

GUIDANCE FOR ARTIFICIAL INTELLIGENCE - DEEPFAKES

What are Deepfakes?

Deepfakes are videos, photos, or audio recordings that use artificial intelligence to make it look or sound like someone is saying or doing something that is not real.

Deepfake Techniques

- **Face Swapping:** Replacing a person's face with another in a video or photo.
- **Voice Cloning:** Using AI to mimic someone's voice.
- **Lip Syncing:** Making a person's lips appear to match different audio.

Dangers of Deepfakes

While deepfakes can be used for harmless or even educational purposes, they can also be used unethically in the following ways:

- **Identity theft:** Deepfakes can impersonate individuals in videos or audio recordings, allowing criminals to commit fraud under another's name.
- **Threats/Blackmail:** Deepfakes can create convincing footage or audio that puts the subject in a bad light, which can then be used to threaten or blackmail them.
- **Scams:** Deepfakes can create convincing likenesses of friends, family, or coworkers who claim to be in danger or needing urgent assistance, with the goal of tricking you into sending them money.
- **Cyberbullying:** Deepfakes can be used to create embarrassing or hurtful videos or audio recordings of someone.
- **Propaganda/Fake news:** Deepfakes can be used to create fake videos or audio recordings that trick people and sway opinions on important topics.
- **Election manipulation:** Deepfakes can be used to make candidates look bad or confuse voters by creating fake speeches or interviews.
- **Nonconsensual Intimate Images:** Deepfake photos and videos that remove the subject's clothing without their knowledge or consent.

Content adapted from: <https://www.aiforeducation.io/s/Deepfakes-Classroom-GuideDiscussion.pdf>

How to Identify Deepfakes

Detecting deepfakes involves looking for inconsistencies and anomalies in videos, photos, or audio recordings that may not be immediately apparent. It's important to note that deepfake technology is quickly improving, so it may soon be impossible to tell a fake from something real.

- **Look closely at people's faces in the video.** If their skin looks too smooth, their eyes don't blink normally, or their facial expressions seem off, it might be a deepfake.
- If the voice doesn't sound quite right, like it's **too robotic or the tone doesn't match how the person usually speaks**, it could be a deepfake.
- Pay attention to **how the lips move**. If they're not syncing up with the words perfectly, or the movement looks strange, the video might be manipulated.
- Consider the video's **overall appearance**. Sometimes, the area around the person or behind them might look **blurry or glitchy**. These can be signs the video has been altered.
- If the **shadows on the person's face or in the background don't match up with the lighting** in the rest of the video, it could be a clue of a deepfake.

What Can You Do?

- **Question everything:** AI-generated content has already flooded the internet, so be cautious. Always investigate if a suspicious image, video, or audio clip is real.
- **Look for the clues described above:** glitches and blurs, lip sync issues, etc.
- **Check the source:** be extra cautious with content from unknown websites or accounts.

Remember: We must practice discernment in all areas of life. Pray for wisdom, and keep the above information in mind when using your mobile devices.

"Prove all things; hold fast that which is good." - 1 Thessalonians 5:21

"The simple believeth every word: but the prudent man looketh well to his going." - Proverbs 14:15

