

Deepfake Literacy

Category: Artificial Intelligence in Education

Deepfake literacy is the capacity to understand, evaluate and respond to synthetic or manipulated media that convincingly represents a real person as saying or doing something they did not. It combines technical awareness with attention to identity, consent, context, circulation and potential harm.

A deepfake cannot be assessed reliably by looking for glitches alone. Compression, editing and improving generation make surface cues uncertain. Verification should therefore investigate provenance, seek the original, compare authoritative coverage and consider who benefits from the claim or its rapid online circulation worldwide ([Roe et al., 2024](#)).

Verification and Response

1. **Pause:** Resist rapid sharing, particularly when content is emotionally charged or reputationally damaging.
2. **Trace:** Find the earliest available upload, fuller recording, date and stated source.
3. **Corroborate:** Check trusted reporting, official accounts and independent evidence outside the post.
4. **Assess harm:** Consider consent, impersonation, harassment, fraud and disproportionate effects on targeted people.
5. **Respond:** Avoid amplification, preserve evidence where appropriate and report through relevant institutional channels.

Examples in Practice

A student receives a video apparently showing a university leader making an inflammatory statement. Rather than reposting it, the student searches for the original event, checks the university's verified channels and consults reliable reporting. The class discusses both authenticity and the harms caused by circulation.

Why It Matters

Deepfakes can produce deception and uncertainty even when viewers remain unconvinced, weakening trust in genuine evidence. Targeted media-literacy education can improve protective judgement, but educators should avoid teaching that every unfamiliar recording is false or that detection tools are infallible ([Hwang et al., 2021](#)).

Further Reading

6. Roe, J., Perkins, M., & Furze, L. (2024). [Deepfakes and higher education: A research agenda and scoping review of synthetic media](#). *Journal of University Teaching & Learning Practice*, 21(10).
7. Hwang, Y., Ryu, J. Y., & Jeong, S.-H. (2021). [Effects of disinformation using deepfake: The protective effect of media literacy education](#). *Cyberpsychology, Behavior, and Social Networking*, 24, 188–193.
8. Vaccari, C., & Chadwick, A. (2020). [Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news](#). *Social Media + Society*, 6(1).

Related Terms: Synthetic Media Literacy | Disinformation | Digital Verification | Online Safety