



UNIVERSITÄT
LEIPZIG

The **Deepfake** Project

Twelve key findings on the impact and perception of deepfake technology from five years of research

Prof. Dr. Alexander Godulla
Prof. Dr. Christian P. Hoffmann
Daniel Seibert, M.A.

deepfake-project.com

Study 1: Systematic literature analysis and status quo of deepfake research

1.

Research on deepfakes is primarily characterized by studies from the field of computer science, which deal with the development of deepfake detection tools, and studies from the field of law, which are dedicated to the development of legal regulations for dealing with harmful uses of deepfakes.

2.

As awareness of the technology grows, more and more studies in the humanities and social sciences are focusing on the phenomenon and investigating, for example, the impacts of deepfakes on audiences and their ability to recognize them.

Study 2: Quantitative audience survey on deepfakes in Germany

3.

At the time the study was conducted (5-11 October 2022; n = 1421), most respondents stated that they had not yet encountered a deepfake.

75.5% of respondents do not believe they have encountered a deepfake in the past 12 months; 24.5% of respondents state that they have already encountered a deepfake in that time.

4.

The majority of respondents have little or no knowledge of deepfakes and are not aware of the possibilities of the technology and its potential areas of application.

Variable		Percentage (%)
Deepfake Knowledge	No prior knowledge	67.1
	Low	11.3
	Medium	17.4
	High	4.2

Note: The level of knowledge about deepfakes was measured using 13 knowledge questions. No prior knowledge of deepfakes includes respondents who stated in the questionnaire that they did not know what a deepfake was as well as respondents who did not answer any of the quiz questions correctly. Answering one to four questions correctly was considered a low level of knowledge, five to seven correct answers a medium level of knowledge and eight or more correct answers a high level of knowledge.

5.

Characteristics of *digital inequalities* in the knowledge and perception of deepfake technology emerge.

- (a) Our study points to socio-demographic differences: younger respondents, those with a higher educational attainment and males are more likely to understand the technology, its potentials and applications.
- (b) The results of our study also show that an individual's digital skills have a strong connection to knowledge of deepfake technology.
- (c) The level of knowledge about deepfakes is, in turn, significantly related to the perception of risks and opportunities of the technology. A nuanced understanding of deepfake technology could therefore increase the likelihood of deriving personal benefits from its use, while understanding its dangers could strengthen individuals' ability to mitigate some of the risks of the technology.

6.

A majority of respondents perceive the emergence of deepfake technology as a major threat.

Variable		Percentage (%)
Perceived risks	Low	5.8
	Medium	24.3
	High	69.9
Perceived potentials	Low	26.3
	Medium	1.9
	High	31.7

Note: Respondents' attitudes towards deepfakes were measured using 21 statements (twelve risks, nine opportunities) on areas of application of the technology. Respondents were asked to rate their agreement with these statements (do not agree at all to fully agree). A low perception of risks includes respondents who did not agree with any of the risk statements, a medium perception of risks includes those who agreed to one to six of these statements and a high perception of risks includes those who agreed to seven to twelve of these statements. A low perception of potentials includes respondents who did not agree with any of the opportunity statements, a medium perception of potentials includes those who agreed to one to four of these statements and a high perception of potentials includes those who agreed to five to nine of these statements.

I have concerns...	Completely or rather disagree	Neither nor	Completely or rather agree
about falling for a deepfake.	6.6	26.1	67.3
about becoming a victim of a harmful deepfake (e.g. pornographic deepfakes, scams on the phone).	18	31.3	50.7
about deepfakes influencing public opinion (e.g. being used to influence election campaigns).	3.6	16.9	79.5
that soon I won't be able to trust any video content.	6.7	25.8	67.5
that the line between reality and fiction is blurring.	3.9	18.6	77.5
that deepfakes will harm our democracy (e.g. by spreading misinformation/fake news).	4.2	18.7	77.1
that anyone is capable of creating deepfakes.	12.9	29.4	57.7
that manipulation of personal data (e.g. audio or images of me) will become easier.	6.1	20.6	70.4
that I will have to protect my personal data even more in the future (e.g. as it can be manipulated for harmful purposes).	5.2	20.7	74.1
about falsely spreading deepfakes.	25.8	31.9	42.3
that detecting deepfakes will become increasingly difficult.	2.7	10.7	86.6
about the increasing prevalence of deepfakes.	3	16.7	80.3

7.

The benefits and potentials of deepfakes are only recognized to a limited extent by respondents.

In my opinion...	Completely or rather disagree	Neither nor	Completely or rather agree
deepfakes can also be used for beneficial purposes.	18.1	37.1	44.8
deepfakes provide me with benefits (e.g. by creating funny videos of me and my friends).	27.2	39.9	32.9
there are advantages in the fact that any person is able to create deepfakes.	44.5	38.0	17.5
deepfakes provide benefits to the film industry (e.g. making action scenes more realistic or displaying people who already died in movies).	8.5	28.5	63
deepfakes also benefit the creation of pornographic content with the consent of those involved (e.g. if users voluntarily upload their photos to an app and share them to create pornographic content).	30.9	41	28.1
deepfakes provide benefits to the art industry (e.g. creating new artwork in the style of famous artists).	19	39.5	41.5
deepfakes offer advantages for the fashion industry (e.g. virtual try-on clothes).	12.8	35.6	51.6
deepfakes provide benefits to the video game industry (e.g. personalized avatars in games).	12	32.6	55.4
deepfakes foster creativity in media content production.	20.5	41.5	38

8.

Most respondents do not feel equipped to deal with deepfakes.

- (a) Most respondents rate their own ability to recognize deepfakes and distinguish them from real videos as very low.
- (b) The majority of respondents are worried about falling for a deepfake.
- (c) Only one third of respondents are confident that they can use the internet and fact-checking sites to identify a deepfake.
- (d) Respondents who frequently use the internet and social media and have a high interest in new technologies rate their ability to identify deepfakes higher than those who are less engaged in the digital world and less interested in new technologies.

Deepfake detection self-efficacy		Percentage (%)
I am confident that I am able to recognize a deepfake by myself.	Completely or rather disagree	62.9
	Neither nor	27.2
	Completely or rather agree	9.9
I am confident that I can distinguish between a real video and a deepfake video.	Completely or rather disagree	60.8
	Neither nor	27.0
	Completely or rather agree	12.2
I am confident that I can recognize a deepfake video by searching the internet for information (e.g., fact-checking sites).	Completely or rather disagree	35.3
	Neither nor	32.7
	Completely or rather agree	32.0
I am confident that I would not fall for the content of a deepfake video.	Completely or rather disagree	53.8
	Neither nor	29.2
	Completely or rather agree	17.0

9.

Respondents' perception of being exposed to deepfakes online reinforces a sense of media cynicism among them (an exaggerated form of media criticism with conspiratorial tendencies).

Respondents who rate their ability to detect deepfakes higher are more likely to develop a sense of media cynicism when they encounter deepfakes. It is possible that those who believe they understand the new deepfake technology better are either more aware of its potential (negative) impacts in the digital world or are more concerned (e.g. about misuse of the technology by powerful – political or media – institutions).

Study 3: Qualitative expert interviews in the fields of journalism and technology in Germany

10.

The perception of deepfake technology among experts in the fields of journalism and technology (n = 20) is predominantly characterized by risks for journalistic practice (survey period 09/2022-05/2023).

Challenges in dealing with the technology relate to more intensive fact-checking requirements in the future and are characterized by concerns about the compatibility of the technology with basic journalistic principles (accuracy, truthfulness, transparency).

11.

Following the experts, the potential of deepfake technology for journalistic practice primarily relates to personalizing content and simplified forms of creating and editing journalistic content.

Deepfakes are predominantly perceived as a supporting technology that can speed up every day work processes such as the editing and visualization of content. The technology is perceived to be less relevant regarding the development of new, complex journalistic formats and products. The potential of the deepfake technology is argued to be relevant in the automation of personalized, location- and interest-related content in the future (e.g. traffic, weather, sport).

12.

To date, media organizations and journalistic practice are lacking guidelines and strategic approaches for dealing with the risks and potentials of deepfake technology.

- (a) Knowledge about deepfakes and competences in dealing with artificial intelligence-based technologies need to be promoted in order to establish a broader understanding of the advantages and disadvantages of deepfake technology.
- (b) Guidelines for dealing with deepfakes should create an understanding of the boundaries of reality and fiction in the representation of (especially visual) content generated using deepfake technology. This allows to determine which synthetic content and forms of presentation are suitable for journalistic practice, considering basic journalistic principles.
- (c) In terms of fact-checking, new procedures should be developed, and responsibilities clarified to facilitate the identification of potentially harmful deepfakes.