Information, Communication & Society

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rics20

Deepfakes in the context of AI inequalities: analysing disparities in knowledge and attitudes

Daniel Bendahan Bitton, Christian Pieter Hoffmann & Alexander Godulla

INFORMATION, COMMUNICATION & SOCIETY
https://doi.org/10.1080/1369118X.2024.2420037

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Deepfakes in the context of AI inequalities: analysing disparities in knowledge and attitudes

Daniel Bendahan Bitton, Christian Pieter Hoffmann and Alexander Godulla

Institute for Communication and Media Studies, Leipzig University, Leipzig, Germany

ABSTRACT
This study contributes to research on digital inequalities in the context of artificial intelligence by examining user perceptions of deepfake technology. We focus on the stratification of deepfake knowledge and attitudes towards deepfakes as critical elements of technology access. Based on a survey of 1,421 German internet users, we analyse the role of sociodemographic variables, digital skills, and personal innovativeness in predicting deepfake knowledge. We then examine the role of deepfake knowledge in users' assessments of risks and potentials associated with the technology. Our results point to a generally low level of knowledge and a strong focus on risks in internet users' perceptions of deepfakes. We find that age, gender and educational attainment predict knowledge about deepfakes. Digital skills, personal innovativeness, and social media use also positively relate to deepfake knowledge. This knowledge, in turn, is shown to play a role in users' positive attitude towards the technology. While age plays only a minor role, female gender strongly relates to low knowledge and negative attitudes towards deepfakes. We thus find evidence of a sizeable gender divide in user access to the novel deepfake technology.

ARTICLE HISTORY
Received 8 April 2024
Accepted 12 September 2024

KEYWORDS
Deepfakes; digital divide; AI divide; gender; digital skills; attitudes

Introduction

Research on the 'digital divide' builds on the assumption that socioeconomic inequalities are manifested in unequal access to and participation in the digital world (Hargittai, 2002). Studies in this field examine the interplay between personal categories, such as age, sex, or health, positional categories, such as education or income, and technology access, as well as outcomes of technology use (Helsper, 2017a; Van Deursen & Van Dijk, 2019; Van Dijk, 2005). Exclusion from the use of digital technologies is problematized because it is associated with fewer socioeconomic opportunities (Van Deursen et al., 2017; Van Deursen & Helsper, 2018). As technologies evolve, new manifestations of digital inequalities emerge (Lutz, 2019; Vartanova & Gladkova, 2019) – such as the 'AI divide' (Carter et al., 2020). Those with little affinity for new digital technologies or more sensitive to their disadvantages are in danger of missing out on their benefits, while others are

CONTACT Daniel Bendahan Bitton daniel.bendahan@uni-leipzig.de Nikolaistr. 27-29, 04109 Leipzig, Germany

eager to embrace new technologies – and may thus be more exposed to their specific risks (Kurpicz-Briki et al., 2023).

Since the emergence of pornographic deepfakes in 2017 on Reddit, deepfake technology has attracted tremendous interest in research and practice (Godulla et al., 2021). Using existing audio-visual recordings as training material, deepfakes can depict individuals in scenarios that never actually occurred, creating the illusion that they took actions that never happened or expressed thoughts they never articulated (Vaccari & Chadwick, 2020).

Consequently, deepfakes pose new challenges to internet users – due to the high quality of their manipulations and users' willingness to believe audio-visual content. Numerous risks associated with deepfakes have been discussed: the technology could be used for identity theft and phishing tactics (Citron & Chesney, 2019), targeted advertising and surveillance (Cardenuto et al., 2023), reputational attacks (Godulla et al., 2021), as well as for political influence campaigns (Dobber et al., 2020; Hameleers et al., 2023). On the other hand, deepfakes could have constructive uses, for example in educational programmes, in the creative industries and for entertainment (Kolagati et al., 2022). Access to deepfake technology may thus be associated with advantages to some, while exclusion from its use could disadvantage other individuals.

This study examines deepfakes in the context of the AI divide. To date, there is still little research on AI-related inequalities (e.g., Cardenuto et al., 2023; Carter et al., 2020; Kurpicz-Briki et al., 2023). We address this gap and apply a digital inequalities perspective to explore predictors of internet users' *knowledge* of and *attitudes* towards the AI-based deepfake technology. In his sequential model of digital technology access, Van Dijk (2005) proposes 'motivational access' as the most basic step of technology access. If users do not expect benefits from the use of a technology, they will not seek to master the required material and skills access. Based on Rogers (1983) 'Innovation-Decision Process', Eichhorn et al. (2022) argue that motivational access presupposes knowledge. We therefore examine both users' deepfake knowledge and their perceptions of the technology's risks and opportunities. Based on a survey of German internet users ($n = 1421$), we focus on the role of sociodemographics, digital skills and personal innovativeness. Hence, this study addresses the following research questions:

RQ1: How are internet users' sociodemographics, digital skills and personal innovativeness related to their knowledge of deepfake technology?

RQ2: What role does deepfake knowledge play in internet users' attitudes towards deepfake technology?

We find that most respondents have little knowledge of deepfakes. Youth, high educational attainment, male gender, digital skills, and frequent usage of social media predict knowledge of deepfake technology. Furthermore, deepfake knowledge is shown to play a small role in user attitudes towards deepfakes, mostly the perception of the technology's opportunities. We also find that age plays only a minor role in user attitudes towards deepfakes, but attitudes are heavily gendered.

We argue that deepfakes should be examined in the context of the AI divide as their utilization necessitates an understanding of this emerging technology as well as its specific opportunities and risks (e.g., Kitsara, 2022; Yu, 2020). Additionally, we propose that deepfakes could affect digital inequalities because malicious use of the technology could harm vulnerable users while those in a position of relative privilege might employ the technology to achieve beneficial outcomes (Kitsara, 2022).

Literature review

Digital inequalities: the roles of attitudes and knowledge

Originally, the term 'digital divide' focused on inequalities in participation in the digital society due to unequal access to the internet, network infrastructures, and related information and communication technologies (ICT) (Van Dijk, 2013), with interest in the issue increasing in the early 2000s. It has since become understood that technology access comprises a complex, multi-levelled phenomenon, encompassing not just material access to technology, but also motivational or skills access (Van Deursen et al., 2017; Van Dijk, 2005). Technology access can be affected by various personal categories, such as age, gender, or health, as well as positional categories, such as education or income (Van Dijk, 2005) that may intersect in affecting digital inclusion (Tsatsou, 2022). Numerous studies in the field examine the potential impact of socioeconomic predictors on inequalities in technology usage (Cotter & Reisdorf, 2020; Hargittai & Walejko, 2008; Lutz et al., 2014; Van Deursen et al., 2014; 2017). However, Helsper (2017a) points out that these personal categories need to be understood as relative and should be examined in their specific social and temporal contexts. In the present case, we focus on the rise of a novel technology that poses distinct challenges to various user groups.

A key component in the dynamics of digital inequality is users' motivation to employ a technology (Van Deursen et al., 2017; Van Dijk, 2005; 2006). Not all individuals are equally motivated to either explore a new technology or employ an established one (Reisdorf et al., 2019). Helsper (2017b) argues that perceptions of technology need to be taken into consideration when examining digital inequalities. While it is impossible to know for sure what drives users to embrace or avoid new technology, several studies analyze individual factors bolstering use motivations. For example, user concerns or lack of trust are argued to be crucial factors in adopting and using emerging AI technologies (Bozic, 2023). As Wang and Liu (2022) note, the concepts of *motivation* and *attitude* are sometimes used interchangeably in the digital inequalities literature. Reisdorf and Groselj (2017) operationalize motivational access by asking users about their opinions regarding ICT. Other studies similarly assess user attitudes to examine the role of motivation in technology access (Van Deursen & Helsper, 2015b; Van Deursen & Van Dijk, 2015).

Motivation to use technology also relates to the users' interest and their socialization and attitudes towards 'what is appropriate for certain people in certain groups' (Van Deursen et al., 2017, p. 469). Hence, cultural resources are argued to be indicators of Internet and social-media use. For example, users might engage more with digital technology if they find that it enhances efficiency or perceive it as an enjoyable. In contrast, those who find digital technology to be a source of danger and feel uncomfortable using it might feel that it 'is out of their control, potentially, controlled by others' (Reisdorf et al., 2019, p. 87). While users' attitudes towards a technology can be based on actual experiences of its benefits and/or harms (Hargittai & Hinnant, 2008; Van Dijk, 2013), they may also be influenced by public or media framing of the technology. In many cases, users' attitudes are derived from observing public or expert discourses on the opportunities and risks associated with a technology (Morrow, 2022).

Based on Rogers (1983) 'Innovation-Decision Process', Eichhorn et al. (2022) argue that motivational access presupposes knowledge. The less users know about a technology, the more tenuous their attitudes towards it might be. In other words, any additional knowledge gained about the technology could potentially alter user attitudes and behaviour. For example, Cotter and Reisdorf (2020) find that knowledge on algorithms shapes users' expectations and behaviour when consuming information online (e.g., critical reflecting on representations and the truthfulness of information online). Similarly, without awareness of the technology, users are unlikely to develop the necessary motivation to use it. Motivation presupposes at least a basic level of familiarity (cf. Eichhorn et al., 2022). Conversely, the more knowledge of a technology users possess, the more their attitudes towards it will be substantive (whether affirmative or dismissive), and the more able they will be to develop a motivation to use or avoid it.

The present study will examine the role of knowledge and attitudes in the unequal access to deepfakes as an emerging new AI technology. This study thus contributes to the emergent field of 'AI inequality' (e.g., Cardenuto et al., 2023; Carter et al., 2020; Kurpicz-Briki et al., 2023). Research on the emerging AI divide is still rare and can be described, for example, as 'inequality that concerns accessing and using AI-enabled technology and that technology's impacts' (Carter et al., 2020, p. 261). While some studies discuss the extent to which the use of AI technology could weaken the digital divide, most research focuses on the emergence of new inequalities (e.g., Bozic, 2023; Kitsara, 2022; Lainjo, 2023; Yu, 2020).

On the one hand, AI technology might help the underprivileged overcome inequalities – for example, through improved learning processes and skills development (Lainjo, 2023) or the establishment of new business and economic growth (Bozic, 2023; Kitsara, 2022). The attainment of personal and economic benefits through AI technology use, however, depends on a person's awareness and digital literacy (Kitsara, 2022; Yu, 2020). Hence, socioeconomic benefits derived from AI might be limited to those who are motivated and possess the necessary knowledge to handle the respective technology (Bozic, 2023; Lainjo, 2023).

On the other hand, 'the digital divide [might increase] socioeconomic inequities by widening economic opportunity gaps' (Lainjo, 2023, p. 4971). Due to the increasing use of AI, existing inequalities could be reinforced. For example, training data might ingrain biases against disadvantaged groups, thus perpetuating those biases in the content generated using AI algorithms (Gao et al., 2023; Simon & Isaza-Ibarra, 2023). Furthermore, the use of AI for personalization and automatization might impact the way content is created and disseminated, potentially facilitating the consumption of congenial content and contributing to so-called 'echo chambers' (Arguedas & Simon, 2023; Jungherr & Schroeder, 2023; Trattner et al., 2022). These dynamics may reinforce pre-existing socioeconomic inequalities by boosting the views and preferences of those in a position of power and privilege, in particular (cf., D'Ignazio & Klein, 2020; Marwick, 2022). Such inequalities may render emerging AI technologies less attractive or useful to those at the margins. Therefore, transparency in the use of AI-generated or – edited content and data protection are of essential importance for users, especially when the technology is applied in strategic communication and targeting (Gao et al., 2023; Jungherr & Schroeder, 2023).

Deepfakes: familiarity, opportunities and risks

Deepfakes constitute a type of synthetic media that make use of AI-based deep-learning technology and allow users to alter existing recordings of individuals and objects and to generate new ones (Caldera, 2019). The creation of deepfakes is not limited to photo and video manipulations but also encompasses edits to or the generation of audio files. As a result, the technology enables users to visualize incidents that did not actually take place, depict individuals in any situation imaginable and have them say any sentence (Vaccari & Chadwick, 2020).

To date, the majority of deepfakes on the internet are of a pornographic nature, most of them depicting (famous) women (Simonite, 2019). While pornographic deepfakes might be seen as an opportunity to fulfil sexual fantasies (Öhman, 2020), their use without the consent of the individual depicted constitutes an ethical and legal challenge. If women are objectified in pornographic deepfakes and placed in a position of non-consensual (sexual) exploitation, the technology could further reinforce gender-specific differences in visual information (Wagner & Blewer, 2019). In this context, Paris (2021) describes the emergence of 'image-based sexual abuse', as 'social media and face swapping rely on the consent that is barely granted' (p. 7). Paris (2021) analyzed 200 examples of manipulated images and videos, finding that harmful deepfakes and cheap fakes are used to harass, silence, and coerce especially women, LGBTQ individuals and people of colour. This malicious use of the technology can therefore contribute to inequality: Vulnerable individuals may experience harm from the use of the technology, while powerful entities (e.g., entertainment and technology companies, wealthy individuals, political parties) might benefit from its use (Paris, 2021).

Recently, deepfake technology is increasingly expanding into other societal and political spheres beyond pornography. Technological advancements and increasing ease of access facilitate the spread of deepfakes, raising the likelihood that users encounter them online, or actually attempt to generate deepfakes themselves (Seibert, 2023). Previous studies found that most internet users still lack familiarity with the technology or the skills necessary to recognize a deepfake (Bray et al., 2023; Lewis et al., 2022). While more and more studies in the social sciences examine audiences' skills in detecting deepfakes and the technology's impact on, for example, political opinions (e.g., Dobber et al., 2020), thus far there are hardly any studies exploring internet users' attitudes towards the technology.

Initial studies indicate that users' perceptions of deepfake technology are characterized by uncertainty – which, once exposed to and made aware of deepfakes, extends to mistrust towards all (even real) audiovisual content online (Vaccari & Chadwick, 2020). Internet users describe the technology's advancements as scary (Lee et al., 2021) and consider its use for the creation of pornographic content especially harmful (Kugler & Pace, 2021). Users often associate deepfakes with misinformation and fear the technology's effect on public discourse and democratic processes (Citron & Chesney, 2019). Furthermore, the technology raises privacy concerns as users might fall victim to a harmful deepfake if their data are used for identity theft, financial fraud, targeted advertising, or surveillance (Cardenuto et al., 2023; Citron & Chesney, 2019; Kugler & Pace, 2021).

While public attention appears to focus more on dangers associated with deepfake technology, some potential benefits can also be identified. For example, deepfake technology enables its users to experiment with creating synthetic audiovisual content of themselves for personal creative and entertainment purposes. Since the technology allows content to be delivered in a creative and engaging manner, it is also suitable for industry sectors such as film, art, fashion, and video games (e.g., Kolagati et al., 2022). Finally, in contrast to the non-consensual use of an individual's face to create pornographic deepfake content, the technology might also be integrated in a consensus-based platform for creating pornographic content using the faces and bodies of users who authorized the use of their data for this purpose (Raffaghello et al., 2019). The present study will examine both users' knowledge of deepfakes and their attitudes towards the technology based on an assessment of its opportunities and risks.

Research model

We argue that deepfakes can and should be examined in the context of AI inequalities, as sociodemographic factors might affect users' knowledge of this emerging AI technology. Knowledge, in turn, may shape user attitudes and perceptions of the technology's risks and potentials. As argued above, knowledge and attitudes (or motivation) are critical elements in user access to a technology such as deepfakes (Cotter & Reisdorf, 2020; Van Dijk, 2005; 2006).

Previous studies of digital inequalities have shown that men and younger individuals, and those with higher educational attainment and more financial resources are more likely to access technology (e.g., Van Deursen et al., 2014, 2017; Van Dijk, 2005, 2013; Vartanova & Gladkova, 2019). We propose that the same will apply to knowledge of deepfakes, as males and young individuals are more likely to explore the novel technology, and resources, such as education and income, facilitate familiarization with an emerging technology such as this:

H1: Male gender, youth, education, and income are positively related to knowledge about deepfakes.

Similarly, we expect digital skills among internet users to be a predictor of knowledge about deepfake technology. Skills play a critical role in allowing users to make use of those digital technologies they have material access to (Hargittai, 2002; Hargittai & Hinnant, 2008; Lutz, 2019; Van Dijk, 2006). Digital skills have been shown to facilitate the adoption of new technology, which allows users to benefit from the technology's use (Carter et al., 2020; Madden et al., 2017). We therefore expect that those high in digital skills are more likely to be aware of new technology and have a higher knowledge about emerging deepfake technology.

H2: Digital skills are positively related to knowledge about deepfakes.

In addition, we expect general motivations to explore new technology among internet users to play a role in this context. Previous studies have pointed out the role of personal innovativeness in user adoption of ICT (e.g., Ahn et al., 2007). Findings indicate that individuals who are open to and interested in new technologies in general feel more competent in dealing with them (Bakke & Henry, 2015), are more likely to adopt them (Chahal & Rani, 2022), and find it easier to understand their structures and internalize their functions (Rasimah et al., 2011). In consequence, users who are more comfortable in the digital world and participate in digital activities enjoy real life advantages over those who don't (Park, 2017).

H3: Personal innovativeness is positively related to knowledge about deepfakes.

Finally, as pointed out by Eichhorn et al. (2022), knowledge access is related to motivational access (cf. Cotter & Reisdorf, 2020; Van Dijk, 2005). In other words, familiarity with or knowledge of a technology can influence users' attitudes towards the technology, and thus their willingness to use it. As discussed above, users still struggle to reliably identify deepfakes (Bray et al., 2023; Lewis et al., 2022), and initial encounters tend to induce insecurity or mistrust (Vaccari & Chadwick, 2020). Risks associated with the technology, such as misinformation, privacy concerns, fraud, and reputational risks, are more salient in public discourse (Citron & Chesney, 2019; Kugler & Pace, 2021; Lee et al., 2021). We therefore expect those with little knowledge of the technology to harbour more negative attitudes towards it, while, conversely, those with more deepfake knowledge are more likely to also take potential benefits of the technology into account.

H4: Deepfake knowledge is positively related to positive attitudes towards deepfake technology.

In consequence, inequalities in deepfake knowledge might lead to inequalities in deepfake perceptions and cause unequal usage and outcomes of the technology (Hargittai & Hinnant, 2008; Van Deursen & Helsper, 2015a; Van Dijk, 2013).

Methods

Sample

Data collection was carried out in cooperation with a certified market research institute in October 2022. Respondents were invited by email and received a small monetary compensation upon completing the survey (1664 respondents). To ensure equivalence with the overall population in Germany, we defined quotas for age, gender, and state of residency (within Germany). For data analysis we only considered respondents who had fully completed the questionnaire (1421 respondents). Among the respondents, 49.3 percent are men and 50.7 percent are women,[1] and the average age is 48 years.[2] Most of the respondents have a monthly net income between 1500 and 3000 euros (44.6%); followed by those with an income below 1500 euros (29.9%) or above 3000 euros (25.5%).[3] Furthermore, educational attainment was measured by the highest educational qualification obtained, with around half of the respondents having a low[4] educational attainment (46.9) and about one quarter each having a medium (28.6) or high educational attainment (24.5).

Measures

At the beginning of the questionnaire, participants were divided into two groups by asking: *Do you know what a 'deepfake' is?* If they answered in the affirmative ($n = 480$), they were asked to describe what they think a deepfake is and to answer 13 knowledge questions. These respondents were then provided with a short definition of what deepfakes are (see Appendix). Afterwards, the respondents watched a deepfake example of Morgan Freeman, simulated by voice actor Boet Schouwink using deepfake technology (Diep Nep, 2021). The group of participants who answered that they did not know what a deepfake is, were immediately provided with the definition and the video example. After watching the video, all respondents were asked to assess the extent to which they agree with a set of risks and opportunities associated with deepfake technology (as identified in the literature review). Then, respondents evaluated their digital skills and personal innovativeness as well as their internet application use and social media use. Finally, further sociodemographic characteristics were determined.

To assess deepfake knowledge, the survey employed an original measure based upon prior literature analyses (Godulla et al., 2021). The measure provided participants with 13 correct or incorrect statements about deepfakes. The participants were asked to indicate whether they considered these statements to be true or false (see Appendix, Table A5). To evaluate attitudes towards deepfakes, participants were confronted with a set of 21 statements describing risks and opportunities associated with deepfakes (see Appendix), and were asked to rate their agreement with these statements on a five-point Likert scale (do not agree at all to fully agree).

Digital skills were assessed following Hargittai and Hsieh (2012) by asking the participants to rate their knowledge of six internet specific terms (e.g., *PDF, Cache, Spyware*) on a five-point Likert scale from not good at all to very good. The degree of participants' innovativeness in new technology was evaluated following Raaij and Schepers (2008, p. 846). Participants were asked to rate their personal innovativeness considering four items (e.g., *when I hear about a new technology, I look for ways to try it)* on a five-point Likert scale (do not agree at all to fully agree).

Furthermore, participants were asked about their internet application use (e.g., *information services, messaging services, online banking*) and social media use (e.g., *Facebook, YouTube, TikTok*) based on a six-point Likert scale (never to several times a day). Finally, while quotas for specific demographic characteristics (gender, age, and federal state of residence) were determined at the beginning of the questionnaire, further characteristics such as educational attainment and income were queried at the end of the questionnaire. An overview of all items used to measure participants' digital skills, personal innovativeness and perception of deepfake risks and opportunities is provided in the Appendix (Table A1).

Mean indices. Mean indices were calculated to measure participants' internet and social media use, digital skills, personal innovativeness and perception of deepfake risks and opportunities. The variables generated for internet application usage, social media usage, digital skills, personal innovativeness and sociodemographic variables were used for a linear regression on deepfake knowledge using the software SPSS Statistics 2021 (version 28.0.1.0). The variables for both deepfake risk perception and deepfake opportunity perception were used as dependent variables in subsequent analyses.

Results

Analysis of knowledge about deepfakes

Initial descriptive findings show that the majority of participants (67.1%) stated not to know what a deepfake is and therefore did not answer the knowledge questions. In addition, findings of the knowledge questions show that more than ten percent have little knowledge about deepfakes. Hence, in total, more than three quarters of the respondents have no or little knowledge about deepfakes. In contrast, only about one fifth have a medium or high level of knowledge about the technology (see Table 1, $M = 1.72$, SD = 2.71).

Table 1. Differentiation of deepfake knowledge.

Variable		Frequency (%)
Deepfake knowledge	None	954 (67.1)
	Low	161 (11.3)
	Medium	247 (17.4)
	High	59 (4.2)

Note: $N = 1421$.

Answering one to four knowledge questions correctly was coded as a low level of knowledge (11.3%), five to seven correct answers as a medium knowledge level (17.4%) and eight or more correct answers as a high level of knowledge (4.2%).

To analyse predictors of deepfake knowledge, a multiple linear regression[5] was conducted (see Table 2). To assess deepfake knowledge, a summative index of correctly answered knowledge questions was calculated.

Results largely show support for H1, as youth, male gender and educational attainment are positively related with deepfake knowledge, while income is not. Age has only a small effect size, while gender is the strongest predictor among the sociodemographic variables. We also find support for H2, as digital skills are strongly and significantly related to deepfake knowledge. Additionally, results show that personal innovativeness does also relate positively to deepfake knowledge (H3). When we control for internet application and social media usage, the latter is strongly and positively related to deepfake knowledge. Overall, the regression model explains 21 percent (corrected $R^2 = .21$) of the variance in the level of knowledge about deepfakes ($F(8, 1286) = 42.85$; $p < .001$).

Assessments of risks and opportunities of deepfakes

Turning to the risks associated with deepfakes (see Appendix Table A2), respondents generally fear that the prevalence of deepfakes will increase (80.3%) and that their detection will become more and more difficult (86.6%). In consequence, 77.5% of respondents have concerns that the line between reality and fiction is increasingly blurring due to the development of the technology. In addition, a large proportion of respondents agree the technology creates risks such as (negative) influences on public opinion (e.g., election campaigns: 79.5%) and harm to democracy (e.g., due to misinformation: 77.1%). Interestingly, only a relatively low percentage of respondents report concerns about accidently

Table 2. Linear regression on deepfake knowledge.

Predictors	*b*	*SE B*	ß
Age	−.02***	.01	−.12
Gender	−.51***	.14	−.09
Educational level	.29**	.09	.09
Income	.02	.03	.01
Digital skills	.56***	.08	.20
Personal innovativeness	.19*	.08	.07
Internet application usage	−.07	.14	−.02
Social media usage	.54***	.12	.18
Intercept	−.18	.59	

Note: $R^2 = .21$, $N = 1295$, deviating N due to random missingness. Dependent variable = Deepfake knowledge. Gender coded as male (1) and female (2). * $p < .05$. ** $p < .01$. *** $p < .001$.

spreading a deepfake (42.3%). Further analyses revealed significant differences between male and female respondents when it comes to the perceived risks of becoming a victim of a harmful deepfake oneself and having to protect personal data more strictly in the future (see Appendix Table A4). In both cases, women expressed a stronger fear of becoming a victim of a harmful deepfake and of protecting personal data more strictly.

In contrast to these risk perceptions, respondents' views of the opportunities offered by deepfake technology are more hesitant (Appendix Table A3). The use of deepfakes finds most approval in creative industries such as the film industry (63%), video game (55.4%), fashion (51.6%), and art industry (41.5). Focusing on the media industry, only one third of respondents agree that deepfakes may be used to foster creativity in media content production (38%). Furthermore, respondents rarely perceive deepfake technology as a benefit for their personal usage (32.9%). A subsequent analysis has shown that there are significant differences between male and female respondents regarding the possibility of creating consensual pornographic deepfakes. In this context, male respondents more strongly agreed on the potential of using the technology to create consensus-based pornographic content compared to female respondents (see Appendix Table A4).

Attitudes towards deepfake technology

Having established predictors of deepfake knowledge and explored the audiences' assessments of the risks and potentials of deepfake technology, we proceed by examining the extent to which differences in knowledge about deepfakes relate to user attitudes towards the technology by conducting a series of multiple linear regressions and a *t*-test (see Tables 3–5).

First, a linear regression analysis on the perception of deepfake risks was conducted to assess predictors of negative attitudes towards the technology (Table 3). With regards to H4, results show that knowledge about deepfakes is not significantly related to negative assessments of the technology. Focusing on sociodemographic characteristics, female gender strongly and negatively relates to risk perceptions, while no significant influence is shown for age, educational level, or income. Interestingly, digital skills and internet application use are also positively related to risk perceptions. No significant relationships were found for respondents' personal innovativeness and social media use.

Table 3. Linear regression on risk perception.

Predictors	*b*	*SE B*	ß
Age	.00	.00	.07
Gender	.31***	.05	.20
Educational level	−.02	.03	−.02
Income	−.01	.01	−.03
Digital skills	.06*	.03	.07
Personal innovativeness	−.01	.03	−.01
Internet application usage	.09*	.04	.09
Social media usage	−.03	.04	−.04
Deepfake knowledge	.00	.01	.01
Intercept	2.87***	1.81	

Note: R^2 = .04, N = 1294, deviating N due to random missingness. Dependent variable = Deepfake risk perception. Gender coded as male (1) and female (2). * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4. Linear regression on opportunity perception.

Predictors	*b*	*SE B*	ß
Age	−.01**	.00	−.09
Gender	−.14**	.05	−.08
Educational level	.02	.03	.02
Income	−.04***	.01	−.09
Digital skills	.05	.03	.06
Personal innovativeness	.13***	.03	.14
Internet application usage	.11*	.05	.10
Social media usage	.10*	.04	.11
Deepfake knowledge	.02*	.01	.07
Intercept	2.44***	.19	

Note: $R^2 = .18$, $N = 1294$, deviating N due to random missingness. Dependent variable = Deepfake opportunity perception. Gender coded as male (1) and female (2). * $p < .05$. ** $p < .01$. *** $p < .001$.

Turning to the (rarer) positive assessments of deepfake technology (Table 4), we find that deepfake knowledge is indeed weakly but positively related to deepfake opportunity perceptions. We thus find mixed results for H4. Youth, (female) gender and income are negatively related to a positive assessment of deepfake technology, but education is not. After gender, personal innovativeness is the strongest predictor of opportunity perceptions in the model. Internet application use and social media use are both also positively related to a positive assessment of deepfakes.

To complement the analysis of risk and opportunity perceptions towards deepfakes, participants' overall perception of the technology was assessed in the questionnaire using a slider ranging from dangerous (0) to promising (100). Results show that respondents rate the technology as more dangerous than promising overall ($M = 26.80$, SD = 21.92). We divided the sample into subsamples, one with low deepfake knowledge (no familiarity and up to four correct knowledge questions: 78.5%) and one with relatively high deepfake knowledge (more than four correct answers: 21.5%). The results of a t-test for independent samples (Table 5) show that there is a highly significant ($t(701) = -3.59$, $p < .001$) difference in the overall assessment of the deepfake technology. Respondents with a low knowledge of deepfakes ($M = 25.34$, SD = 20.74) perceive the technology to be more dangerous compared to respondents with a high knowledge ($M = 30.18$; SD = 24.12). This further indicates that deepfake knowledge does play a role in user attitudes towards the technology (H4).

Additionally, further analyses on gender differences in the overall assessment of deepfakes were conducted (see Appendix Tables A6–A8). First, the results of a t-test show that female respondents (M = 29.50, SD = .89) perceive the technology to be more dangerous compared to male respondents (M = 24.15, SD = .76). These results are highly significant ($t(1398) = 4.59$, $p < .001$). Second, the impact of deepfake knowledge and digital skills on the technology's overall assessment were analyzed separately for male and female respondents. The results of two t-tests show that male respondents with high deepfake

Table 5. Differences in overall deepfake assessment based on the level of deepfake knowledge.

t-test	Knowledge	*N*	*M*	SD	*T*	df	*p*	Mean difference	SEM
	Low	980	25.34	20.74	−3.59	701	<.001	−4.83	1.35
	High	422	30.18	24.12					

Note: Deviated *N* due to random missingness.

knowledge ($M = 32.34$, SD = 1.43) and high digital skills ($M = 32.11$, SD = 1.33) perceive the technology to be less dangerous compared to those with low knowledge ($M = 27.55$, SD = 1.11) and low digital skills ($M = 27.13$, SD = 1.16). The results of these t-tests are highly significant ($t(691) = 2.63$, $p < .01$; $t(691) = 2.83$, $p < .01$). When focusing on female respondents, no significant results were found for deepfake knowledge. However, female respondents with high digital skills ($M = 25.73$, SD = 1.14) compared to low ($M = 22.51$, SD = 1) perceive the technology to be slightly less dangerous ($t(697) = 2.12$, $p < .05$).

Discussion

This study addresses digital inequalities in the context of AI by examining the role of sociodemographic factors, digital skills, and personal innovativeness in knowledge about deepfakes (RQ1), the attitudes of internet users towards deepfake technology, as well as the impact of deepfake knowledge on the perception of risks and opportunities of the technology (RQ2). Examining knowledge about deepfakes and attitudes towards the technology is important because knowledge is a prerequisite for technology access (Eichhorn et al., 2022), as are attitudes that motivate users to access the technology (Van Dijk, 2005). Lack of knowledge and negative attitudes therefore predict avoidance of a novel technology, such as deepfakes, thus contributing to an AI divide (Carter et al., 2020).

First, we find that in our sample of German internet users, knowledge about deepfake technology and its applications is rather low. This finding directionally corresponds with previous studies showing that users struggle to identify deepfakes (Bray et al., 2023; Lewis et al., 2022). When examining the role of sociodemographic factors in this context, we find deepfake knowledge to be more prevalent among those of lower age, higher educational attainment and male gender, as well as higher digital skills and personal innovativeness. Furthermore, frequent usage of multiple social media platforms predicts knowledge about deepfakes. These findings are largely in line with previous research on determinants of digital inequalities (Hargittai, 2002; Hargittai & Walejko, 2008; Van Dijk, 2006). However, the role of gender is markedly large in our examination of deepfakes, while age plays a surprisingly small role. The role of gender in the development and proliferation of AI technology has been examined repeatedly, with studies finding a 'lack of diversity in both data and developers, programmer bias, and the existing gender bias in society, now amplified through AI' (Nadeem et al., 2020, p. 1). Deepfakes, in particular, with their roots and frequent use in the context of pornography may repel women from familiarizing themselves with this technology. As noted above, digital inequalities studies need to take the perceptions of specific technologies, such as deepfakes, into consideration (Helsper, 2017b).

Our findings demonstrate that, currently, the perception of deepfakes is dominated by a focus on risks rather than benefits. Risks that participants perceive as particularly urgent include data protection concerns, deepfake detection, the technology's use to influence the public opinion and the difficulty to distinguish between reality and fiction. This dominance of concerns might be due to audio-visual content traditionally being ascribed high levels of trustworthiness and reliability (Vaccari & Chadwick, 2020). Altering audio-visual media is thus likely to trigger uncertainty among audiences. While most of the respondents are concerned that, in general, it might become increasingly difficult to distinguish between reality and fiction, relatively few of them fear accidently spreading a deepfake themselves. A possible reason for this result could be a third person effect that has repeatedly been observed in studies of mis- and disinformation (Altay & Acerbi, 2023). Accordingly, respondents might estimate a stronger influence of hyper-realistic, authentic appearing deepfakes on the deception of others than on themselves.

Results of a t-test indicate that users with higher levels of deepfake knowledge have a less critical attitude overall towards the technology. Having conducted regression analyses on both negative and positive attitudes towards deepfakes, we find that knowledge does not significantly relate to negative attitudes – possibly because negative attitudes are almost ubiquitous – but do (weakly) predict positive attitudes. So, overall, we find some evidence that deepfake knowledge plays a role in users' attitudes towards the technology (cf., Eichhorn et al., 2022) – which is important as attitudes relate to motivation to access the technology (Van Dijk, 2005). These relationships may still be weak as overall knowledge in the population is very low and attitudes are overwhelmingly critical. Effect sizes may thus change over time as the technology proliferates and users become more accustomed to this AI application.

When comparing results for the regression models on negative vs. positive attitudes, some differences emerge: Digital skills positively relate to negative attitudes, so more skilled users may be better able to conceive of AI technology dangers. Frequent usage of multiple internet applications is positively related to both positive and negative attitudes towards deepfake technology. Frequent internet users may thus have more differentiated views of the technology's opportunities and risks. Social media usage, however, is only related to the perception of deepfake opportunities. Age is negatively related to the perception of opportunities, but overall plays a relatively minor role. Gender emerges as by far the strongest predictor of concerns. Female respondents are particularly afraid of falling victim to a harmful deepfake. Additionally, female gender is strongly negatively related to deepfake knowledge and negatively predicts positive attitudes towards deepfakes. As noted above, this may be due to the gendered manifestation and impact of AI technology, in general, and the particular challenges of deepfake technology, in particular (Raffaghello et al., 2019). Furthermore, knowledge about the technology is a predictor for male respondents' overall assessment of deepfakes, while it does not influence female respondents' overall perception of the technology. Hence, female respondents seem to be more sceptical about deepfakes in general, regardless of their knowledge about the technology.

This study is subject to some limitations: We present data from a cross-sectional study, therefore our findings are based on correlational evidence. The measures applied and some of our findings are limited by the novelty of the technology under investigation. We have developed original measures for deepfake knowledge and attitudes based on a review of the available literature, which is still limited in scope. As discussed above, few German internet users have yet knowingly encountered deepfakes, so overall levels of familiarity are low, which may also affect user attitudes towards a largely unfamiliar technology. Future studies should attempt to replicate the present findings when users have had a chance to learn more about deepfakes. Given our findings on the important role of gender in knowledge of and attitudes towards deepfakes, future studies should expand this research to examine sexual identity and race. Previous studies have documented harmful uses of the technology targeted at minorities and marginalize groups

(Paris, 2021). Such examinations would allow for an intersectional analysis of digital inequalities in the context of deepfakes (cf., Tsatsou, 2022).

Nevertheless, this study contributes to emergent research on the AI divide (Carter et al., 2020). To our knowledge, no previous study has examined deepfakes through a digital inequalities lens (cf., Lutz, 2019; Vartanova & Gladkova, 2019). Previous research on deepfakes has focused on human or technological deepfake detection (cf., Godulla et al., 2021), rather than knowledge or attitudes. Also, most previous studies have focused on risks associated with deepfakes, for example in the political domain (Dobber et al., 2020; Hameleers et al., 2023; Vaccari & Chadwick, 2020), with few taking a broader view and examining both positive and negative attitudes. From an AI divide perspective in particular, future studies might explore how deepfakes can support the accessibility of information (e.g., services for individuals with disabilities) through, for example, speech synthesis (Bendel, 2017). In other words, the technology may be helpful in ameliorating some inequalities.

Notes

1. The questionnaire provided three gender options (male, female, diverse). Since only two participants reported a diverse gender, these persons were not included in the analysis.
2. The youngest participant was 18 and the oldest 74 years of age.
3. Income was assessed using nine categories (in Euro; 0–499; 500–999, 1000–1499, 1500–1999, 2000–2499, 2500–2999, 3000–3499, 3500–3999, 4000+).
4. Low educational attainment: no qualification, still at school, lower or intermediate secondary school leaving certificate; medium educational attainment: general higher education entrance qualification; high educational attainment: university degree, doctorate.
5. Note: In the linear multiple regression, respondents who stated that they did not know what a deepfake was were also included (0-coded). For the analysis it was checked that all prerequisites for multiple linear regression were met (linear relationship between variables, independence of residuals, multicollinearity, homoscedasticity, normal distribution of residuals).

Notes on contributors

Daniel Bendahan Bitton (M.A.) is a research associate at the departments of Empirical Communication and Media Research and Online Communication at the Institute for Communication and Media Studies (IfKMW) at Leipzig University. His research interests include the reception and effects of media content on audiences, technological impact assessment, fundamental research on deepfakes and science communication [email: daniel.bendahan@uni-leipzig.de].

Christian Pieter Hoffmann is a professor of Communication Management at the Institute for Communication and Media Studies (IfKMW) and for Political Communication at the Institute for Political Science at Leipzig University. His research interests include online communication, financial communication, and political communication. Scientific works by Christian concentrate on (political) online participation, privacy, online trust, and cynicism [email: christian.hoffmann@uni-leipzig.de].

Alexander Godulla is a professor of Empirical Communication and Media Research at the Institute for Communication and Media Studies (IfKMW) at Leipzig University. His research interests include the reception and effects of media content on audiences, international journalism research, technological impact assessment in public communication and visual communication research [email: alexander.godulla@uni-leipzig.de].

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project is co-financed with tax revenue based on the budget approved by the state parliament of Saxony.

ORCID

Daniel Bendahan Bitton http://orcid.org/0000-0001-7258-8753
Christian Pieter Hoffmann http://orcid.org/0000-0002-5282-6950
Alexander Godulla http://orcid.org/0000-0002-1011-0639

References

Ahn, T., Ryu, S., & Han, I. (2007). The impact of web quality and playfulness on user acceptance of online retailing. *Information & Management*, *44*(3), 263–275. https://doi.org/10.1016/j.im.2006.12.008

Altay, S., & Acerbi, A. (2023). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*, *9*(1), 1–13.

Arguedas, A. R., & Simon, F. M. (2023). Automating democracy: Generative AI, journalism, and the future of democracy. https://ora.ox.ac.uk/objects/uuid:0965ad50-b55b-4591-8c3b-7be0c587d5e7

Bakke, S., & Henry, R. (2015). Unraveling the mystery of new technology use: An investigation into the interplay of desire for control, computer self-efficacy, and personal innovativeness. *AIS Transactions on Human-Computer Interaction*, *7*(4), 270–293. https://doi.org/10.17705/1thci.00075. https://aisel.aisnet.org/thci/vol7/iss4/3/

Bendel, O. (2017). The synthetization of human voices. *AI & Society*, *34*(1), 83–89. https://doi.org/10.1007/s00146-017-0748-x

Bozic, V. (2023). Strengthening EFL learners' autonomy: exploring the Voscreen app and its impact on encouragement and engagement in technology-based language learning. *Language Education & Technology (LET Journal)*, *3*(2), 96–109. http://www.langedutech.com/letjournal/index.php/let

Bray, S. D., Johnson, S. D., & Kleinberg, B. (2023). Testing human ability to detect 'deepfake' images of human faces. *Journal of Cybersecurity*, *9*(1), 1–18. https://doi.org/10.1093/cybsec/tyad011

Caldera, E. (2019). Reject the evidence of your eyes and ears: Deepfakes and the law of virtual replicants. *Seton Hall Law Review*, *50*(1), 177–205. https://scholarship.shu.edu/shlr/vol50/iss1/5/

Cardenuto, J. P., Yang, J., Padilha, R., Wan, R., Moreira, D., Li, H., & Rocha, A. (2023). The age of synthetic realities: Challenges and opportunities. arXiv preprint (pp. 1–27). https://arxiv.org/pdf/2306.11503.pdf

Carter, L., Liu, D., & Cantrell, C. (2020). Exploring the intersection of the digital divide and artificial intelligence: A hermeneutic literature review. *AIS Transactions on Human-Computer Interaction*, *12*(4), 253–275. https://doi.org/10.17705/1thci.00138

Chahal, J., & Rani, N. (2022). Exploring the acceptance for e-learning among higher education students in India: Combining technology acceptance model with external variables. *Journal of Computing in Higher Education*, *34*(3), 844–867. https://doi.org/10.1007/s12528-022-09327-0

Citron, D. K., & Chesney, R. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, *107*, 1753–1820.

Cotter, K., & Reisdorf, B. C. (2020). Algorithmic knowledge gaps: A new horizon of (digital) inequality. *International Journal of Communication*, *14*(21), 745–765. https://ijoc.org/index.php/ijoc/article/view/12450/2952

Diep Nep. (2021, July 29). This is not Morgan Freeman – A look behind the Deepfake Singularity [video]. YouTube. https://www.youtube.com/watch?v=F4G6GNFz0O8

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.
Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2020). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics*, *26*(1), 69–91. https://doi.org/10.1177/1940161220944364
Eichhorn, T., Jürss, S., & Hoffmann, C. P. (2022). Dimensions of digital inequality in the sharing economy. *Information, Communication & Society*, *25*(3), 395–412. https://doi.org/10.1080/1369118X.2020.1791218
Gao, B., Wang, Y., Xie, H., Hu, Y., & Hu, Y. (2023). Artificial intelligence in advertising: Advancements, challenges, and ethical considerations in targeting, personalization, content creation, and ad optimization. *Sage Open*, *13*(4), 1–20. https://doi.org/10.1177/21582440231210759
Godulla, A., Hoffmann, C. P., & Seibert, D. (2021). Dealing with deepfakes – An interdisciplinary examination of the state of research and implications for communication studies. *Studies in Communication and Media*, *10*(1), 72–96. https://doi.org/10.5771/2192-4007-2021-1-72
Hameleers, M., van der Meer, T. G., & Dobber, T. (2023). They would never say anything like this! Reasons to doubt political Deepfakes. *European Journal of Communication*, *0*, 1–15. https://doi.org/10.1177/02673231231184703
Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4), 1–18. https://doi.org/10.5210/fm.v7i4.942
Hargittai, E., & Hinnant, A. (2008). Digital inequality: Differences in young adults' use of the Internet. *Communication Research*, *35*(5), 602–621. https://doi.org/10.1177/0093650208321782
Hargittai, E., & Hsieh, Y. P. (2012). Succinct survey measures of web-use skills. *Social Science Computer Review*, *30*(1), 95–107. https://doi.org/10.1177/0894439310397146
Hargittai, E., & Walejko, G. (2008). The participation divide: Content creation and sharing in the digital age. *Information, Community and Society*, *11*(2), 239–256. https://doi.org/10.1080/13691180801946150
Helsper, E. J. (2017a). The social relativity of digital exclusion: Applying relative deprivation theory to digital inequalities. *Communication Theory*, *27*(3), 223–242. https://doi.org/10.1111/comt.12110
Helsper, E. J. (2017b). A socio-digital ecology approach to understanding digital inequalities among young people. *Journal of Children and Media*, *11*(2), 256–260. https://doi.org/10.1080/17482798.2017.1306370
Jungherr, A., & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, *33*(2-3), 164–173. https://doi.org/10.1093/ct/qtad006
Kitsara, I. (2022). Artificial intelligence and the digital divide: From an innovation perspective. In A. Bounfour (Ed.), *Platforms and artificial intelligence: The next generation of competences* (pp. 245–265). Springer International Publishing.
Kolagati, S., Priyadharshini, T., & Rajam, V. M. A. (2022). Exposing deepfakes using a deep multilayer perceptron–convolutional neural network model. *International Journal of Information Management Data Insights*, *2*(1), 1–8. https://doi.org/10.1016/j.jjimei.2021.100054
Kugler, M. B., & Pace, C. (2021). Deepfake privacy: Attitudes and regulation. *Northwestern University Law Review*, *116*(3), 611–680. https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/illlr116&id=628&men_tab=srchresults
Kurpicz-Briki, M., Glauser, C., & Schmid, L. (2023). Vom digital divide zum cyber divide HMD Praxis der Wirtschaftsinformatik, 1-17. https://doi.org/10.1365/s40702-023-00971-3
Lainjo, B. (2023). The global social dynamics and inequalities of artificial intelligence. *International Journal of Innovation Scientific Research and Review*, *5*(8), 4966–4974. http://www.journalijisr.com/sites/default/files/issues-pdf/IJISRR-1306.pdf
Lee, Y., Huang, K. T., Blom, R., Schriner, R., & Ciccarelli, C. A. (2021). To believe or not to believe: framing analysis of content and audience response of top 10 deepfake videos on YouTube. *Cyberpsychology, Behavior, and Social Networking*, *24*(3), 153–158. https://doi.org/10.1089/cyber.2020.0176

Lewis, A., Vu, P., & Chowdhury, A. (2022). *Do content warnings help people spot a deepfake? Evidence from two experiments*. The Royal Society. https://royalsociety.org/-/media/policy/projects/online-information-environment/do-content-warnings-help-people-spot-a-deepfake.pdf

Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and bit data. *Human Behavior and Emerging Technologies*, *1*(2), 141–148. https://doi.org/10.1002/hbe2.140

Lutz, C., Hoffmann, C. P., & Meckel, M. (2014). Beyond just politics: A systematic literature review of online participation. *First Monday*, *19*(7), 141–148. https://doi.org/10.5210/fm.v19i7.5260

Madden, M., Gilman, M., Levy, K., & Marwick, A. (2017). Privacy, poverty, and big data: A matrix of vulnerabilities for poor Americans. *Washington University Law Review*, *95*, 53–126. https://openscholarship.wustl.edu/law_lawreview/vol95/iss1/6/

Marwick, A. E. (2022). Privacy without power: What privacy research can learn from surveillance studies. *Surveillance & Society*, *20*(4), 397–405. https://doi.org/10.24908/ss.v20i4.16009

Morrow, E. (2022). Priming privacy: The effect of privacy news consumption on privacy attitudes, beliefs, and knowledge. *Journal of Broadcasting & Electronic Media*, *66*(5), 772–793. https://doi.org/10.1080/08838151.2022.2138888

Nadeem, A., Abedin, B., & Marjanovic, O. (2020). Gender bias in AI: A review of contributing factors and mitigating strategies. *ACIS 2020 Proceedings*, *27*, 1–12. https://aisel.aisnet.org/acis2020/2

Öhman, C. (2020). Introducing the pervert's dilemma: A contribution to the critique of Deepfake Pornography. *Ethics and Information Technology*, *22*(2), 133–140. https://doi.org/10.1007/s10676-019-09522-1

Paris, B. (2021). Configuring fakes: Digitized bodies, the politics of evidence, and agency. *Social Media + Society*, *7*(4), 1–13. https://doi.org/10.1177/20563051211062919

Park, S. (2017). *Digital capital*. Palgrave Macmillan.

Raaij, E. M., & Schepers, J. J. L. (2008). The acceptance and use of a virtual learning environment in China. *Computers & Education*, *50*(3), 838–852. https://doi.org/10.1016/j.compedu.2006.09.001

Raffaghello, I., Kastalio, L., Kalf, S., & Paisley, E. (2019). What does a feminist approach to deepfake pornography look like? https://mastersofmedia.hum.uva.nl/blog/2019/10/24/what-does-a-feminist-approach-to-deepfake-pornography-look-like/

Rasimah, C. M. Y., Ahmad, A., & Zaman, H. B. (2011). Evaluation of user acceptance of mixed reality technology. *Australasian Journal of Educational Technology*, *27*(8), 1369–1387. https://doi.org/10.14742/ajet.899

Reisdorf, B. C., Blank, G., & Dutton, W. H. (2019). Internet cultures and digital inequalities. In M. Graham (Ed.), *Dutton society and the internet: How networks of information and communication are changing our lives* (pp. 80–85). Oxford University Press.

Reisdorf, B. C., & Groselj, D. (2017). Internet (non-)use types and motivational access: Implications for digital inequalities research. *New Media & Society*, *19*(8), 1157–1176. https://doi.org/10.1177/1461444815621539

Rogers, E. M. (1983). *Diffusion of innovations*. Free Press.

Seibert, D. (2023). Deepfakes – Temporary hype or long-term innovation driver? In A. Godulla, & S. Boehm (Eds.), *Digital disruption and media transformation: How technological innovation shapes the future of communication* (pp. 143–152). Springer VS.

Simon, F. M., & Isaza-Ibarra, L. F. (2023). AI in the news: Reshaping the information ecosystem? https://ora.ox.ac.uk/objects/uuid:9947240c-06d3-42c2-9a23-57ff4559b63c/files/sj3860857t

Simonite, T. (2019, July 10). 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*. https://www.wired.com/story/most-deepfakes-porn-multiplying-fast/

Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., & Moe, H. (2022). Responsible media technology and AI: Challenges and research directions. *AI and Ethics*, *2*(4), 585–594. https://doi.org/10.1007/s43681-021-00126-4

Tsatsou, P. (2022). Vulnerable people's digital inclusion: Intersectionality patterns and associated lessons. *Information, Communication & Society*, *25*(10), 1475–1494. https://doi.org/10.1080/1369118X.2021.1873402

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. https://doi.org/10.1177/2056305120903408

Van Deursen, A., & Helsper, E.. (2015a). The third-level digital divide: Who benefits most from being online?. *Communication and Information Technologies Annual*, *10*, 29–52.

Van Deursen, A., & Helsper, E. (2015b). A nuanced understanding of internet use and non-use among the elderly. *European Journal of Communication*, *30*(2), 171–187. https://doi.org/10.1177/0267323115578059

Van Deursen, A., & Helsper, E. (2018). Collateral benefits of Internet use: Explaining the diverse outcomes of engaging with the Internet. *New Media & Society*, *20*(7), 2333–2351. https://doi.org/10.1177/1461444817715282

Van Deursen, A., Helsper, E., Eynon, R., & Van Dijk, J. (2017). The compoundness and sequentiality of digital inequality. *International Journal of Communication*, *11*, 452–473.

Van Deursen, A., & Van Dijk, J. (2015). Toward a multifaceted model of internet access for understanding digital divides: An empirical investigation. *The Information Society*, *31*(5), 379–391. https://doi.org/10.1080/01972243.2015.1069770

Van Deursen, A. J., & Van Dijk, J. A. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media & Society*, *21*(2), 354–375. https://doi.org/10.1177/1461444818797082

Van Deursen, A., van Dijk, J., & Helsper, E. (2014). Investigating outcomes of online engagement. Media@LSE. http://eprints.lse.ac.uk/id/eprint/59994

Van Dijk, J. A. (2005). *The deepening divide*. Sage.

Van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4–5), 221–235. https://doi.org/10.1016/j.poetic.2006.05.004

Van Dijk, J. A. (2013). A theory of the digital divide. In M. Ragnedda, & G. W. Muscher (Eds.), *The digital divide. The internet and social inequality in international perspective* (pp. 29–52). Routledge.

Vartanova, E., & Gladkova, A. (2019). New forms of the digital divide. In J. Trappel (Ed.), *Digital media inequalities: Policies against divides, distrust and discrimination* (pp. 193–213). Nordicom. https://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-11846

Wagner, T. L., & Blewer, A. (2019). The word real is no longer real: Deepfakes, gender, and the challenges of AI-altered video. *Open Information Science*, *3*(1), 32–46. https://doi.org/10.1515/opis-2019-0003

Wang, L., & Liu, C. (2022). The role of motivation in the digital divide: The moderating effect of device access. *Mobile Media & Communication*, *10*(3), 406–420. https://doi.org/10.1177/20501579211060621

Yu, P. K. (2020). The algorithmic divide and equality in the age of artificial intelligence. *Florida Law Review*, *72*(2), 331–390.

Appendix

Deepfake definition (as provided in the questionnaire):

> From this point on, we understand a deepfake to be manipulated audiovisual media content, e.g., a video that has been generated with the help of artificial intelligence. Deepfakes allow individuals to be depicted in any situation imaginable, even if it did not actually take place. In addition, existing audio recordings of a person's voice can be manipulated. Deepfakes can therefore make the person being depicted say any sentence. Low-quality fakes, so-called 'cheapfakes', which are not created with AI, are not understood as deepfakes. This includes, for example, videos taken out of context or

videos edited together to give a false impression of what happened. Similarly, filters found on Instagram or TikTok are not considered to be deepfakes (e.g., changing eye or hair colour, dog ear filters).

Table A1. Measurement instrument.

Construct	Item	Mean	SD
Digital skills	*Knowledge of the Internet terms …*		
($\alpha = 0.88$)	Advanced Search	3.35	1.16
Adopted from Hargittai and Hsieh	PDF	3.81	1.11
(2012)	Spyware	2.70	1.30
($M = 3.11$ and SD = 1.27)	Wiki	3.05	1.37
	Cache	2.76	1.36
	Phishing	2.98	1.33
Personal innovativeness	When I hear about a new technology, I look for ways to try it out.	2.71	1.18
($\alpha = 0.83$) Adopted from Raaij and Schepers (2008)	In my circle of friends, I'm usually the first person to try out a new technology.	2.25	1.24
($M = 2.63$ and SD = 1.22)	I'm generally reluctant to try out new technologies.[a]	2.92	1.22
	I like experimenting with new technologies.	2.64	1.22
Deepfake risk perception	*I have concerns …*		
($\alpha = 0.90$), Developed following Godulla et al. (2021)	About falling for a deepfake.	3.81	1.07
($M = 3.79$, SD = 1.08)	About becoming a victim of a harmful deepfake (e.g., pornographic deepfakes, scams on the phone).	3.25	1.28
	About deepfakes influencing public opinion (e.g., being used to influence election campaigns).	4.07	1.02
	That soon I won't be able to trust any video content.	3.81	1.08
	That the line between reality and fiction is blurring.	4.04	1.00
	That deepfakes will harm our democracy (e.g., by spreading misinformation/fake news).	4.05	1.03
	That anyone is capable of creating deepfakes.	3.52	1.13
	That manipulation of personal data (e.g., audio or images of me) will become easier.	3.92	1.06
	That I will have to protect my personal data even more in the future (e.g., as it can be manipulated for harmful purposes).	3.99	1.04
	About falsely spreading deepfakes.	2.59	1.40
	That detecting deepfakes will become increasingly difficult.	4.28	.92
	About the increasing prevalence of deepfakes.	4.09	.96
Deepfake opportunity perception	*In my opinion …*		
($\alpha = 0.89$), Developed following	Deepfakes can also be used for beneficial purposes.	3.07	1.23
Godulla et al. (2021) and Seibert (2023)	Deepfakes provide me with benefits (e.g., by creating Funny videos of me and my friends).	2.62	1.24
($M = 2.89$, SD = 1.20)	There are advantages in the fact that any person is able to create deepfakes.	1.99	1.08
	Deepfakes provide benefits to the film industry (e.g., making action scenes more realistic or displaying people who already died in movies).	3.59	1.20
	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).	2.96	1.22
	Deepfakes provide benefits to the art industry (e.g., creating new artwork in the style of famous artists).	3.30	1.20
	Deepfakes offer advantages for the fashion industry (e.g., virtual try-on clothes).	3.39	1.22
	Deepfakes provide benefits to the video game industry (e.g., personalized avatars in games).	2.91	1.19
	Deepfakes foster creativity in media content production.	2.21	1.23

[a]Reverse coded.

Table A2. Perception of risks of deepfake technology.

I have concerns …	Completely or rather disagree	Neither nor	Completely or rather agree	Total
About falling for a deepfake.	90 (6.6)	356 (26.1)	916 (67.3)	1362 (100)
About becoming a victim of a harmful deepfake (e.g., pornographic deepfakes, scams on the phone).	224 (18)	390 (31.3)	633 (50.7)	1247 (100)
About deepfakes influencing public opinion (e.g., being used to influence election campaigns).	49 (3.6)	231 (16.9)	1090 (79.5)	1370 (100)
That soon I won't be able to trust any video content.	91 (6.7)	351 (25.8)	919 (67.5)	1361 (100)
That the line between reality and fiction is blurring.	54 (3.9)	256 (18.6)	1067 (77.5)	1377 (100)
That deepfakes will harm our democracy (e.g., by spreading misinformation/fake news).	58 (4.2)	257 (18.7)	1058 (77.1)	1373 (100)
That anyone is capable of creating deepfakes.	172 (12.9)	393 (29.4)	771 (57.7)	1336 (100)
That manipulation of personal data (e.g., audio or images of me) will become easier.	83 (6.1)	282 (20.6)	1001 (70.4)	1366 (100)
That I will have to protect my personal data even more in the future (e.g., as it can be manipulated for harmful purposes).	71 (5.2)	285 (20.7)	1019 (74.1)	1375 (100)
About falsely spreading deepfakes.	248 (25.8)	307 (31.9)	408 (42.3)	963 (100)
That detecting deepfakes will become increasingly difficult.	37 (2.7)	149 (10.7)	1201 (86.6)	1387 (100)
About the increasing prevalence of deepfakes.	42 (3)	230 (16.7)	1109 (80.3)	1381 (100)

Table A3. Perception of opportunities of deepfake technology.

In my opinion …	Completely or rather disagree	Neither nor	Completely or rather agree	Total
Deepfakes can also be used for beneficial purposes.	219 (18.1)	451 (37.1)	543 (44.8)	1213 (100)
Deepfakes provide me with benefits (e.g., by creating funny videos of me and my friends).	289 (27.2)	424 (39.9)	349 (32.9)	1062 (100)
There are advantages in the fact that any person is able to create deepfakes.	353 (44.5)	302 (38.0)	139 (17.5)	794 (100)
Deepfakes provide benefits to the film industry (e.g., making action scenes more realistic or displaying people who already died in movies).	110 (8.5)	371 (28.5)	818 (63)	1299 (100)
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).	259 (30.9)	344 (41.0)	236 (28.1)	839 (100)
Deepfakes provide benefits to the art industry (e.g., creating new artwork in the style of famous artists).	225 (19)	468 (39.5)	491 (41.5)	1184 (100)
Deepfakes offer advantages for the fashion industry (e.g., virtual try-on clothes).	162 (12.8)	451 (35.6)	654 (51.6)	1267 (100)
Deepfakes provide benefits to the video game industry (e.g., personalized avatars in games).	153 (12)	416 (32.6)	706 (55.4)	1275 (100)
Deepfakes foster creativity in media content production.	244 (20.5)	493 (41.5)	452 (38)	1189 (100)

Table A4. Gender differences in risk and opportunity perception.

Variable	Gender	Completely or rather disagree	Neither nor	Completely or rather agree	Chi^2/V
Risk of falling victim to a harmful deepfake	Male	131 (21.8)	210 (35)	259 (43.2)	$Chi^2(2) = 27.9$,
	Female	93 (14.4)	179 (27.8)	373 (57.8)	$p < .001$, V = .15
Risk of having to protect personal data even more in the future	Male	49 (7.3)	152 (22.7)	469 (70)	$Chi^2(2) = 16.9$,
	Female	22 (3.1)	133 (18.9)	548 (78)	$p < .001$, V = .11
Potential of creating consensus-based pornographic deepfakes	Male	137 (29.3)	172 (36.8)	159 (34)	$Chi^2(2) = 18.9$,
	Female	122 (33)	172 (46.5)	76 (20.5)	$p < .001$, V = .15

Table A5. Knowledge questions about deepfake technology.

Item	Correct	Wrong	Don't know
Most of the deepfakes circulating on the internet are pornographic in nature. [correct]	106 (22.1)	184 (38.3)	190 (39.6)
The term deepfake applies exclusively to manipulated video content. [wrong]	168 (35)	198 (41.2)	114 (23.8)
Deepfakes are created using artificial intelligence. [correct]	315 (65.6)	65 (13.5)	100 (20.1)
Deepfakes can be used to visualize events that did not actually take place. [correct]	414 (86.3)	27 (5.6)	39 (8.1)
Existing recordings of a person cannot be manipulated using deepfake technologies. [wrong]	72 (15)	334 (69.6)	74 (15.4)
The victims of harmful deepfakes are mainly politicians. [wrong]	156 (32.5)	180 (37.5)	144 (30)
Almost anyone with Internet access can create a deepfake. [correct]	282 (58.8)	81 (16.8)	117 (24.4)
With the help of appropriate software, deepfakes can be easily recognized as such at an early stage. [wrong]	222 (46.3)	92 (19.2)	166 (34.5)
Only experts with access to the relevant technologies can create deepfakes. [wrong]	101 (21)	288 (60)	91 (19)
Deepfakes have been circulating on the internet for around ten years. [wrong]	158 (32.9)	82 (17.1)	240 (50)
The first video labelled as a deepfake was circulated via the platform Reddit. [correct]	98 (20.4)	31 (6.5)	351 (73.1)
The creation and distribution of deepfakes has been prohibited by law in Germany since 1 July 2022. [wrong]	104 (21.7)	68 (14.2)	308 (64.1)
Technologies used to create deepfakes can be used in areas such as the film industry or the fashion industry. [correct]	373 (77.7)	38 (7.9)	69 (1.4)

Table A6. Overall assessment of deepfakes based on gender.

t-test	Gender	*N*	*M*	SD	*T*	df	*p*	Mean difference	SEM
	Male	693	29.50	.89	4.59	1.398	<.001	5.35	1.17
	Female	707	24.15	.76					

Table A7. Overall assessment of deepfakes based on knowledge and digital skills of male respondents.

t-test		*N*	*M*	SD	*T*	df	*p*	Mean difference	SEM
Knowledge	Low	405	27.55	1.11	2.63	691	<.01	−4.69	1.78
	High	288	32.34	1.43					
Digital skills	Low	363	27.13	1.16	2.83	691	<.01	−4.97	1.76
	High	330	32.11	1.33					

Note: To assess the level of digital skills, respondents were divided into two groups of approximately equal size based on their mean values of all digital skills items (low digital skills: $M \leq 3.4$; high digital skills: $M > 3.4$).

Table A8. Overall assessment of deepfakes based on knowledge and digital skills of female respondents.

t-test		*N*	*M*	SD	*T*	df	*p*	Mean difference	SEM
Knowledge	Low	533	23.62	.83	1.23	705	.11	−2.17	1.77
	High	174	25.79	1.78					
Digital skills	Low	346	22.51	1	2.12	697	<.05	−3.21	1.51
	High	361	25.73	1.14					

Note: To assess the level of digital skills, respondents were divided into two groups of approximately equal size based on their mean values of all digital skills items (low digital skills: $M \leq 2.8$; high digital skills: $M > 2.8$).