You Watch: Connect lesson to background knowledge about deepfakes and get students' attention 2. Video: Show the pedagogy.cloud video explaining the ethical considerations in the topic of deepfakes. 3. Case Study: Detail a real-world scenario that discusses a news website posting a controversial deepfake video. 4. Simulation: Lead students through an interactive activity exploring the possible ethical considerations in determining how to respond when a deepfake video is posted. 	

5. **Discussion:** Ask whole-class questions to reflect on experience and consider perspectives.
6. **Assessment:** Verify student understanding with an exit ticket

1. Before You Watch

Video Introduction: Show a [short clip](#) that has a real video and a deepfake side by side. Ask students to compare the two to see if they can identify which one is the deepfake, and how they know. Ask how much students have seen deepfakes, and if they have any specific stories.

While You Watch: Mention these topics and questions for students to look out for as they watch the video:

- Define “deepfake”
- Note the reasons for why a deepfake in the video impacted media or truth.
- What is a concern about deepfakes past the basic fact that it is not real?

2. Video Summary

The video highlights the profound implications of deepfakes in distorting reality and creating confusion between truth and fabrication, especially in journalism, politics, and personal relationships. A case study of a fabricated speech by a mayor demonstrates the potential damage to reputations, public trust, and media credibility. Ethical concerns are raised, focusing on issues of consent and authenticity, and the necessity of detecting and preventing such manipulations. While acknowledging the positive uses of deepfakes in entertainment and education, the video emphasizes the need for balancing innovation with ethical considerations. It concludes by advocating for collective responsibility in media literacy and critical engagement with digital content, underscoring the importance of maintaining authenticity and trust in our digital world.

3. Case Study

Distribute or read Case Study handout.

Summary: This case study examines the consequences of a local news website unknowingly publishing a deepfake video of a mayor's speech, leading to intense public debate and division over its authenticity. The situation escalates as the community is split, with some believing the mayor's denial and others convinced by the video, ultimately raising concerns about the erosion of public trust and the difficulty in distinguishing truth in the digital age. The incident results in the news organization issuing an apology and adopting stricter content verification processes, sparking wider discussions about ethical journalism, the need for public education in digital literacy, and regulation of deepfake technology.

4. Simulation

Group Formation: Divide students into small groups, each representing a different stakeholder in the deepfake scenario above (media organization, government body,

tech company).

If desired, students could be given time on computers to research more information about deepfake technology and proposed solutions.

Within each group, students could optionally be assigned specific roles for the initial discussion:

- Media Organization: Editor-in-Chief; Investigative Journalist; Social Media Manager who posted the video in the first place; Copy Editor / Fact Checker; Legal Advisor
- Government Body: The Mayor; City Council Member(s); County Judge; Public Relations Officer; Legal Counsel
- Tech Company: CEO; AI Researcher; Public Relations Manager; Security Expert; Legal Advisor

Scenario Presentation: Present a detailed scenario (see below) involving the release of a deepfake video, similar to the case study provided earlier. Outline the challenges and decisions each group must face. Each group is tasked in a handout with determining the appropriate response to the deepfake video, and establishing policies going forward.

Group Role Play: Each group role-plays their response to the scenario, making decisions based on their research and discussion. Encourage groups to explore different outcomes that could result from their decisions.

If available, students could use computers or tablets to research this issue to help them determine appropriate responses for their organizations.

Whole Class Discussion: Groups share their strategies and outcomes with the class. Discuss the real-world implications of deepfakes and the importance of ethical decision-making. Facilitate a discussion on the various group approaches, their effectiveness, and ethical considerations.

5. Discussion

These questions are designed to be used in whole-class discussion. Ask questions that relate most effectively to the lesson.

1. What are deepfakes, and how can they be identified?
2. How do deepfakes affect our perception of truth in media?
3. What responsibilities do media organizations have in verifying content?
4. What laws or regulations, if any, should govern the use of deepfakes in media?
5. How much of the blame for this controversy is on the public for believing that deepfakes are real?
6. Are there legitimate uses for deepfake technology? How can the line be drawn between legitimate uses and inappropriate uses?

6. Assessment

Exit Ticket: Provide a prompt for students to reflect on their learning, such as:

- How has your understanding of deepfakes and their impact on media changed?
- What is one thing you can do to be a more discerning media consumer?

Sources to Learn More

- [Ctrl Shift Face](#) - A YouTube channel that makes deepfakes replacing actors' faces with other people's faces in movie scenes
- [Deepfake video](#) depicted Ukrainian president Volodymyr Zelenskyy telling his soldiers to lay down their weapons and surrender – Seen as a brand new “battleground” in war

Case Study: A Deepfake Dilemma

Introduction: This case study explores the impact of deepfake technology on public trust and media integrity. It centers around a local news website that unknowingly publishes a deepfake video of the mayor's speech.

Scenario: In a small yet politically active community, a local news website releases a video depicting the mayor giving a highly controversial speech, which includes some incendiary remarks and policy proposals. The authenticity of the video rapidly becomes a topic of heated debate among the public. The mayor vehemently denies making such a speech, claiming the video is a fabrication. An investigation is launched, which later confirms that the video is indeed a deepfake, a sophisticated AI-generated falsification.

Community Reaction: The community is sharply divided in its response to the incident. Some residents are quick to believe the mayor's denial, citing his previous track record and character, while others view the video as credible, leading to a surge in mistrust towards the mayor. Some people saw the video but not the response, and fully believe the mayor should be kicked out of office. The revelation that the video is a deepfake only adds to the confusion, with a segment of the population questioning the ability to discern truth in the digital age. Social media platforms become battlegrounds for debates on the implications of AI in media and the erosion of public trust.

The Debate: Opinions vary widely among community members and experts. Some argue that the news organization should be held accountable for not verifying the authenticity of the video before its release, highlighting the need for stricter standards in journalism. Others focus on the dangers posed by deepfake technology, calling for regulations to manage its use. Meanwhile, a few advocate for greater public education on digital literacy to combat the spread of such deceptive content.

Conclusion: The news organization behind the release asserts that they were not aware of the video's fabricated nature at the time of publication. The case culminates in the news organization issuing a public apology and implementing stricter verification processes for content. It sparks a broader discussion on ethical journalism and the need for awareness and regulations surrounding deepfake technology.

Questions

- How did the release of the deepfake video affect the public's trust of both the mayor and the local news organization? What could be long-term consequences of this video?
- Would you support government regulation of news and social media content for the purpose of preventing deepfakes from tricking people?

Simulation: Organizational Policy Meetings

Deepfakes are hyper-realistic videos created using artificial intelligence and machine learning technologies. They can make people appear to say or do things they never did, posing significant challenges to truth and trust in digital media.

In a small community, a local news website publishes a video showing the mayor giving a controversial speech, filled with provocative remarks. Upon finding out about the controversy, the mayor denies ever making such a speech, asserting the video is a fake. Subsequent investigation reveals the video to be a deepfake, leading to public debate and confusion over the authenticity of digital content and the integrity of both the mayor and the media.

General Instructions

You will be discussing the issue of deepfake videos within a particular organization's policy meeting. Each group will use their discussion time to develop a strategic approach based on their roles and perspectives. Consider the ethical, legal, and societal implications of your decisions. Your aim is to collaboratively form a comprehensive response to the deepfake scenario presented.

The organizations are listed on the back of this page. If your teacher decides to use individual roles, those are explained as well.

Questions to help guide your discussion

Ethical: What standards of truth and authenticity should exist in digital media? What ethical responsibilities do your organization, and the other organizations, have in preventing the spread of false material? How much should right/wrong guide the response to this incident?

Legal: Who is responsible for the false content? Should anyone be held liable in court? How might existing laws apply to deepfakes?

Societal: How do deepfakes affect public trust in media, government, and technology? What are the potential long-term impacts of deepfakes? How can society adapt to challenges posed by deepfakes? What will help the public be better informed?

Group and Individual Roles

Group 1: The media organization that posted the video

Purpose in Discussion:

Develop a strategy to verify and report on deepfake content, balancing ethical reporting and audience engagement.

(Optional) Roles:

Editor-in-Chief: Oversees publication decisions and sets journalistic standards.

Investigative Journalist: Pursues complex stories, seeking to uncover the truth.

Social Media Manager: Manages digital content, responsible for initially posting the video.

Copy Editor / Fact Checker: Verifies the accuracy and integrity of news content.

Legal Advisor: Provides legal guidance on publication ethics and potential liabilities.

Group 2: The local government

Purpose in Discussion: Formulate a response plan and policy recommendations to address the challenges posed by deepfakes.

(Optional) Roles:

The Mayor: Faces the direct impact of the deepfake and must address public concerns.

City Council Member(s): Deliberate on the implications of the deepfake for local governance. Considers the mayor's long-term chances of being able to lead the city in the aftermath of the video.

County Judge: Considers the legal aspects and potential legal actions related to the deepfake.

Public Relations Officer: Manages communication and public perception of the government's response.

Legal Counsel: Advises on legal strategies and implications for the government.

Group 3: Tech company that made the technology

Purpose in Discussion:

Propose technological solutions and corporate strategies to detect, prevent, and address the implications of deepfakes.

(Optional) Roles:

CEO: Leads the company, making strategic decisions in response to the deepfake challenge.

AI Researcher: Focuses on developing technology to detect and counteract deepfakes.

Public Relations Manager: Handles external communications and maintains the company's public image.

Security Expert: Advises on digital security measures related to deepfake detection and prevention.

Legal Advisor: Provides legal advice, ensuring compliance and addressing regulatory issues.

Video Script for Animator

Hello Young Innovators! Today we're discussing the ethics of deepfakes.

Title screen

In an era where seeing is believing, what happens when the videos we watch can be manipulated to show something that never happened? Welcome to the world of deepfakes.

Socrate narrating throughout Opening shot of Socrate looking at a tablet playing a video]

Deepfakes are videos created using artificial intelligence to replace a person's likeness and voice with someone else's. It's a technology that has both fascinated and alarmed people worldwide. Imagine watching a video of a famous actor saying words they never uttered, or a politician making a speech they never gave. That's the power, and the peril, of deepfakes.

[Tablet with a video playing takes up much of the screen, The face in the video is then replaced by another completely different face]

The implications for media are profound. Deepfakes can distort reality, creating confusion about what's real and what's fabricated. In journalism, politics, and even personal relationships, the line between truth and falsehood becomes blurred.

[Show a group of people watching a screen, and reacting with surprise or shock - Socrate standing off to the side and looks at video once or twice. Example - https://media.istockphoto.com/id/1210861885/photo/group-of-people-watching-the-phone-screen.jpg?s=1024x1024&w=is&k=20&c=CEq_6LNVyPLQTKeNWTqvFmVhuPJ9uWYkgILg9eqt6KQ=]

Consider the case of a local mayor accused of giving a controversial speech. A video circulates, tensions rise, but then it's revealed to be a deepfake. The damage, however, is already done. Trust in media dwindles, and the fabric of community trust is strained.

[A politician is giving a speech at a lectern while a video camera is recording; Then the same video appears on cell phones, but image is changed somehow. Example - <https://www.wikihow.com/images/thumb/a/a6/Run-for-Mayor-Step-2-Version-3.jpg/v4-460px-Run-for-Mayor-Step-2-Version-3.jpg.webp> - needs camera set up taking video]

The ethical considerations of deepfakes are immense. These videos may show someone who didn't know or approve of the message being delivered. This can lead to widespread misinformation and manipulation. The lines between reality and fabrication are blurred. How can truth and factual accuracy be preserved in the digital age?

[Side-by-side footage of a real video and a deepfake: Two separate screens of a person talking, where everything is the same except the face on one of the videos is clearly different]

Detecting deepfakes is a race against technology. As deepfake technology evolves, so do the methods to detect them. Tech companies, governments, and individuals must collaborate in order to foster transparency and verification.

[Socrat watching a large screen. Videos appear on screen one at a time and are “evaluated,” as a green checkmark or a red X superimposes on the screen]

Media organizations bear a responsibility to verify content. But we, as consumers, must also be vigilant. Critical thinking, fact-checking, and awareness of deepfakes are essential tools in discerning truth from fabrication.

[Visuals of a busy newsroom where workers are looking at videos, depicting the challenge faced by media. Example - [0https://thecictravels.cic.sc.edu/wp-content/uploads/2019/05/meoli2-1-1024x623.jpg](https://thecictravels.cic.sc.edu/wp-content/uploads/2019/05/meoli2-1-1024x623.jpg)]

Deepfakes aren't all bad. They have applications in entertainment, education, and more. But the journey towards harnessing their potential without sacrificing trust is a complex one. It requires ongoing dialogue, legal consideration, technological innovation, and societal awareness.

[A group of people talking around a table - Socrat could be one of the people at the table. Example - https://cloudinary.hbs.edu/hbsit/image/upload/s--AolbzAdd--/f_auto,c_fill,h_375,w_750,/v20200101/F42D9CA3707828FD74484C2B9EE51CFF.jpg]

Deepfakes challenge our perception of reality. As we navigate this digital landscape, our collective responsibility is to engage with media thoughtfully and critically. It's not just about technology; it's about preserving trust, authenticity, and the very nature of truth in our interconnected world.

[Show visuals from earlier shots in the video, including Tablet, Mayor, Meeting]

Let's discuss: If you were the subject of a deepfake video made without your consent, what concerns would you have regarding your personal rights, reputation, and privacy? What steps would you want to be able to take to remedy the situation?

[Discussion Question: If you were the subject of a deepfake video made without your consent, what concerns would you have regarding your personal rights, reputation, and privacy? What steps would you want to be able to take to remedy the situation?]



Question text appears on screen

Video Script for Narrations

Hello Young Innovators! Today we're discussing the ethics of deepfakes.

In an era where seeing is believing, what happens when the videos we watch can be manipulated to show something that never happened? Welcome to the world of deepfakes.

Deepfakes are videos created using artificial intelligence to replace a person's likeness and voice with someone else's. It's a technology that has both fascinated and alarmed people worldwide. Imagine watching a video of a famous actor saying words they never uttered, or a politician making a speech they never gave. That's the power, and the peril, of deepfakes.

The implications for media are profound. Deepfakes can distort reality, creating confusion about what's real and what's fabricated. In journalism, politics, and even personal relationships, the line between truth and falsehood becomes blurred.

Consider the case of a local mayor accused of giving a controversial speech. A video circulates, tensions rise, but then it's revealed to be a deepfake. The damage, however, is already done. Trust in media dwindles, and the fabric of community trust is strained.

The ethical considerations of deepfakes are immense. These videos may show someone who didn't know or approve of the message being delivered. This can lead to widespread misinformation and manipulation. The lines between reality and fabrication are blurred. How can truth and factual accuracy be preserved in the digital age?

Detecting deepfakes is a race against technology. As deepfake technology evolves, so do the methods to detect them. Tech companies, governments, and individuals must collaborate in order to foster transparency and verification.

Media organizations bear a responsibility to verify content. But we, as consumers, must also be vigilant. Critical thinking, fact-checking, and awareness of deepfakes are essential tools in discerning truth from fabrication.

Deepfakes aren't all bad. They have applications in entertainment, education, and more. But the journey towards harnessing their potential without sacrificing trust is a complex one. It requires ongoing dialogue, legal consideration, technological innovation, and societal awareness.

Deepfakes challenge our perception of reality. As we navigate this digital landscape, our collective responsibility is to engage with media thoughtfully and critically. It's not just about technology; it's about preserving trust, authenticity, and the very nature of truth in our interconnected world.

Let's discuss: If you were the subject of a deepfake video made without your consent, what concerns would you have regarding your personal rights, reputation, and privacy? What steps would you want to be able to take to remedy the situation?