



The Online Safety Agency

## Deepfake Bullying - Students

Deepfake technology utilises AI to manipulate images and videos with such precision that it becomes difficult to distinguish the fake from the real. This capability isn't limited to tech experts; it's now accessible to anyone with a smartphone and internet connection, including students. The result? There is a troubling rise when teachers' and students' images are manipulated into compromising or embarrassing content. The sources for these images are often innocuously public, such as school websites or social media accounts, making almost anyone a potential target.

The implications of deepfake misuse extend far beyond the school gates. When a person's image is manipulated to depict them in situations or saying things that never happened, it doesn't just impact their social standing; it can lead to serious legal consequences. Deepfake bullying blurs the lines between mischief and defamation, prompting a need for legal clarity and protective measures.

Understanding the gravity of this issue is the first step. Educators and students must be equipped with strategies to protect their digital identities. From understanding the basics of deepfake technology and its potential for defamation to adopting practical measures for digital safety, the educational community must be proactive.

### What are deepfakes?

Deepfakes are synthetic media in which a person in an existing image or video is replaced with someone else's likeness using advanced artificial intelligence (AI) and machine learning technologies. The term "deepfake" is a blend of "deep learning" and "fake," reflecting the deep learning algorithms that drive the technology. These algorithms, particularly those based on Generative Adversarial Networks (GANs), can analyse the target person's facial expressions, voice, and mannerisms and generate highly realistic images or videos that mimic the target.

### How Deepfakes are Created

Deepfakes are created using AI software that requires a dataset of images or video clips of the person to be mimicked. The more data available, the more convincing the deepfake can be. The process involves two main steps:

1. **Training** - The AI analyzes the target person's characteristics from the collected dataset. It learns how the person looks from different angles, how they speak, and how their face moves when they express emotions.

2. **Generation** - The AI then applies this learned information to an existing video or image of another person, superimposing the target's likeness to create a seemingly authentic output where the target appears to say or do things they never actually did.

## Uses and Implications

Deepfakes have a range of applications, both positive and negative:

**Entertainment** - In movies and video games, deepfakes can be used to enhance special effects, de-age actors, or bring deceased actors back to virtual life.

**Education** - They can create engaging historical documentaries where historical figures appear to be speaking directly to the audience.

**Misinformation and Cybersecurity** - On the darker side, deepfakes can be used to create false narratives, manipulate public opinion, impersonate individuals for fraudulent purposes, or create non-consensual adult content.

**Political Manipulation** - Deepfakes pose a significant threat in the realm of politics, where they can be used to fabricate speeches or actions of public figures to influence voter perception and sway elections.

## Strategies for Protection

### Mitigating Deepfake Risks

**Avoid Direct Gaze in Photos** - Opt for poses that don't provide a clear, frontal view of your face, making it harder for deepfake technology to capture your likeness accurately.

**Exercise Your Right to Opt-Out** - If you're uncomfortable with your image being posted online by your school or workplace, request not to be included in publicly shared photos.

**Embrace Avatars** - Use avatars as a creative and safe way to represent yourself online, maintaining your privacy while engaging in online spaces.

**Be Mindful of What You Share** - Think carefully before posting images online. Apply stringent privacy settings to control who can see your posts.

**Understand and Utilize Photo Use Agreements** - Be clear on how your images will be used and seek adjustments to agreements that don't align with your comfort level.

**Education and Awareness** - Understanding the nature of deepfakes, recognising the signs of image-based abuse, and knowing when and how to report misuse are crucial skills for both educators and students.

**Legal Perspectives and Support** - In some scenarios, legal action for defamation could be a recourse for victims of deepfake bullying. Understanding the legal framework surrounding defamation, image-based abuse, and the misuse of AI technology is essential. Schools should provide support for educators and students navigating these challenges, offering guidance on reporting misuse and seeking legal advice when necessary.

As we venture further into the digital age, the need for even more comprehensive online safety education becomes increasingly apparent. Protecting our digital identities and ensuring a safe online environment for both educators and students requires a collective effort. By fostering an atmosphere of respect, awareness, and proactive protection, we can navigate the complexities of AI technology and its impact on our lives. Together, we can mitigate the risks and embrace the benefits of the online world.

**[www.safeonsocial.com](http://www.safeonsocial.com) - [www.theonlinesafetyagency.com](http://www.theonlinesafetyagency.com)**

No part of this e-book/cheat sheet or its associated modules may be reproduced or transmitted by any person or entity in any form or by any means, electronic or otherwise including photocopying, recording or scanning or by any information storage without prior permission from Safe on Social Media Pty Ltd other than the licensor who is licensed to use this information in newsletters and in print and has been granted permission from the publisher under an annual license.

The publisher, authors, licensee, licensor and their respective employees or agents will not accept responsibility for injuries or damage, physical or emotional occasioned to any person as a result of a social media use or any other activities described in this e-book/cheat sheet, very attempt has been made to ensure that the information in this e-book/cheat sheet is accurate,

it is the nature of social media to be constantly changing.

Therefore, Safe on Social Media Pty Ltd gives no guarantees and accepts no responsibility to the completeness or accuracy of the contents of this guide.