How Deepfake Technology Harms Society

Manuel Cortez | DTSC 690 |Eastern University

What are Deepfakes?

(What the Heck Is a Deepfake? | Information Security at UVA, U.Va., n.d.)
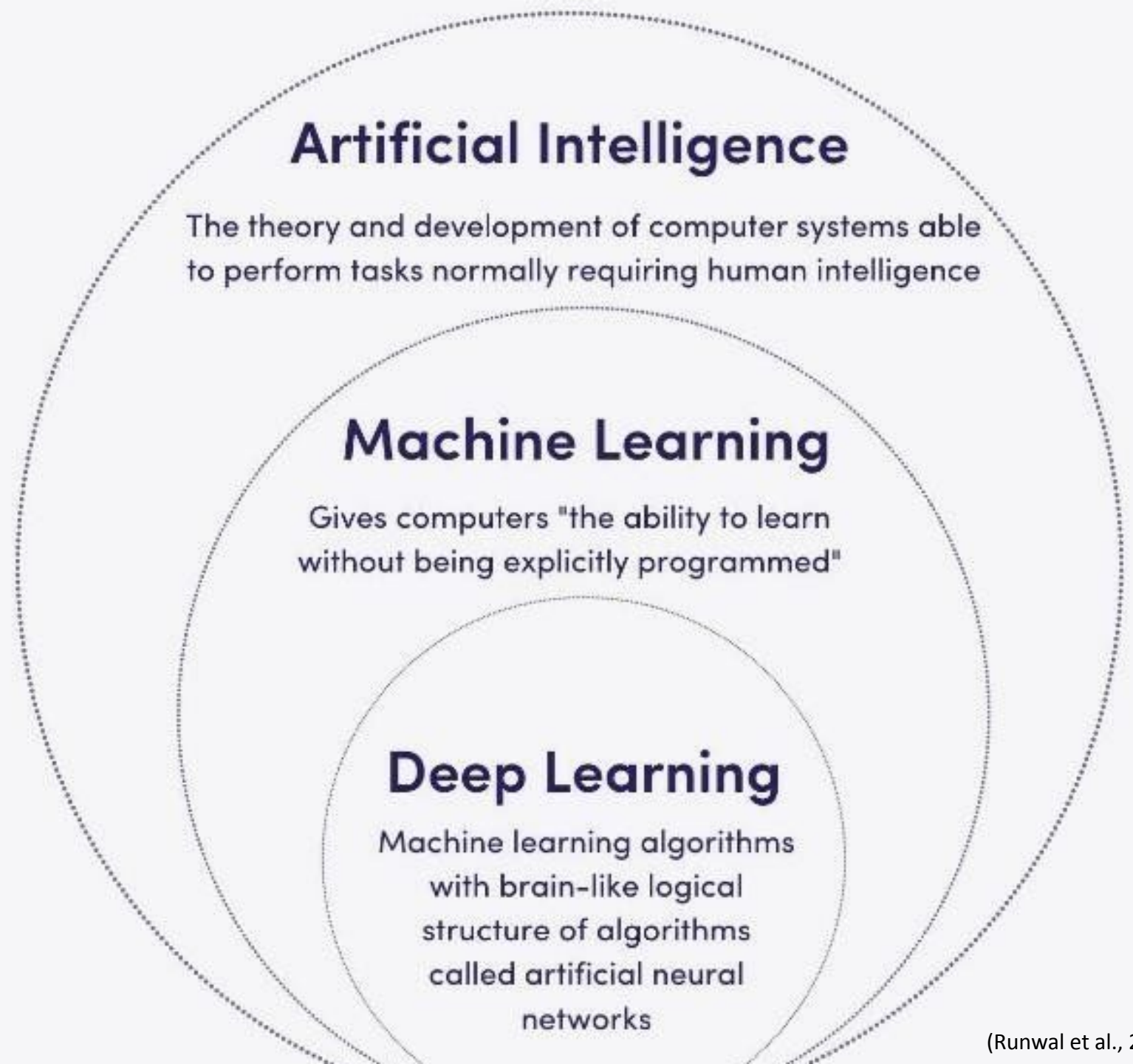
- Synthetic Media
- Artificial Images and Audio

("In Event of Moon Disaster: AI Nixon Merge," 2022)

How do Deepfakes work? *(What Is Deepfake and How to Use It*, n.d.)

- Deep learning
  - Neural Networks
    - Autoencoders
    - GAN
    - First Order Motion Models
- Optimized for Image Processing

Artificial Intelligence

The theory and development of computer systems able to perform tasks normally requiring human intelligence

Machine Learning

Gives computers "the ability to learn without being explicitly programmed"

Deep Learning

Machine learning algorithms with brain-like logical structure of algorithms called artificial neural networks

(Runwal et al., 2023)

Why it matters?

- AI is increasing as a whole
  - Deepfakes as a subcategory are following that rise
  - 900% increase in deepfakes (Letzing, 2021)
- Lack of media literacy
  - 4 out 10 adults (Amanda Marsden, 2022)
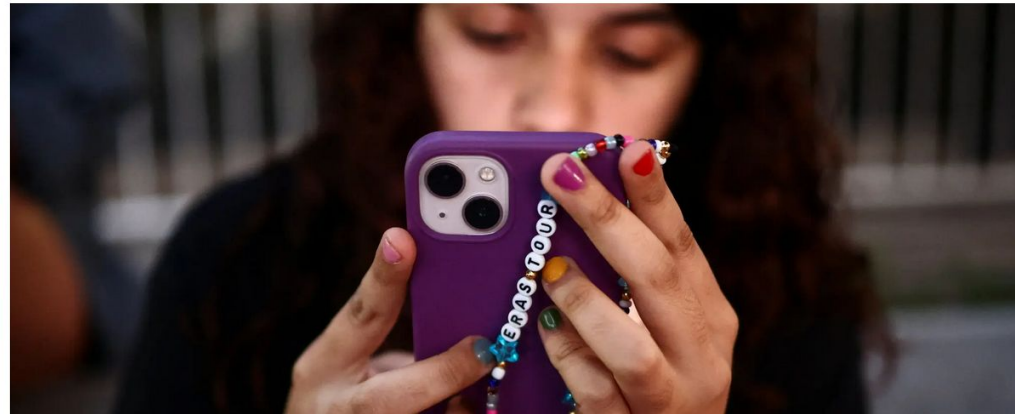
The New York Times

Explicit Deepfake Images of Taylor Swift Elude Safeguards and Swamp Social Media

Fans of the star and lawmakers condemned the images, probably generated by artificial intelligence, after they were shared with millions of social media users.

Share full article   259

(Conger & Yoon, 2024)

Where is the ethical dilemma?

- Erosion of trust
    - Weaken social pillars
    - Propagation of false information
    - Truth value of videos are lowered (Fallis, 2021)
- Unconsensual images/videos
    - Loss of privacy and autonomy

Harm on Women

- A general loss of autonomy (United Nations Population Fund, 2021)
  - Only about 80 percent of countries have laws supporting sexual health and well-being.
  - Twenty countries or territories have "marry-your-rapist" laws, where a man can escape criminal prosecution if he marries the woman or girl he has raped.
  - More than 30 countries restrict women's right to move around outside the home.

Harm on Women

- Tech-facilitated gendered violence (Tech-Facilitated Gender-Based Violence, n.d.)
  - 38 percent of women have had personal experiences of online violence
  - 85 percent of women who spend time online have witnessed digital violence against other women
- Deepfake etymology (What Is Deepfake and How to Use It, n.d.)
- Deepfake and revenge pornography

Arguments for Deepfakes

- Creativity within entertainment
  - Beatles song (Benchetrit, 2023)
- Improving…
  - Education
    - *MyHeritage* (Ofgang, 2022)
  - Translations
    - *Flawless* (Vincent, 2021)
- Simulations/data for medical research (UCL, 2019)

Possible Solutions

- Current and Legal Solutions (Kindred, 2023)|(Kelleher, 2023)...
  - Laws right now fall on a couple of fronts
  - Not well-suited for this technology
  - Revenge Porn
  - Intellectual Copyright, Harassment and Tort claims

Possible Solutions

- Technological Solutions…
  - AI/ML detection tools
  - Blockchain (Cheikosman et al., 2021)
  - But, can it keep up…

Moral Obligations…

- … to use technology appropriately
- … enact laws that protect marginalized individuals
- … create technology that creates safeguards against misuse

References

Amanda Marsden. (2022, September 7). *National Survey Finds Most U.S. Adults Have Not Had Media Literacy Education in High School*. Medialiteracynow.org. https://medialiteracynow.org/nationalsurvey2022/

Benchetrit, J. (2023, November 8). *Here comes the song: A new Beatles record is here, thanks to artificial intelligence*. CBC. https://www.cbc.ca/news/entertainment/new-beatles-song-1.7015861

Cheikosman, E., Hewett, N., & Gabriel, K. (2021, October 12). *Blockchain can help combat threat of deepfakes. Here's how*. World Economic Forum. https://www.weforum.org/agenda/2021/10/how-blockchain-can-help-combat-threat-of-deepfakes/

Conger, K., & Yoon, J. (2024, January 26). Explicit Deepfake Images of Taylor Swift Elude Safeguards and Swamp Social Media. *The New York Times*. https://www.nytimes.com/2024/01/26/arts/music/taylor-swift-ai-fake-images.html

References

Fallis, D. (2020). The Epistemic Threat of Deepfakes. *Philosophy & Technology*, *34*. https://doi.org/10.1007/s13347-020-00419-2

Frequently asked questions: Tech-facilitated gender-based violence. (n.d.). UN Women – Headquarters.

     https://www.unwomen.org/en/what-we-do/ending-violence-against-women/faqs/tech-facilitated-gender-based-violence

In Event of Moon Disaster: AI Nixon merge. (2022). In *YouTube*. https://www.youtube.com/watch?v=SZzn2G1KhFY

Kelleher, K. (2023, August 10). *Revenge Porn, deep fakes, AI, Boston University, Boston University Law School, Rebecca A. Delfino,*

     artificial intelligence» Dome | Blog Archive | Boston University. Sites.bu.edu.

     https://sites.bu.edu/dome/2023/08/10/revenge-porn-and-deep-fake-technology-the-latest-iteration-of-online-abuse/

References

Kindred, M. (2023, August 2). *Deepfakes: The Effect on Women and Potential Protections*. University of Cincinnati Law Review Blog.

      https://uclawreview.org/2023/08/02/deepfakes-the-effect-on-women-and-potential-protections/

Letzing, J. (2021, April 1). *How to tell reality from a deepfake?* World Economic Forum.

      https://www.weforum.org/agenda/2021/04/are-we-at-a-tipping-point-on-the-use-of-deepfakes/

Ofgang, E. (2022, November 21). *How to Teach With Deep Fake Technology*. TechLearningMagazine.

      https://www.techlearning.com/news/how-to-teach-with-deep-fake-technology

Runwal, R., Purve, S. J., & Chandak, M. (2023). Classification of Handwritten Digits on the web using Deep Learning. *International Journal of*

      Next-Generation Computing. https://doi.org/10.47164/ijngc.v14i1.1003

References

UCL. (2019, November 29). *Opinion: How the technology behind deepfakes can benefit all of society*. UCL News.

    https://www.ucl.ac.uk/news/2019/nov/opinion-how-technology-behind-deepfakes-can-benefit-all-society

United Nations Population Fund. (2021, April 14). *Nearly half of all women are denied their bodily autonomy, says new UNFPA report,*

    My Body is My Own. Www.unfpa.org.

    https://www.unfpa.org/press/nearly-half-all-women-are-denied-their-bodily-autonomy-says-new-unfpa-report-my-body-my-own

Vincent, J. (2021, May 18). *Deepfake dubs could help translate film and TV without losing an actor's original performance*. The Verge.

    https://www.theverge.com/2021/5/18/22430340/deepfake-dubs-dubbing-film-tv-flawless-startup

References

What Is Deepfake and How to Use It. (n.d.). DiscoverDataScience.org.

   https://www.discoverdatascience.org/articles/everything-you-need-to-know-about-how-to-use-deepfake/

What the heck is a deepfake? | Information Security at UVA, U.Va. (n.d.). Security.virginia.edu. https://security.virginia.edu/deepfakes