UNIVERSITÄT
LEIPZIG

DEEPFAKE IMPACT? ANALYSING THE SUSTAINABILITY OF NEWS ORGANISATIONS IN THE ERA OF ARTIFICIAL INTELLIGENCE (AI)

Jönköping, October 28, 2021

Daniel Seibert, Alexander Godulla, Rosanna Planer
Leipzig University

DEEPFAKES

Deepfakes can be defined as an output of “artificial intelligence-based image synthesis technique that involves creating fake but highly realistic video content“ (CDEI, 2019). Using deepfake technology, it is possible to represent a person in any situation or environment imaginable (CDEI, 2019) and use existing recordings of a person's voice to have them say any sentence (Vaccari & Chadwick, 2019).

1. DEEPFAKE DEFINITION

1. DEEPFAKE DEFINITION

- Face replacement
- Face re-enactment
- Face generation
- Speech synthesis (CDEI, 2019)

Trends

- prevalence of deepfakes is rising extremely (Simonite, 2019)
- to date: majority of deepfakes consist of pornographic content of women (Simonite, 2019)

1. DEEPFAKE DEFINITION

Key challenges

- high credibility and audiences' trust in truthfulness of audio-visual content (Fehrensen & Täubner, 2019; Sundar et al., 2021)
- Deepfake creation is getting easier
- the (viral) distribution of Deepfakes through social network sites
- Deepfakes may enter broader social, political, and economic contexts

RQ: WHAT CHALLENGES DO JOURNALISM AND NEWS ORGANISATIONS FACE DUE TO THE CREATION OF DEEPFAKES?

2. METHODOLOGY: LITERATURE REVIEW

2. METHODOLOGY: LITERATURE REVIEW

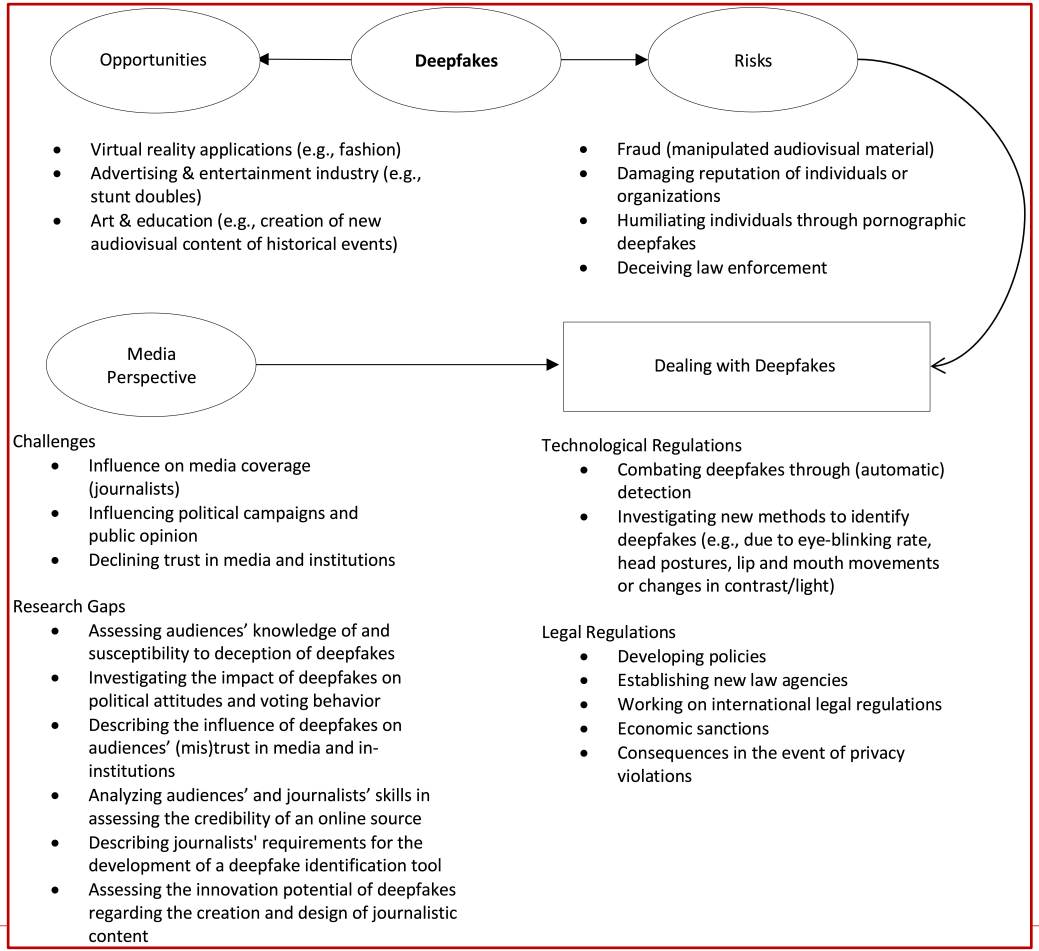
- period: April 27th to May 11th, 2020
- keywords: deepfake / deep fake
- time frame: publications from 2018 to 2020
- journal articles and conference papers
- sample size: 37 articles

Items	Prevalence
Analytical approach	Conceptual: 20 Empirical: 17
Type of deepfake approach investigated	Focus on risks: 31 Focus on opportunities: 6

3. RESULTS

3. RESULTS

Figure 1. Challenge Model of Deepfake Impact

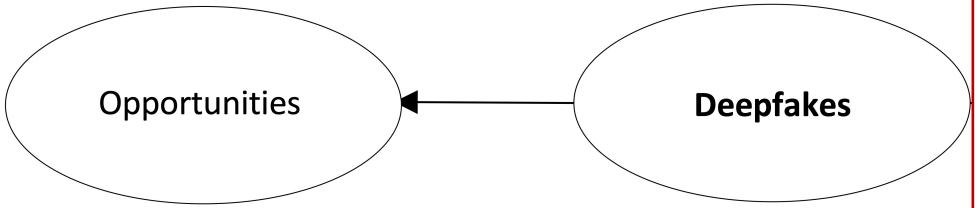


- **Addressing challenges and revealing gaps in research in order to support media organisations**

3. RESULTS

- **Opportunities and risks**

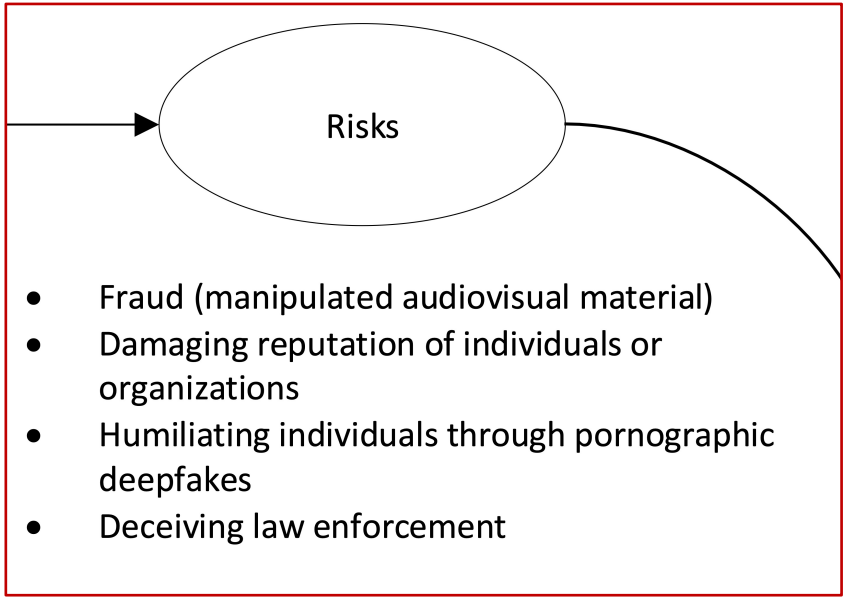
Figure 1. Challenge Model of Deepfake Impact



- Virtual reality applications (e.g., fashion)
- Advertising & entertainment industry (e.g., stunt doubles)
- Art & education (e.g., creation of new audiovisual content of historical events)

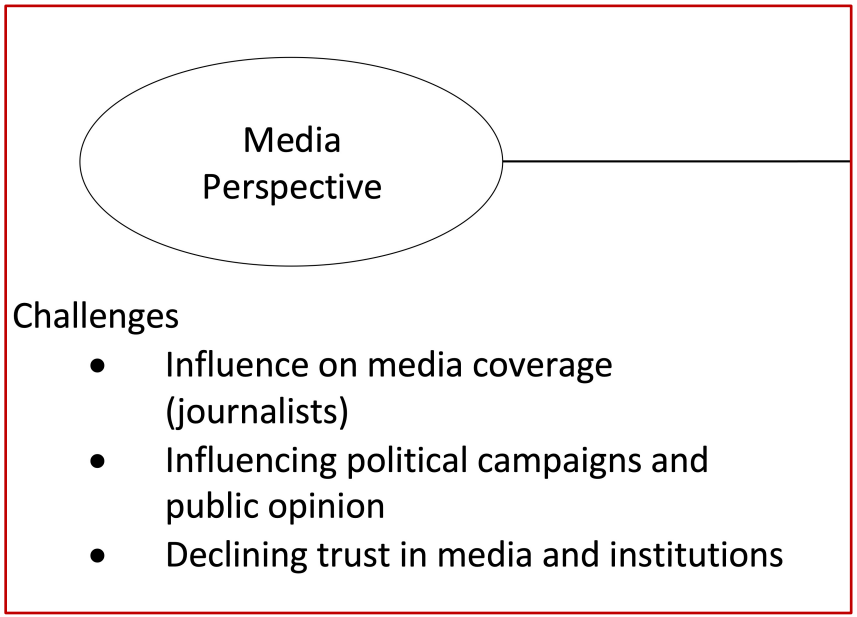
3. RESULTS

- **Opportunities and risks**



3. RESULTS

- **Applying the media perspective**



3. RESULTS

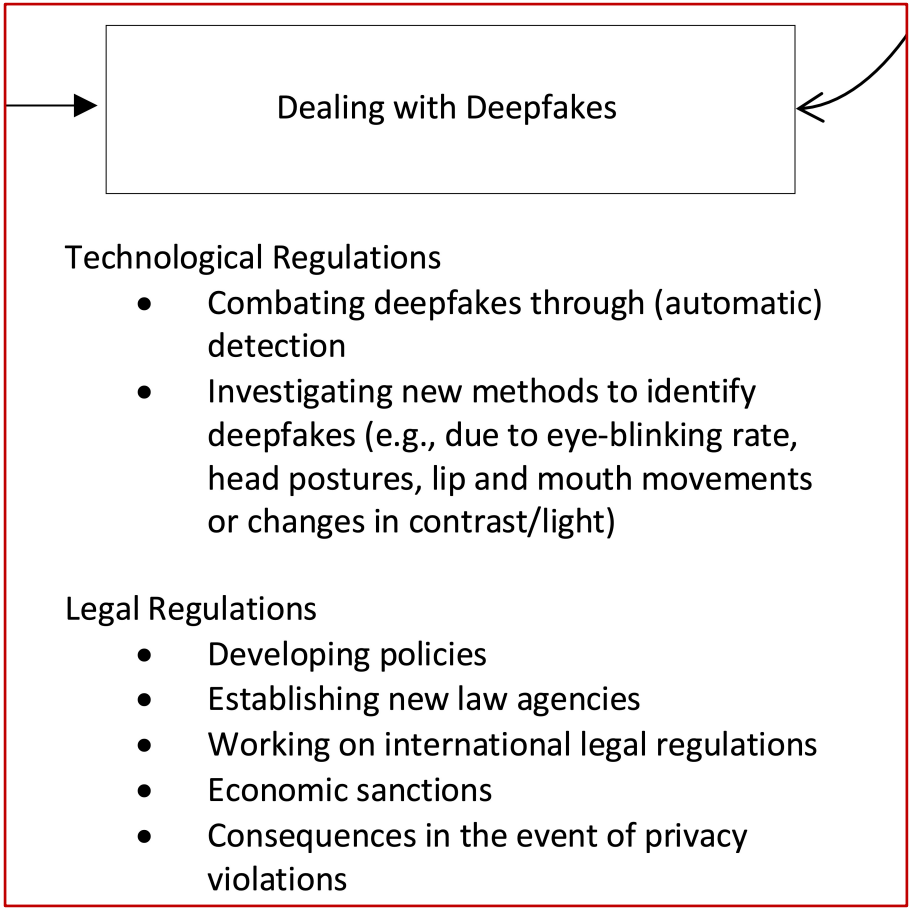
➢ **Applying the media perspective**

Research Gaps

- Assessing audiences' knowledge of and susceptibility to deception of deepfakes
- Investigating the impact of deepfakes on political attitudes and voting behavior
- Describing the influence of deepfakes on audiences' (mis)trust in media and in-institutions
- Analyzing audiences' and journalists' skills in assessing the credibility of an online source
- Describing journalists' requirements for the development of a deepfake identification tool
- Assessing the innovation potential of deepfakes regarding the creation and design of journalistic content

3. RESULTS

- **Dealing with deepfakes**



4. OUTLOOK

4. OUTLOOK

- Further developing technological and legal regulations
- Expanding social sciences research
- Applying an interdisciplinary approach considering media law, media economics and media technology
- Exploring potentials of deepfakes in the context of education or science communication

REFERENCES

REFERENCES

CDEI - The Centre for Data Ethics and Innovation (2019). Deepfakes and Audio-visual Disinformation. *CDEI Snapshot Series.* Retrieved from: https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai

Fehrensen, M. & Täubner, M. (2019*).* Warum wir nicht glauben sollten, was wir sehen. Gefälscht wurde schon immer. Sogenannte Deepfakes aber weisen in eine neue Dimension der Manipulation [Why we should not believe what we see. Counterfeiting has always been used. But so-called deepfakes point to a new dimension of manipulation]. *Brand eins*. Retrieved from: https://www.brandeins.de/magazine/brand-eins-wirtschaftsmagazin/2019/gefuehle/social-media-warum-wir-nicht-glauben-sollten-was-wir-sehen?utm_source=zeit&utm_medium=parkett

Godulla, A., Hoffmann, C. P., & Seibert, D. (2021). Dealing with deepfakes–an interdisciplinary examination of the state of research and implications for communication studies. SCM Studies in Communication and Media, 10*(1),* 72-96. DOI: 10.5771/2192-4007-2021-1-72

Simonite, T. (2019). Most Deepfakes Are Porn, and They're Multiplying Fast. Researchers worry that doctored videos may disrupt the 2020 election, but a new report finds that 96 percent of deepfakes are pornographic. WIRED. Retrieved from: https://www.wired.com/story/most-deepfakes-porn-multiplying-fast/

Sundar, S. S., Molina, M. D., & Cho, E. (2021). Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps?. Journal of Computer-Mediated Communication 00, 1-19. DOI: 10.1093/jcmc/zmab010

Turton, W. & Martin, A. (2020, January 7). How Deepfakes Make Disinformation More Real Than Ever. Bloomberg. Retrieved from: https://www.bloomberg.com/news/articles/2020-01-06/how-deepfakes-make-disinformation-more-real-than-ever-quicktake

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)*, 1-13. DOI: 10.1177/2056305120903408

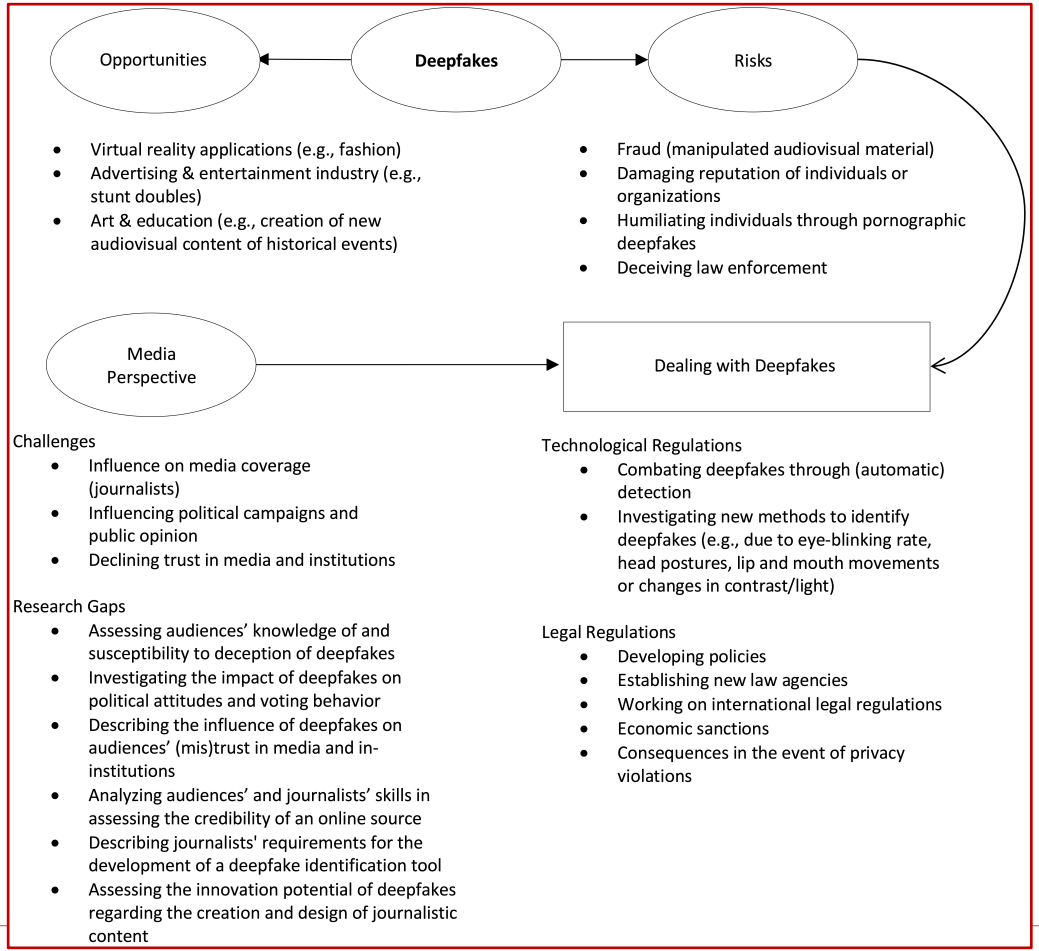
UNIVERSITÄT
LEIPZIG

THANK YOU VERY MUCH FOR YOUR ATTENTION!

Daniel Seibert
Leipzig University, Institute for Communication and Media Studies

daniel.seibert@uni-leipzig.de
www.uni-leipzig.de

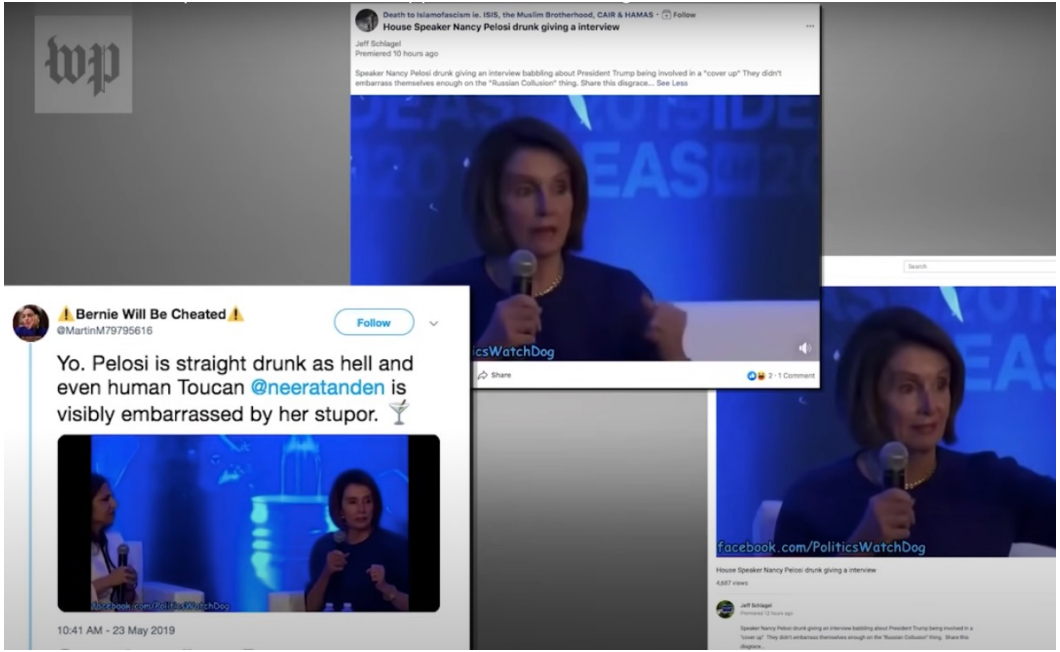
Figure 1. Challenge Model of Deepfake Impact



DEEPFAKE EXAMPLE

Nancy Pelosi deepfake
(Turton & Martin, 2020)

- artificially slowed-down video spread on Facebook
- challenge: deepfake identification and platform regulations



LITERATURE REVIEW

TABLE 3. SUMMARY AND METHODOLOGICAL PREFERENCES OF THE EXAMINED EMPIRICAL STUDIES (N=17).

Item	Prevalence
Examined groups of people	Experiments: 1 study: n= 30; 1 study: n=2.005 No groups of people were examined: 15
Main method applied	Content analysis: 1 Experiment: 16
Geographical place(s) of data collection	Canada: 2 China: 1 France: 1 Germany: 1 Jordan: 1 Korea: 1 the Netherlands: 1 Norway: 1 Serbia: 1 United Arabian Emirates: 1 United Kingdom: 1 United States: 5
Dataset of deepfakes examined	Mean: n=240; 1 study: n=2; 1 study: n=3; 1 study: n=4; 1 study: n=8; 1 study: n=48; 2 studies: n=49; 1 study: n=100; 1 study: n=175; 1 study: n=600; 1 study: n=620;
Different type of dataset	1 study: n=640; 1 study: n=820; 3 studies: not applicable 1 study: n=84 (news articles)