

Do Now

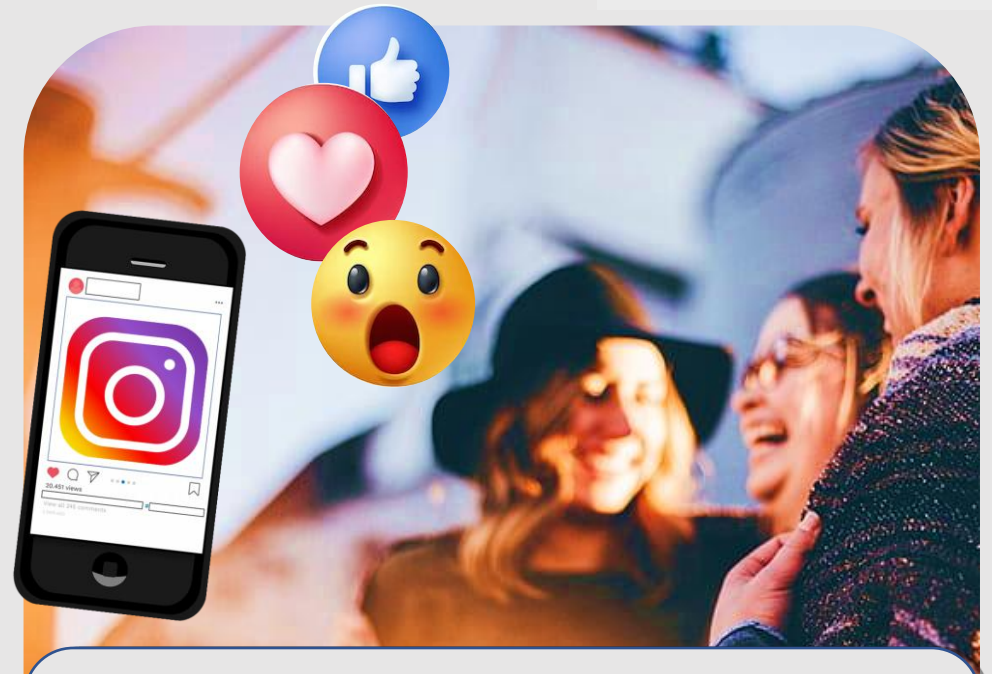
Lesson 4: : malicious use, the law and staying safe



Three girls from Amy's class were in the school corridor, laughing at a video which was sent to them in a group message. When Amy walked over, the three looked terrified and quickly put their phones away.

Amy's heart sank. She heard a bit of the video, and it sounded like her voice. But she didn't remember saying those words.

Discuss: What might the three girls have been watching? What should Amy do about this?



Define the term AI and what we might mean when we talk about 'deepfakes.'

AI is..... When we talk about deepfakes we mean...

Deepfakes: malicious use, the law and staying safe

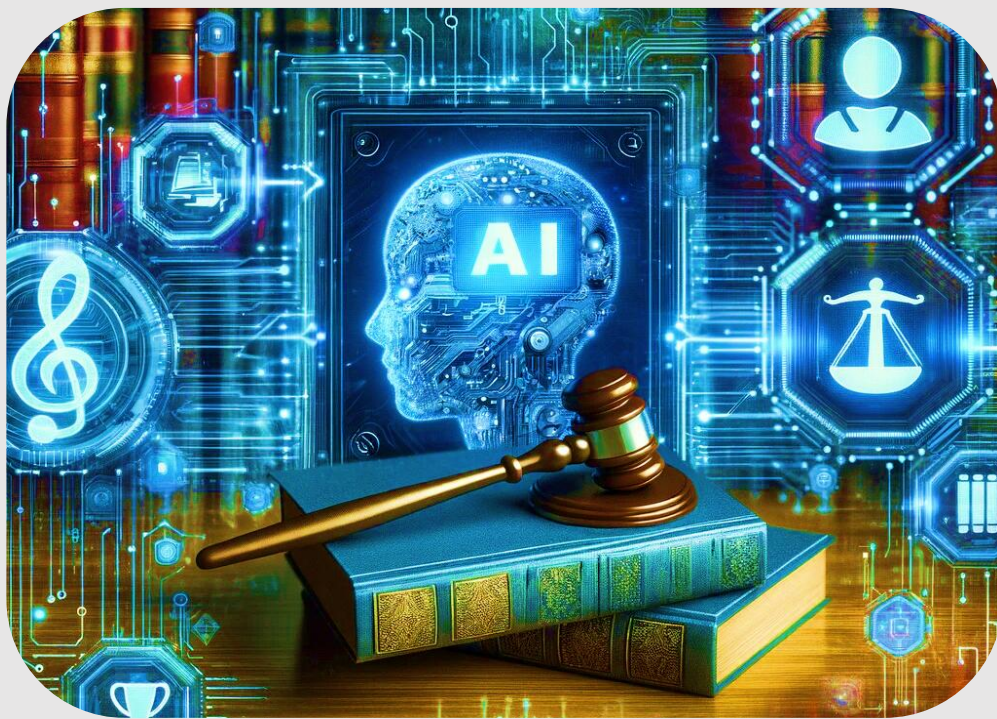
Learning outcomes:

Explain how new technology produced by artificial intelligence, such as deepfakes, can be used to malicious ends, as well as for entertainment purposes.

Describe the legal penalties for malicious deepfakes and where to go for help and support.

AI: Artificial intelligence is the science of making machines that can think like humans. This could be used to create images, text, video footage or music.

Deepfakes : An image or recording that has been convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said.



Did you know?

It was only April 2024 that the UK Government outlawed the creation and sharing of explicit deepfakes. **Amy's predicament could have been much worse...**

<https://www.youtube.com/watch?v=eZon6XQoYv8>



Amy was right to be worried. She reported what she heard to her Head of Year, and the girls were called into a meeting.

All were unaware of the new UK law against the creation of malicious deepfakes, which are often **'immoral, often misogynistic, and a crime.'**

(UK Department of Justice).

Fortunately, after seeing the video contents, the HoY found it was not an explicit video, but it was poking fun at Amy and using her voice and image without consent.

The girls were punished through the school's anti-bullying policy.

Deepfakes: malicious use, the law and staying safe

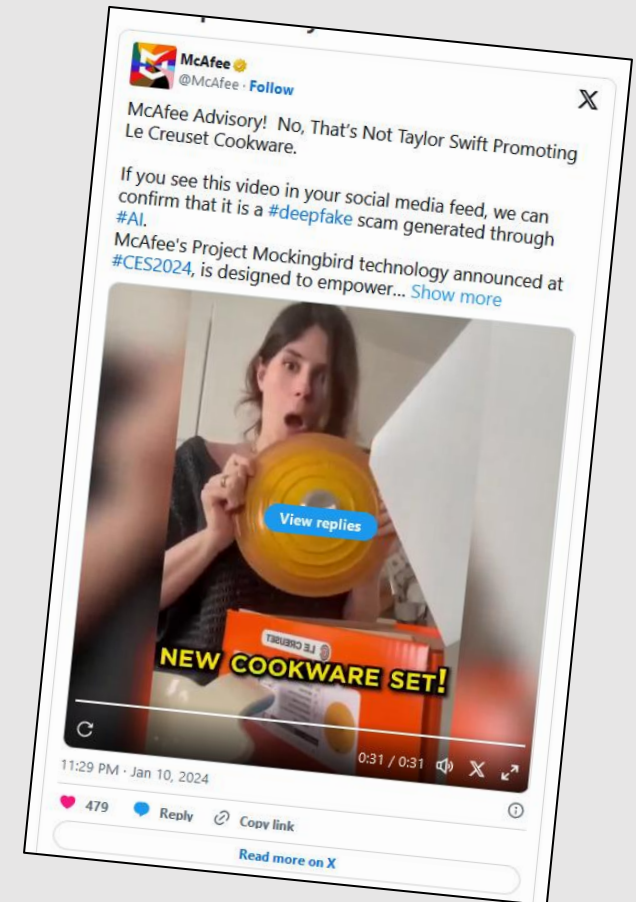


How convincing are deep fakes?

Watch the second video on this page, of Taylor Swift supposedly advertising cookware here:

<https://www.creativebloq.com/features/deepfake-examples>

DISCUSS: As you watch, be ready to feedback:
How might this have misled people?
Could you tell it was fake? How?



Deepfakes: malicious use, the law and staying safe



Deepfakes – Can You Trust What You See Online?

What is a Deepfake?

A deepfake is a photo, video or audio recording that has been altered using Artificial Intelligence (AI) to make it appear that someone has said or done something they never actually said or did.

Modern deepfakes can be incredibly convincing. AI can learn a person's facial expressions, voice and mannerisms from existing videos and then generate entirely new content that looks real.

This creates a major challenge: if we can no longer trust our eyes and ears, how do we know what is real?

Why Are Deepfakes a Problem?

Deepfakes can be used to:

- Spread fake news and misinformation
- Damage someone's reputation
- Create fake celebrity endorsements
- Manipulate public opinion
- Commit fraud and scams
- Harass or embarrass individuals online

As the technology improves, it is becoming increasingly difficult to tell the difference between real and fake content.

Discussion Question

If a video looks completely real, should we automatically believe it?

Deepfakes: malicious use, the law and staying safe



Staying Safe in a World of Deepfakes

Deepfakes Can Cause Real Harm

Deepfakes have been used to create fake political speeches, fraudulent phone calls and even fake explicit images of individuals without their consent.

In some cases, criminals have used AI-generated voices to trick people into transferring large amounts of money, believing they were speaking to a trusted colleague or family member.

Because of these risks, the UK Government has introduced new laws making the creation of sexually explicit deepfakes without consent a criminal offence.

What Can We Do?

The most important skill is critical thinking.

Before believing or sharing content online, ask yourself:

- Where has this come from?
- Is it from a trusted source?
- Can the information be verified elsewhere
- Does anything seem unusual or inconsistent?

If you come across a harmful deepfake:

- ✓ Do not share it
- ✓ Report it to the platform
- ✓ Tell a trusted adult
- ✓ Contact the police if someone has been targeted or harmed

Key Message

Deepfakes are becoming more sophisticated every year. The best defence is not technology – it is learning to question, verify and think critically about what we see online.

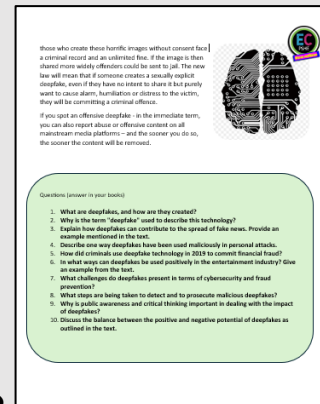
Will it become impossible to tell the difference between real and fake content in the future? If so, what should we do about it?

Belong - Care - Aspire - Succeed

Now answer the questions in your booklets.



1. What are deepfakes, and how are they created?
2. Why is the term "deepfake" used to describe this technology?
3. Explain how deepfakes can contribute to the spread of fake news. Provide an example mentioned in the text.
4. Describe one way deepfakes have been used maliciously in personal attacks.
5. How did criminals use deepfake technology in 2019 to commit financial fraud?
6. In what ways can deepfakes be used positively in the entertainment industry? Give an example from the text.
7. What challenges do deepfakes present in terms of cybersecurity and fraud prevention?
8. What steps are being taken to detect and to prosecute malicious deepfakes?
9. Why is public awareness and critical thinking important in dealing with the impact of deepfakes?
10. Discuss the balance between the positive and negative potential of deepfakes as outlined in the text.



How are deepfakes helping to aid the spread of misinformation?

<https://www.youtube.com/watch?v=wxEpIn8VMWw>

You have been given a bingo card. Whilst listening carefully to the clip, you need to **cross out** all the statements about how deepfakes are contributing to the spread of fake news and misinformation.

There are incorrect answers mixed in, so watch out. The first person to collect all the answers shouts 'BINGO' and if they are correct,



Joe Biden, Keir Starmer and Sadiq Khan have all been victims of deepfakes.	Deepfakes just save time. Politicians are often busy, so the tech allows teams to work on pre-prepared speeches for them.	In Indonesia, long-dead politicians have been brought back to life through deepfakes to endorse current political candidates.
There is huge potential to use deepfakes to undermine trust in politicians and institutions.	Everyone can spot a deepfake, so they're no real threat to true and reliable sources.	Not many countries aside from Western ones even use deepfake technology.
If reliable sources use deepfakes, this can make people trust a technology which is often used maliciously.	Deep fake tech is popular with everyone in the politics industry.	The mayor of New York deepfaked himself, so he could speak to the people of New York in their own languages.

Deepfakes: malicious use and staying safe



Useful helplines and charities:

[Young Minds](#). Child and adolescent mental health charity for teens struggling with any subject.

Call: 0808 802 5544

[Teen Line | Teens Support hotline - Connect, talk, get help!](#) Teen Line's highly trained teen listeners provide support, resources and hope to any teen who is struggling.

[SANE](#). National out-of hours mental health charity offering emotional support, guidance and information. Call: 0300 304 7000 (4.30pm to 10.30pm)

[Mental Health Foundation](#). Information and support for anyone with mental health problems or learning disabilities.

[Rights online \(coe.int\)](#) Your rights online as a young person using social media sites



Need Support? You're Not Alone

If anything in today's lesson has affected you, or you want to talk to someone, there is help available.



Mental Health & Low Mood

YoungMinds – <https://www.youngminds.org.uk>

Kooth – <https://www.kooth.com>

Mind – <https://www.mind.org.uk>

Talk to Someone

Your Tutor or Head of Year – We're here to help.

Wellbeing Team and School Nurse

Report a Concern on Satchel

Safeguarding Team with the Purple lanyards

Healthy Lifestyle

NHS Every Mind Matters – <https://www.nhs.uk/every-mind-matters>

Change4Life – <https://www.nhs.uk/change4life>

Apps That Can Help

Calm – For mindfulness and sleep.

Headspace – Meditation and stress relief.

Clear Fear – Manage anxiety (designed for young people).

MeeTwo – Anonymously talk to other teens, moderated by experts.

Eating Concerns

Beat Eating Disorders –

<https://www.beateatingdisorders.org.uk>

NHS Live Well – Eating Disorders –

<https://www.nhs.uk/mental-health/conditions/eating-disorders/>



Belong - Care - Aspire - Succeed