

2021-10-26

Check Your Tech - The Ethics of Deepfakes in a Political Context

Dympna O'Sullivan

Technological University Dublin, dympna.osullivan@tudublin.ie

Damian Gordon

Technological University Dublin, Damian.Gordon@tudublin.ie

Ioannis Stavrakakis

Technological University Dublin, ioannis.stavrakakis@tudublin.ie

See next page for additional authors

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomcon>



Part of the [Computer Sciences Commons](#)

Recommended Citation

O'Sullivan, Dympna; Gordon, Damian; Stavrakakis, Ioannis; and Collins, Michael, "Check Your Tech - The Ethics of Deepfakes in a Political Context" (2021). *Conference papers*. 359.

<https://arrow.tudublin.ie/scschcomcon/359>

This Article is brought to you for free and open access by the School of Computer Sciences at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin.

For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)

Authors

Dympna O'Sullivan, Damian Gordon, Ioannis Stavrakakis, and Michael Collins

Check Your Tech - The Ethics of Deepfakes in a Political Context

Dympna O'Sullivan, Damian Gordon, Ioannis Stavrakakis, Michael Collins

Technological University of Dublin (Ireland)

Dympna.OSullivan@TUDublin.ie, Damian.X.Gordon@TUDublin.ie,

Ioannis.Stavrakakis@TUDublin.ie, Michael.Collins@TUDublin.ie,

ABSTRACT

Deepfakes are image, audio, or video files that are either manipulated or completely synthesised (Westerlund, 2019). Modern deepfakes use artificial intelligence techniques to more easily automate this fabrication process, but such content manipulations have a long history. For example, photo manipulation has existed for almost as long as the history of photography itself, with early examples from the 1860s where there is clear evidence that photographs of both incumbent U.S. president Abraham Lincoln, and future president Ulysses S. Grant, were manipulated and composited (LoC1, 2021; LoC2, 2021). This association between politics and image manipulation has continued throughout the decades, with notable examples including the removal of Soviet Commissar Nikolay Yezhov from photographs with the General Secretary of the Soviet Union, Joseph Stalin, when he fell out of political favour in the 1930s, to Italian Prime Minister Benito Mussolini having a horse handler removed from a photograph to make himself appear more heroic in the 1940s (Fineman, 2012). More recently during the 2004 U.S presidential campaign of John Kerry in 2004, a photograph showing Kerry with Jane Fonda speaking at an anti-war rally from the 1970s was spread on social media sites, and although it was quickly identified as having been faked, it has been cited as one of the reasons for his election loss (Roland-Shearer, 2009; Qureshi and Deriche, 2015).

Deepfake software is becoming widely available, and in spite of the fact that such software can have very serious implications for politics, for example, Dobber, *et al.* (2021) have shown that it is possible to change voters' views on a particular politician using deepfake technology, it is nonetheless largely unregulated. It is, therefore, clear such technology adds a worrying new dimension to the types of "dirty tricks" the political parties can undertake; not only can they create fabricated digital content of their political opponents, but they also have plausible deniability if harmful content of their own candidates is uncovered, as they can claim these are deepfakes. Researchers like Korus and Memon (2019) advocate the development of AI systems to detect altered digital content, and the watermarking of content (either visibly or steganographically).

Some answers to this issue may be found by investigating the works of postmodern philosophers like Derrida, Foucault, and Lyotard, who argue that truth is not an absolute and universal concept, but rather is contingent on context and never complete (Hicks, 2004). A recurring motif in postmodern philosophy, best exemplified by Jean Baudrillard 1981 thesis "Simulacra and Simulation", is the notion of recontextualising or "remixing" aspects of society in self-conscious appropriations or parodies. Their works would suggest that the organisations that create deepfake software should also be the ones who are (legally) required to develop software tools to detect those same deepfakes, and those organisations should carefully reflect on the ethical implications of their software.

This issue is one of grave concern, and is one of a rapidly growing number of computer ethics issues that have been emerging recently, to such an extent that a number of third-level institutes across Europe are collaborating to explore some of these key ethical challenges, and to develop educational content that is both based on pedagogically sound principles, and motivated by international exemplars of best practice to highlight these matters as part of the Erasmus+ Ethics4EU project

(O’Sullivan and Gordon, 2020). One specific development that is being undertaken is the creation of a lesson focusing on programming issues, and concentrating specifically on the ethics of developing software systems that can have a negative impact on people’s lives.

References

Baudrillard, J. (1981) *Simulacra and Simulation*. University of Michigan press.

Dobber T, Metoui N, Trilling D, Helberger N, de Vreese C. (2021) “Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?” *The International Journal of Press/Politics*. 26(1), pp. 69-91. doi:10.1177/1940161220944364

Hicks, S. R. C. (2004) *Explaining Postmodernism: Skepticism and Socialism from Rousseau to Foucault*. Scholargy Publishing, Inc..

Korus, P. and Memon, N. (2019) "Content Authentication for Neural Imaging Pipelines: End-To-End Optimization of Photo Provenance in Complex Distribution Channels," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 8613-8621, doi: 10.1109/CVPR.2019.00882.

Fineman, M. (2012) *Faking it: Manipulated photography before Photoshop*. Metropolitan Museum of Art, 2012.

LoC1, Library of Congress, Item 2003654314, <https://www.loc.gov/pictures/item/2003654314/>

LoC2, Library of Congress, Civil War Glass Negatives and Related Prints: Solving a Civil War Photograph Mystery, <https://www.loc.gov/pictures/collection/cwp/mystery.html>

O’Sullivan, D., Gordon, D. (2020) “Check Your Tech – Considering the Provenance of Data Used to Build Digital Products and Services: Case Studies and an Ethical CheckSheet”, IFIP WG 9.4 European Conference on the Social Implications of Computers in Developing Countries, 10th–11th June 2020, Salford, UK.

Qureshi, M. A., and Deriche, M. (2015). A bibliography of pixel-based blind image forgery detection techniques. *Signal Processing: Image Communication*, 39, 46-74.

Roland Shearer, R. (2009) “Fake Anti-War protest photo of Fonda and Kerry represents classic propaganda fauxtography”, *iMediaEthics*, <http://imediaethics.org/fake-anti-war-protest-photo-of-fonda-and-kerry-represents-classic-propaganda-fauxtography/>

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11).