

# Engaging with deepfakes: a meta-synthesis from the perspective of social shaping of technology theory

Engaging with  
deepfakes

Pramukh Nanjundaswamy Vasist and Satish Krishnan  
*Information Systems Area, Indian Institute of Management Kozhikode, India*

Received 27 June 2022  
Revised 4 September 2022  
19 September 2022  
8 October 2022  
Accepted 10 October 2022

## Abstract

**Purpose** – This study aims to establish a comprehensive understanding of the intricacies of how individuals engage with deepfakes, focusing on limiting adverse effects and capitalizing on their benefits.

**Design/methodology/approach** – This study conducted a meta-synthesis of qualitative studies on deepfakes, incorporating study-specific analysis followed by a cross-study synthesis.

**Findings** – Based on the meta-synthesis, the study developed an integrated conceptual framework based on the perspectives from the social shaping of technology theory embedding deepfake-related assertions, motivations, the subtleties of digital platforms, and deepfake-related repercussions.

**Research limitations/implications** – The study offers crucial insights into the evolving nature of deepfakes as a socio-technical phenomenon and the significance of platform dynamics in deepfake production. It enables researchers to comprehend the cascading effects of deepfakes and positions them to evaluate deepfake-related risks and associated mitigation mechanisms.

**Practical implications** – The framework that emerges from the study illustrates the influence of platforms on the evolution of deepfakes and assists platform stakeholders in introducing effective platform governance structures to combat the relentless proliferation of deepfakes and their consequences, as well as providing guidance for governments and policymakers to collaborate with platform leaders to set guardrails for deepfake engagement.

**Originality/value** – Deepfakes have been extensively contested for both their beneficial and negative applications and have been accused of heralding an imminent epistemic threat that has been downplayed by some quarters. This diversity of viewpoints necessitates a comprehensive understanding of the phenomenon. In responding to this call, this is one of the first to establish a comprehensive, theoretically informed perspective on how individuals produce, process, and engage with deepfakes through a meta-synthesis of qualitative literature on deepfakes.

**Keywords** Deepfake, Synthetic media, Fake news, Meta-synthesis, Qualitative study

**Paper type** Research paper

## 1. Introduction

Recent advancements in the field of artificial intelligence (AI) have significantly enhanced individuals' abilities to create, modify, and manipulate digital content (Chesney and Citron, 2018; Öhman, 2020). Deepfakes are a genre of hyper-realistic videos that have been digitally altered to represent individuals saying and doing things that never occurred (Westerlund, 2019). This genre of fakery is of paramount relevance in the current technological revolution and constitutes a danger across sectors (Westerlund, 2019). Amidst the pervasive threat of fake news (Borges-Tiago *et al.*, 2020; Domenico *et al.*, 2021), deepfakes elevate this concern by complicating the ability to distinguish false from authentic content (Breen, 2021; Di Domenico and Visentin, 2020). Deepfakes have been widely touted as the next generation of fake news, which threatens to erode trust in online information and challenge our abilities to decipher true information, particularly when cognitive biases and structural characteristics of new age media, such as echo chambers, are considered (Kalpokas and Kalpokiene, 2022a). This poses an epistemic threat to individuals in acquiring false beliefs (Fallis, 2020). As this post-truth



Satish Krishnan thanks the Indian Institute of Management Kozhikode's Chair Associate Professorship for supporting this research.

---

era dawns upon us, the epistemic choices we make may be increasingly challenged by phenomena such as deepfakes.

Deepfakes are being termed the scariest thing to have ever happened to mankind ([Analytics Insight, 2022](#)), with devastating societal consequences ([Modern Diplomacy, 2022](#)). Politics and pornography have been the primary targets of deepfakes ([Gosse and Burkell, 2020](#)), but the phenomenon has also continued to expand to other domains. For example, performing artists in the entertainment industry are exploring legal options for copyright protection from deepfakes ([Bacchi, 2022](#)), while fraudsters are leveraging celebrity deepfakes to steal cryptocurrencies ([Outlook Money, 2022](#)). In a similar vein, deepfakes are being employed in spear-phishing attacks which wreak havoc on firms ([Stupp, 2019](#)), while they have also been used to alter the course of world events such as Russia's invasion of Ukraine by leveraging deepfaked videos of the Ukrainian president ([CNN, 2022](#)). There is, however, a positive aspect to the deepfake phenomenon. Constructive use cases of deepfakes include the deepfaking of murder victims to nab criminals ([Al-Sibai, 2022](#)), animating storytelling ([Ridden, 2022](#)), voice regeneration of celebrities who have lost their voice ([Daly, 2022a](#)), and narration of bedtime stories to kids in parents' deepfake cloned voices ([Daly, 2022b](#)). While such prosocial applications are being acknowledged, the negative ramifications outweigh the advantages and have prompted a variety of corrective actions in many sectors.

Governments are taking notice and drafting action plans to combat deepfakes ([Kabelka, 2022](#)), while platform players are also taking remedial measures to curb their unchecked growth ([Wiggers, 2022](#)). The latter may have little choice but to combat the spread of deepfakes or attract penalties, especially in light of new legislations in the European Union ([Chee, 2022](#)), while other countries may be expected to follow suit. Scholars are advancing discussions around deepfakes, with several research articles discussing deepfake creation and detection (e.g. [Almars, 2021](#); [Mirsky and Lee, 2021](#); [Tolosana et al., 2022](#)). However, research on fake news in general merits greater attention to comprehend how it evolves alongside its potential ramifications ([Internet Research, 2022](#)). In this regard, the psychological impact of deepfakes is garnering scholarly attention and contested as well, with some highlighting the epistemic threat of deepfakes ([Fallis, 2020](#)) while others argue that such concerns may be overblown ([Harris, 2021](#)). Therefore, it is essential to gain a thorough understanding of the nuances of how individuals engage with deepfakes ([Di Domenico and Visentin, 2020](#)) so that the downside of deepfakes may be mitigated and the positives can be harnessed. To this end, we employ the social shaping of technology (SST) theory ([MacKenzie and Wajzman, 1999](#)) to develop a perspective on how individuals develop, process, and disseminate deepfakes along with governing interventions to mitigate the consequences of deepfakes. Based on prior literature, we refer to this continuum collectively as the deepfake engagement process ([Vasist and Krishnan, 2022a](#)). By conducting a meta-synthesis of 16 qualitative studies, we establish a comprehensive description of this deepfake engagement process.

Meta-synthesis refers to “an exploratory, inductive research design to synthesize primary qualitative case studies for the purpose of making contributions beyond those achieved in the original studies” ([Hoon, 2013, p. 523](#)). It aids in the development of a thorough comprehension of the phenomenon and adds to the formulation of a conceptual framework ([Lazazzara et al., 2020](#)). The current study leverages this approach to assess a selection of prior literature and present a perspective anchored on SST, in which we detail how the assertions and motivations enabled by digital platforms contribute to deepfake generation and dissemination, as well as how these fabrications manifest themselves in a variety of unintended consequences. Through this process, we offer vital insights to a variety of actors, including policymakers, platform players, administrators, and so on, in maneuvering through the complex terrain of epistemic uncertainty created by deepfakes. We contend that this analysis will serve as a robust platform for future scholarly research on deepfakes.

The rest of the paper is organized as follows. In [section 2](#), we discuss the concepts of deepfakes and meta-synthesis as well as the theoretical perspective of SST to set the context. In [section 3](#), we elaborate on the research design and build on the analysis in [section 4](#). In [section 5](#), we detail the nuances of deepfake engagement through the theoretical perspective of SST, and in [section 6](#), we discuss the implications of our work to theory and practice with a specific focus on information systems research alongside highlighting limitations and future research avenues. In conclusion, we reiterate the significance of the meta-synthesis conducted in this study.

## 2. Background

### 2.1 Deepfakes

Deepfake as a term can be traced back to 2017, with its origins associated with technological advancements that enabled face swapping in videos utilizing open-source machine learning (ML) technologies (Cole, 2017). Since then, researchers have defined deepfakes in several ways, referring to it as a “portmanteau of deep learning AI and faked imagery” (Wagner and Blewer, 2019, p. 33) and as a face-swapping technique (Afchar *et al.*, 2018) facilitated by AI (Yang *et al.*, 2018) with recent ones highlighting the deep learning techniques at play (Öhman, 2020). A broader and much synthesized definition refers to it as face-swaps or lip-synced content that is generated with the assistance of AI (Nguyen *et al.*, 2022) while recent studies extend the use of the term beyond individuals to other examples such as cartographic studies (e.g. Zhao *et al.*, 2021).

Deepfakes have mostly been researched for their deceptive capabilities. Deception as a factor is central to fake news as well and hence, we contend that it is important to compare and contrast the two phenomena in light of the current inquiry. Furthermore, this assumes importance in the current context wherein deepfakes are acknowledged as closely related to and discussed alongside fake news (Wu *et al.*, 2022). In this regard, while the current information environment powered by algorithms and automation tends to accelerate the propagation of intentionally created unauthentic information, it is precisely this characteristic that dictates the usage patterns of deepfakes as well (Kalpokas and Kalpokiene, 2022b) and also explains the choice of deepfakes as a medium for generating and spreading false content. Fake news predominantly focuses on the content of messages while “deepfakes create a simulation of the speaker” (Maddalena and Gili, 2020, p. 16), which exacerbate information destabilization with not just the content but also the source of the subject, which is fabricated in the process (Kalpokas and Kalpokiene, 2022c). This amplifies the complexity of fake news and makes it exceedingly indecipherable (Breen, 2021). Furthermore, image fabrication is not entirely new but the use of AI (Schick, 2020) to depict an individual in an act never committed (Kalpokas and Kalpokiene, 2022a) intensifies the fake news crisis (Whittaker *et al.*, 2021).

### 2.2 Social shaping of technology perspective

Social shaping of technology (SST) is a theoretical approach that dates back to 1985 (MacKenzie and Wajcman, 1999) and developed with the aim of comprehending technological change as a social process. The SST approach challenges technological determinism, rejecting its simple causality while recognizing that it has a partial truth in the idea that technology matters (MacKenzie and Wajcman, 1999). The intentional choice of the metaphor “shaping” as opposed to “social construction” underscores the materiality of technology (MacKenzie and Wajcman, 1999). SST scholars contend that the interplay of social forces and the materiality of technical objects merit equal consideration (Howcroft and Taylor, 2022). From the SST perspective, technologies reflect specific forms of power and authority and influence how people work, communicate and consume over time (Winner, 1980).

---

Social science has predominantly witnessed an “artificial gulf” (Williams and Edge, 1996, p. 893) between the social and technical, with the social aspects considered as externalities and the technical aspects regarded as tools and instruments. However, SST is based on the premise that technology and society are mutually inclusive, rather than discrete entities (MacKenzie and Wajcman, 1999), and that each influences the other in interconnected and multifaceted ways as opposed to linear pathways. According to the perspective of SST, social and technical are mutually shaped (MacKenzie and Wajcman, 1999), and each evolves in conjunction with the other. SST shows that technologies are social products, and every stage of evolution of new technologies is contingent upon a range of social factors that influence the nature of the technologies and their ramifications (Williams and Edge, 1996).

SST offers a novel perspective in technological areas where visions are diverse, societal interests are at odds, and the applications and market for the emerging technology are still in a nascent phase (Jørgensen *et al.*, 2009). In particular, the strand of technological frames within the setting of SST underscores the centrality of interpretations of technology for comprehending how technologies are developed, used, and changed (Orlikowski and Gash, 1994). A key implication of SST is that the specifics of a technology’s usage are often not well known at the time of its design and fabrication, rendering it particularly relevant in the context of emerging technologies (Gasser, 1986; Sobreperez, 2012). Such is the nature of deepfakes as a technology, whose prospective use cases may not have been fully understood, yet its malevolent usage has accelerated at an astounding pace. As the potential of deep learning in general and generative adversarial networks in particular, which power deepfakes, becomes increasingly apparent, and as the mainstream adoption of deepfakes increases, we anticipate that deepfakes will be socially shaped to reflect the circumstances of their development and that new pathways will emerge for the phenomenon.

In conclusion, we present examples from prior research to illustrate how the theory has been applied in various contexts to explain the uptake and usage of technologies, as well as highlight how the theory has been utilized in the current study to offer insights on the evolution of deepfakes as a socio-technical phenomenon (see Table 1).

### 2.3 Meta-synthesis

Meta-synthesis as an approach helps in drawing interpretations from a selection of qualitative studies (Major and Savin-Baden, 2012) by providing a means of accumulating wisdom from extant research which forms the foundations of science (Hunt, 1997; Hunter *et al.*, 1982). This approach, depending on how the synthesis is carried out, may be classified as aggregation, interpretation, and translation synthesis (Hoon, 2013). Among the first two approaches, namely aggregation and interpretation analysis, the former is rooted in the positivist paradigm and focused on generating a prediction (Rousseau *et al.*, 2008), while the latter offers an interpretive explanation through an amalgamation of a selection of qualitative studies and develops a model to explicate the findings (Walsh and Downe, 2005). As a result, the synthesis expounds on the patterns within the corpus of studies while preserving the integrity of the studies (Hoon, 2013). The last variant, translation synthesis, is directed towards reconstructing “how the study’s participants constructed their own understandings” of a phenomenon in focus, with the constructed entities serving as data for the synthesis (Hoon, 2013, p. 526). This variant has received considerable interest among scholars in several disciplines (Tranfield *et al.*, 2003). Our study leverages this variant to expound on deepfakes as a phenomenon while the integrity of individual studies comprising the selection is not hampered.

Study	Application of SST
Williams and Edge (1996)	Reviews the literature that explores the social shaping of technology by examining the content of technology and innovation processes
Rohracher (2003)	Highlights the role of a diverse set of actors in the social shaping of two technologies, namely, balanced ventilation systems and smart home technologies, aimed to enhance buildings' environmental performance
Dutton <i>et al.</i> (2004)	Identifies constraints on innovation and changes likely to evolve from the diffusion of virtual learning environments through perspectives from SST
Park and McDowell (2005)	Compares the shaping of a new media technology in two countries and through SST, highlights the relationship between industry, policy, audiences, and the technology
Wilson and Howcroft (2005)	Analyzes a case study involving the success and eventual failure of a nursing information system through SST by leveraging the concept of relevant social groups
Agalianos <i>et al.</i> (2006)	Utilizes SST perspectives to explore the emergence and evolution of Logo, an educational philosophy and an evolving family of computer languages that facilitate its implementation
Howcroft and Light (2010)	Examines the packaged software selection process among organizations and leverages SST to demonstrate how technologies are socially constituted
Bibri and Krogstie (2017)	Uses SST to elaborate on how smart and sustainable cities are shaped by and shape institutional practices and structures
Chang (2019)	Uses SST to explore the usage of robots in organizations and underscores the need for studying human-robot interactions to unravel complex factors driving use and societal perceptions of robotics
Frith and Wilken (2019)	Uses the locative media-startup failures as a point of departure and examines the significance of location-based social networks and reconfigurations made by involved players to the services through SST
Sharma (2020)	Leverages SST to investigate the social shaping of solar microgrid in rural India and the implications of the narrow framing of the socio-technical ensemble by several involved organizations
Muza and Debnath (2021)	Utilizes SST to explore the factors influencing the uptake of electrical appliances in rural Rwanda, as well as the factors contributing to the failure of off-grid renewable energy alternatives
Stroud <i>et al.</i> (2020)	Examines the role of digitalized drones in the manufacturing industry, discusses their potential, and critiques unduly deterministic narratives through the lens of SST
Bibri (2022)	Underscores the need for mutual evolution of technology and society and discusses the social shaping of Metaverse and the trajectory of its innovation processes
Chou (2022)	Utilizes SST to analyze the impact of various social factors in the diffusion of quick response (QR) code payment technology and its commercialization through the case study of related technology from Alipay
Howcroft and Taylor (2022)	Uses SST to investigate what shapes automation and its expected effects along with its implications for the future of work
Ticau and Hadad (2022)	Leverages SST to highlight the dual nature of the relationship between activity trackers and user attitudes, as well as the positive behavioral consequences and difficult-to-control mental states that follow from this interaction
This study	Utilizes SST to develop a comprehensive perspective on the evolving nature of deepfakes. This perspective offers insights on the socio-technical dynamics of assertions and motivations enabled by digital platforms that contribute to deepfake generation and dissemination, their unintended consequences, and governing interventions aimed at mitigating these ramifications. The theoretical foundation of SST sheds light on how the underlying technology that powers deepfakes socially evolves through continual knowledge contribution and feedback-led refinement of assertions, and the meta-synthesis helps progress beyond the individual studies which rarely discussed all these aspects in conjunction with one another

**Table 1.** Illustrations from prior literature applying the theoretical lens of SST

---

### 3. Research design

The research design for our study is carried out in three steps which are elaborated below.

#### 3.1 *Outlining the research question*

A methodical approach to any research endeavor commences with the formulation of an appropriate research question. In the spirit of carrying out this meta-synthesis, we seek to understand how individuals engage with deepfakes throughout the process of developing, viewing, and sharing them, as well as how dissemination and ramifications of deepfakes may be controlled by means of appropriate interventions. In accordance with this goal, the research questions (RQ) guiding this study have been framed as follows:

- RQ1. How do individuals engage with deepfakes in terms of creating, viewing, and disseminating them?
- RQ2. How can dissemination and ramifications of deepfakes be restricted through suitable interventions?

#### 3.2 *Search protocol for locating relevant research*

Identifying relevant studies in the extant literature is a crucial step prior to commencing a meta-synthesis. Scholarly attention on deepfakes has spread across several disciplines, including computer science (Güera and Delp, 2018) and the domain of social sciences with deepfake-related consequences widely discussed in areas such as journalism (Yadlin-Segal and Oppenheim, 2021), politics (Dobber *et al.*, 2021) and ethics (Öhman, 2020). This multidisciplinary nature of deepfake literature led us to execute our search protocol across key databases focused on multidisciplinary research, namely Scopus, Web of Science (WoS), ProQuest, and Ebscohost, which have been leveraged in previous literature reviews, including meta-syntheses (e.g. de Beer and Matthee, 2021; Kaur *et al.*, 2021; Khan and Krishnan, 2021; Riar *et al.*, 2022; Soral *et al.*, 2022; Tandon *et al.*, 2021; Viglia *et al.*, 2022). Guided by our research question to identify the nuances of the process of deepfake engagement, we commenced our meta-synthesis with a combination of keywords that were driven by prior literature reviews on deepfakes (e.g. Carvajal and Iliadis, 2020; Westerlund, 2019). Deepfake as a term is a combination of “deep learning” and “fake” (Mirsky and Lee, 2021), and hence, these were included as additional keywords in the mix. Lastly, deepfakes are also referred to as synthetic media (Leibowicz *et al.*, 2021; Whitaker, 2021), and hence, “synthetic media” was included as an additional search term in the list of keywords.

#### 3.3 *Determining the inclusion and exclusion criteria*

Prior research cautions against selecting so many primary studies that the researchers cannot conduct a deep and extensive study of the findings of each case (Sandelowski and Barroso, 2007). Further, a meta-synthesis necessitates synthesizing the rich insights from each case; thus, a manageable number of studies must be selected to preserve sensitivity towards the analysis and synthesis of the findings of other researchers [1] (Hoon, 2013). Hence, in accordance with prior studies (e.g. Khan and Krishnan, 2021; Lazazzara *et al.*, 2020), we formulated a set of inclusion and exclusion criteria to ensure the corpus of studies was appropriate and representative of the literature on deepfakes. In addition, we included in our corpus research studies employing a variety of qualitative research approaches. Whilst the conventional approach of meta-synthesis was originally established for case studies (Hoon, 2013), it has been expanded to encompass a range of qualitative method-based investigations, including interviews, ethnography, analysis of discussion threads, and social media posts among others (Allen *et al.*, 2016; Athanopoulou and Dopson, 2018; Lazazzara *et al.*, 2020;

---

Soral *et al.*, 2022). In this regard, meta-synthesis entails extracting codes from the results and discussion sections of selected research (Hoon, 2013); hence, it remains unaffected by the qualitative method of inquiry employed in individual studies. Additionally, in line with recent meta-syntheses (e.g. Allen *et al.*, 2016; Xie *et al.*, 2021), we also incorporated qualitative studies assessing user-generated content (UGC) from virtual communities and social media sites. Lastly, with the motive of garnering practical insights on the phenomenon, grey literature was included as part of the review (Adams *et al.*, 2017).

In line with methodological guidance from past literature (e.g. Atkins *et al.*, 2008; Horton, 2020) and the discussion thus far in this section, our criteria for inclusion were (1) studies with a qualitative research design and employing qualitative methods (fully or partially) (2) studies including and interpreting data from interviews, observations, ethnographic approaches, focus group discussions or UGC from discussions on online forums or social media sites (such as Facebook, Twitter, Reddit and so on) (3) studies with deepfakes as the central theme of research (4) studies with clear description and details on sampling method and data collection process. In a similar vein, the criteria for exclusion were (1) studies employing only quantitative research designs (2) technical articles on deepfakes (3) studies positioned as a review (4) studies portraying only a conceptual or theoretical framework.

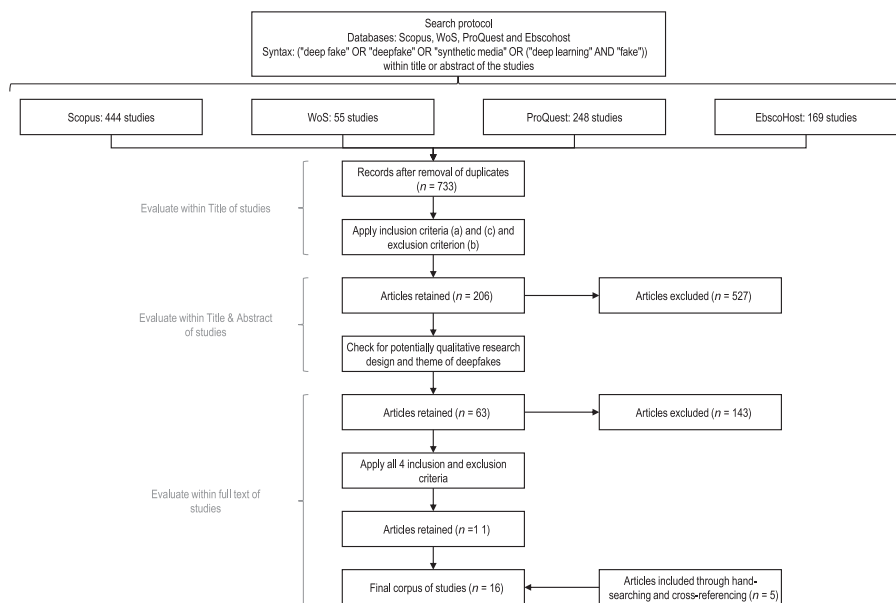
Guided by the research question and search protocol, the preliminary search yielded 444, 55, 248, and 169 studies from Scopus, WoS, ProQuest, and Ebscohost, respectively. The studies were compared across the four databases to yield a final count of 733 unique studies. Both authors independently analyzed these 733 articles. In line with identifying studies relevant to our study objective, we anchored our next step of analysis on the titles of these 733 articles to gain a basic idea of each study and applied the inclusion criteria (1) and (3) and exclusion criterion (2) to distinguish 206 studies. If the research design or the theme of the article could not be determined at this stage or if any disagreement arose, the article was retained for the next step of evaluation.

We then looked into the abstract of the 206 studies to further evaluate their relevance in terms of the possibility of qualitative research design and situated in the context of deepfakes. Not only did we include studies that appeared pertinent, but we also included studies with abstracts that lacked a clearly articulated research design and theme. Moreover, as with the previous step, an article was retained for the next stage of full-text analysis in the event of disagreement between the authors or a lack of clarity around the research design or theme. Through this, 63 studies were identified for the final round of screening in which all the inclusion and exclusion criteria were applied, resulting in 11 studies deemed most suitable for the purpose of the meta-synthesis. Furthermore, in accordance with prior literature (e.g. Lazazzara *et al.*, 2020), each of the 63 articles was reread in its entirety to confirm the robustness of the evaluation and avoid the exclusion of any relevant studies. As a final step, to ensure no crucial literature was excluded, we sourced additional literature by hand-searching, cross-referencing, and including relevant studies from authoritative sources. This resulted in the identification of 5 articles that were aligned with our review's purpose and hence, included in the final corpus of 16 articles. The process is detailed in Figure 1 and Table A1 of Appendix.

## 4. Analysis

### 4.1 Synopsis of reviewed studies

Our review of qualitative deepfake literature reveals a rising trend of publications featuring the phenomenon over the past four years. This is apparent from the emergence of deepfake as a term in 2017 (Cole, 2017). Furthermore, the reviewed articles contain several references to the use of deepfakes to create sexually explicit or pornographic content (e.g. Newton and Stanfill, 2020; Popova, 2019). The creation of such content involves the digital replacement of the face of one person with that of another who is involved in a sexual act. In this article,



**Figure 1.**  
Flow diagram of study  
selection

we refer to these phrases “sexually explicit content” and “pornographic content” as content of an explicit nature. In one of the first qualitative studies on deepfakes, [Fikse \(2018\)](#) employed digital ethnographic methods and participated in online discussion forums, such as the deepfakes club forum, with the aim of identifying harmful usage patterns of deepfake technology and associated socio-technical consequences. In doing so, the author acknowledges the high levels of expertise required to create deepfakes, as well as the fears and concerns that accompany it as it ushers in a post-truth era. However, technology has certainly advanced since then, as evidenced by the abundance of online fabrication methods and the ease with which deepfakes can now be produced ([Gandhi, 2021](#)). In a subsequent study, [Popova \(2019\)](#) examined celebrity deepfake engagement through yet another digital ethnographic approach centered on two deepfake sites. Through the study, the author demonstrated low levels of concern among communities for the private individual behind the deepfake and noted attempts by communities to restrict the content within the community as opposed to disseminating it to a larger audience ([Popova, 2019](#)). In this regard, it should be noted that celebrity deepfakes, particularly those of Hollywood stars, were gaining popularity around this time ([Hern, 2018a](#)), and the social community site Reddit witnessed a significant increase in deepfake community members, exacerbating concerns and eventually leading to Reddit banning the nearly hundred thousand strong community that indulged in celebrity face-swapped deepfakes of an explicit nature ([Farokhmanesh, 2018](#); [Hern, 2018a, b](#)).

In the ensuing years, the phenomenon began to attract the attention of scholarly communities as concerns escalated and community bans followed. A 2019 report on the state of deepfakes indicated that women were the exclusive target of deepfakes ([Sensity, 2019](#); [Wang, 2019](#)). In accordance with this, [Newton and Stanfill \(2020\)](#) shed light on the gendered discourse in deepfake communities through a thematic analysis of deepfake discussions on two GitHub communities and emphasize that while open source platforms are a platform for various technological tools, they also are “a site of toxic geek masculinity” ([Newton and Stanfill, 2020, p. 2](#)). In a similar vein, [Winter and Salter \(2020\)](#) traced the inexorable spread of deepfakes despite regulatory efforts and the concomitant surge in non-consensual deepfaked content.



Amidst the rising sophistication of deepfakes and debates around the ease or challenge in spotting them (e.g. Barber, 2019; Sample, 2020), Thaw *et al.* (2020) conducted a study with students and working professionals and found that the majority of them are unable to detect deepfakes. Additionally, they offered insights on features utilized in evaluating deepfakes that may be considered for use in improving deepfake detection algorithms (Thaw *et al.*, 2020).

In 2021 and 2022 till date, the scholarly contribution to deepfake research has undoubtedly been at its peak, with over two-thirds of reviewed studies published during these years. In this context, Ali *et al.* (2021) highlight the role of learning interventions by introducing a group of middle and high school students in the United States and demonstrate the efficacy of the interventions that enable the students to recognize deepfakes and misinformation in general. While it has been widely accepted in prior literature that people with higher levels of media literacy are better able to identify fake news online (e.g. Cheng and Chen, 2020; Jang and Kim, 2018; Mihailidis and Viotty, 2017), recent research has questioned the generalizability of media literacy by suggesting that some types of literacies, such as information literacy, may be advantageous in comparison to others in identifying and debunking fake news online (Jones-Jang *et al.*, 2021). In this regard, additional research on the intricacies of media literacy and its impact on deepfake detection may be worthwhile to take into consideration. In among the first few studies to focus on deepfakes designed not with malicious intent but for entertainment at best with monetization opportunities, Bode (2021) analyzes a deepfaked video of Keanu Reeves created by Corridor, a YouTube channel, and the role of the networked public in shaping the discourse surrounding the video and its popularity. In a departure from deepfake analysis predominantly centered on western nations, de Seta (2021) discusses deepfakes in China through the Zao app and communities of practice contributing to deepfake fabrication in the region. In this regard, the author draws attention to the impact of the Chinese technology industry on the development of deepfakes and related issues (de Seta, 2021). The study is significant in light of China's strict regulation of information (Guo, 2020), the Zao app's growing popularity and privacy concerns (Doffman, 2019), and China's efforts to combat deepfakes (Reuters, 2019).

Among other studies, Flynn *et al.* (2021) conducted semi-structured interviews with image-based sexual abuse victim-survivors and associated stakeholders to analyze deepfakes victims and the abuse they experience. While earlier works had highlighted the negative effects of deepfakes, this was one of the first to include a first-person description of victims' experiences. Amidst rising concerns over a post-truth era ushered in by deepfakes, Gregory (2022) highlights the efforts of WITNESS, a human rights non-profit organization in mitigating misinformation, restoring truth and curbing the malicious use of deepfakes, including the role of authenticity infrastructure in the process through a case study drawing upon the author's own experience and that of the author's colleagues in the organization. As deepfakes' technological prowess continued to expand, Ghazi-Tehrani and Pontell (2021) explored the potential for deepfake-led phishing attacks, which may be extremely challenging to decipher.

The gendered discourse dominating deepfakes continues to dominate discussions in 2022, with Elmaa (2022) using a thematic analysis of user comments on Reddit to highlight concerns regarding the gendered narrative and lingering ambiguities surrounding computer-generated explicit content, particularly of children. Despite efforts from platforms such as Reddit to ban deepfake communities over the years (e.g. Anderson, 2022; Hern, 2018b), Gamage *et al.* (2022) reveal the pro-deepfake nature of Reddit conversations and the support for the creation and transmission of deepfake artifacts through a thematic analysis of posts and comments on Reddit. Deepfakes are increasingly turning into a nightmare for journalists who are beginning to consider ways to stop their spread (Templeton-Knight, 2021). In this regard, Himma-Kadakas and Ojaments (2022), through semi-structured interviews with a group of journalists in Estonia, contend that deepfakes are among the most challenging media when it comes to detecting them. Sabrina (2022), through interviews with university

---

students in South Korea, highlights concerns about the erosion of public trust and accompanying fears, as well as the alleviation of certain issues due to the emergence of deepfake detecting technologies. [Widder et al. \(2022\)](#), through semi-structured interviews with professionals, examine an essential aspect of deepfakes in the context of ethics in open-source AI communities and the beliefs of community users. In doing so, they contribute to past research, such as that of [Winter and Salter \(2020\)](#), who explore the ethical stance of open source platforms in controlling deepfake content.

A summary of the 16 studies reviewed as part of the meta-synthesis and discussed in this section is detailed out with a brief description of each study, the research method and sample employed for the study, the country of research project based on participant profile or the first author affiliation in case of online community discussions where the location could not be ascertained and key findings of each study (see [Table A2 of Appendix](#)).

#### *4.2 Extracting and coding data for insights*

The first step in conducting a meta-synthesis analysis is to undertake study-specific analysis ([Hoon, 2013](#)), in line with extracting and categorizing evidence captured from each study ([Noblit and Hare, 1988](#)). We analyzed and coded the texts in order to gain insights into the processes underlying the overall deepfake engagement process and its implications, as our objective was to examine the dispositions for individuals to engage in deepfake fabrication and dissemination, as well as how the phenomenon affects our worldview and associated choices. In doing so, we classified the insights gleaned from the studies by focusing primarily on the findings and discussion within each study ([Hoon, 2013](#)). Following such a methodology allows researchers to capture the original substance of the studies and avoid any interpretive bias that could jeopardize the purpose of the meta-synthesis.

In accordance with our research questions, a preliminary coding form was constructed, and four studies were chosen at random to find pertinent items deserving of inclusion in the coding form basis which the form was refined. Utilizing the improved coding form, the remaining articles were coded by both authors individually, as using two coders eliminates errors in recording data and prevents the omission of constructs that are relevant to the study ([Miles and Huberman, 1994](#)). The authors employed an iterative approach ([Locke, 2001](#)), alternating between the emerging insights from the studies and the codes. In the event of discrepancies, we engaged in conversation and addressed any differences in the coding process after reaching consensus. Following the instructions of [Mackey and Gass \(2015\)](#), one of the authors also coded the articles across two different time periods to help establish intra-rater reliability. Then, we began to establish connections between the codes and construct more generalized themes. In accordance with recommendations from ([Hoon, 2013](#)), we recognized commonalities between concepts and organized them into theoretical categories to uncover a set of second-order themes. As we sought causative sequences of variables at the level of individual studies, we constructed study-specific causal network sequences, therefore establishing the linkages. For instance, one of the causal linkages established through this process revealed that “non-malevolent motivations” such as an intent directed towards monetary gains and “pro-deepfake assertions” such as a form of celebrity engagement, supported by the “marketplace ecosystem” with platforms such as YouTube offering monetization opportunities lead to “innocuous fabrications” such as deepfaked videos of Keanu Reeves and may also result in “inadvertent sharing” for its entertainment factor. This process of establishing study-specific causal mechanisms served as the basis for the conceptual framework which we ultimately propose through the analysis.

Next, we moved from the level of individual studies to the level of cross-study analysis. Seeking guidance from prior literature (e.g. [Khan and Krishnan, 2021](#); [Lazazzara et al., 2020](#)), we constructed a within-case processual matrix, within which the rows and columns reflected individual studies and categories determined from the primary coding stage, respectively. This

---

allowed us to determine the associations across categories throughout the studies. As a consequence of this step, the themes were refined further, resulting in a collection of improvised second-order themes and causative networks at the study level. This refined set of second-order themes was combined to form a set of aggregate dimensions, which aided the development of the conceptual framework (see [Table 2](#)).

In accordance with recommendations from [Miles and Huberman \(1994\)](#), we also conducted a cross-study comparison to determine how associations developed over the complete array of studies. As an example, when the causal networks from the studies by [Widder et al. \(2022\)](#) and [Gamage et al. \(2022\)](#) were compared, it revealed similarities in terms of digital spaces offering the necessary “technological expertise” to engage in deepfakes which could be “malicious fabrications.” However, the causal networks in [Widder et al. \(2022\)](#) revealed the possibility of “innocuous fabrications” as well, such as satirical deepfakes, whereas the analysis of the causal mechanisms in the study by [Gamage et al. \(2022\)](#) revealed the possibility of “networked spaces” like Reddit providing advice around legal clearances in deepfake creation as well as a repository of banned deepfakes through a “marketplace ecosystem” and a space to incubate deepfake creators and hone their skills towards deepfake creation through “technology expertise”. By engaging in similar activities, we were able to find new factors and establish interrelationships, thereby influencing the conceptual connections which emerged through a framework (see [Figure 2](#)) connecting deepfake-related assertions, motivations, the nuances of digital platforms, the dissemination patterns, and related consequences. In the following section, we elaborate on each of the components of our conceptual framework and examine the relationships.

## 5. Conceptual framework of deepfake engagement

In this section, we first draw upon the theoretical perspective from SST to develop the conceptual framework of deepfake engagement and then explain each of the components of the framework, which is schematically represented in [Figure 2](#).

### 5.1 Theoretical underpinnings

We utilize the theoretical perspective from SST to explain our framework. As noted previously, the central premise of SST is that the social context in which a particular technology is created has a profound impact on its development ([MacKenzie and Wajcman, 1999](#)). SST provides a novel perspective in the context of technologies where visions are diverse, societal interests are at odds, and the applications and market for emerging technology are still evolving ([Jørgensen et al., 2009](#)). We suggest that this is the case with deepfakes and provide supporting evidence. We contend that the social environment of the origins of deepfakes has shaped the dominant use cases of the technology to date. While deepfakes are a category of AI-generated synthetic media ([Whittaker et al., 2020](#)), the term has a negative connotation since its early association with technological advancements that enabled face swapping in pornographic videos using open-source ML tools ([Cole, 2017](#); [Somers, 2020](#)). Despite the fact that the technology associated with making deepfakes has the potential for constructive applications, they have been largely perceived negatively and are known for their dominant use in political or explicit content ([Chandler, 2020](#)). Experts at large believe that the word “fake” in deepfake is often associated with unlawful and fraudulent acts, thereby lending it a negative connotation and indicating malicious intent ([Huijstee et al., 2021](#)). While this connotation does not hinder the growing number of positive use cases of deepfake technology, the literature is still replete with negative use cases in contrast to the relatively few explorations of positive use cases. However, deepfake as a technology is still evolving, and we anticipate it will be influenced by the social context in which it evolves. In

**Table 2.**  
Deepfake engagement  
process variables

Aggregate dimension	Second-order themes	First-order categories
Motivations	Malevolent motivations	Engaging in political smear campaigns (Sabrina, 2022); the need for control and power (Flynn <i>et al.</i> , 2021); indulging in acts of revenge (Fikse, 2018; Flynn <i>et al.</i> , 2021; Gamage <i>et al.</i> , 2022; Popova, 2019; Winter and Salter, 2020); oriented for cyberbullying (Sabrina, 2022); blackmailing (Sabrina, 2022); carrying out phishing attacks (Ghazi-Tehrani and Pontell, 2021)
	Non-malevolent motivations	Hobby to pursue realism (Widder <i>et al.</i> , 2022); recreational tool (de Seta, 2021); monetization (Bode, 2021; Gamage <i>et al.</i> , 2022); displaying technical expertise (Newton and Stanfill, 2020); demonstration and educational purposes (Fikse, 2018; Widder <i>et al.</i> , 2022)
Dichotomous assertions	Pro-deepfake assertions	Technological inevitability (Widder <i>et al.</i> , 2022); a medium for artistic expression (Ali <i>et al.</i> , 2021; Bode, 2021; de Seta, 2021; Widder <i>et al.</i> , 2022); a form of celebrity engagement (Bode, 2021; Popova, 2019); entertaining interaction (de Seta, 2021); hobby (Bode, 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020); technical expertise (Newton and Stanfill, 2020)
	Anti-deepfake assertions	Concerns around the use of deepfakes to generate explicit content and the normalization of such activity (Eelmaa, 2022; Fikse, 2018; Flynn <i>et al.</i> , 2021; Gamage <i>et al.</i> , 2022; Newton and Stanfill, 2020; de Seta, 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020); gendered discourse and associated violence (Eelmaa, 2022; Flynn <i>et al.</i> , 2021; Gregory, 2022; Newton and Stanfill, 2020; Widder <i>et al.</i> , 2022; Winter and Salter, 2020); a form of cybercrime (Ali <i>et al.</i> , 2021; Eelmaa, 2022; Flynn <i>et al.</i> , 2021; Ghazi-Tehrani and Pontell, 2021); undermining of societal values (Eelmaa, 2022); paucity of ethics and morals (Ali <i>et al.</i> , 2021; Bode, 2021; Gamage <i>et al.</i> , 2022; Newton and Stanfill, 2020; de Seta, 2021; Widder <i>et al.</i> , 2022)
Digital platforms as enablers	Technological expertise	Learning opportunity for developers (Newton and Stanfill, 2020); Reddit as a safe setting for learning and optimizing deepfake generation skills (Gamage <i>et al.</i> , 2022); GitHub as an AI learning enabler (Widder <i>et al.</i> , 2022)
	Networked spaces	Platforms as a means of distributed and collaborative technology production (Fikse, 2018; Newton and Stanfill, 2020; Winter and Salter, 2021); initiation to new tools and techniques (Gamage <i>et al.</i> , 2022; de Seta, 2021); user feedback towards improvisation (Gamage <i>et al.</i> , 2022; Popova, 2019); indexing of videos by the network (Bode, 2021); advice concerning legal clearances (Gamage <i>et al.</i> , 2022)
	Marketplace ecosystem	Archive repositories for banned deepfakes (Fikse, 2018; Gamage <i>et al.</i> , 2022); space for deepfake monetization (Bode, 2021; Gamage <i>et al.</i> , 2022)

(continued)

Aggregate dimension	Second-order themes	First-order categories
Deepfake genres	Malicious fabrications	Political deepfakes (Ali <i>et al.</i> , 2021; Gamage <i>et al.</i> , 2022; Ghazi-Tehrani and Pontell, 2021; Gregory, 2022; Sabrina, 2022; Widder <i>et al.</i> , 2022); deepfakes of an explicit nature (Eelmaa, 2022; Fikse, 2018; Flynn <i>et al.</i> , 2021; Gamage <i>et al.</i> , 2022; Newton and Stanfill, 2020; Popova, 2019; de Seta, 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020); AI-generated fake reviews (Gamage <i>et al.</i> , 2022)
	Innocuous fabrications	Satirical creations (Fikse, 2018; Widder <i>et al.</i> , 2022); positive celebrity engagement (Bode, 2021); revenue generation through advertising and downstream effects (Bode, 2021); educational purposes (Ali <i>et al.</i> , 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020)
Dissemination	Intentional sharing	Inflicting political harm (Gamage <i>et al.</i> , 2022; Sabrina, 2022); Distorting facts (Gregory, 2022); acts of fury and revenge (Flynn <i>et al.</i> , 2021); gratifying desires (Eelmaa, 2022)
	Inadvertent sharing	Inability to recognize deepfakes (Ali <i>et al.</i> , 2021; Thaw <i>et al.</i> , 2020); sharing for the purposes of recreation or entertainment (Bode, 2021; de Seta, 2021); broadcasting under time constraints and a sense of competition (Himma-Kadakas and Ojaments, 2022)
Consequences	Societal	Influence on public opinion and electoral discourse (Fikse, 2018; Gamage <i>et al.</i> , 2022); distortion of social reality (Sabrina, 2022); normalization and desensitization of illegal acts of explicit nature (Eelmaa, 2022; Gamage <i>et al.</i> , 2022); destruction of societal values (Eelmaa, 2022; Gamage <i>et al.</i> , 2022); erosion of ethics and morality (Ali <i>et al.</i> , 2021; Eelmaa, 2022; Gamage <i>et al.</i> , 2022; Ghazi-Tehrani and Pontell, 2021; Newton and Stanfill, 2020; de Seta, 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020)
	Organizational	Targets of spear-phishing (Ghazi-Tehrani and Pontell, 2021); posting of fake reviews (Gamage <i>et al.</i> , 2022); journalists' loss of trust in information (Himma-Kadakas and Ojaments, 2022); loss of credibility for journalist organizations (Himma-Kadakas and Ojaments, 2022)
	Individual	Victimization of deepfake targets and related violence (Eelmaa, 2022; Flynn <i>et al.</i> , 2021); cyberbullying and blackmail (Sabrina, 2022); distrust and skepticism about media (Fikse, 2018; Sabrina, 2022; Widder <i>et al.</i> , 2022)

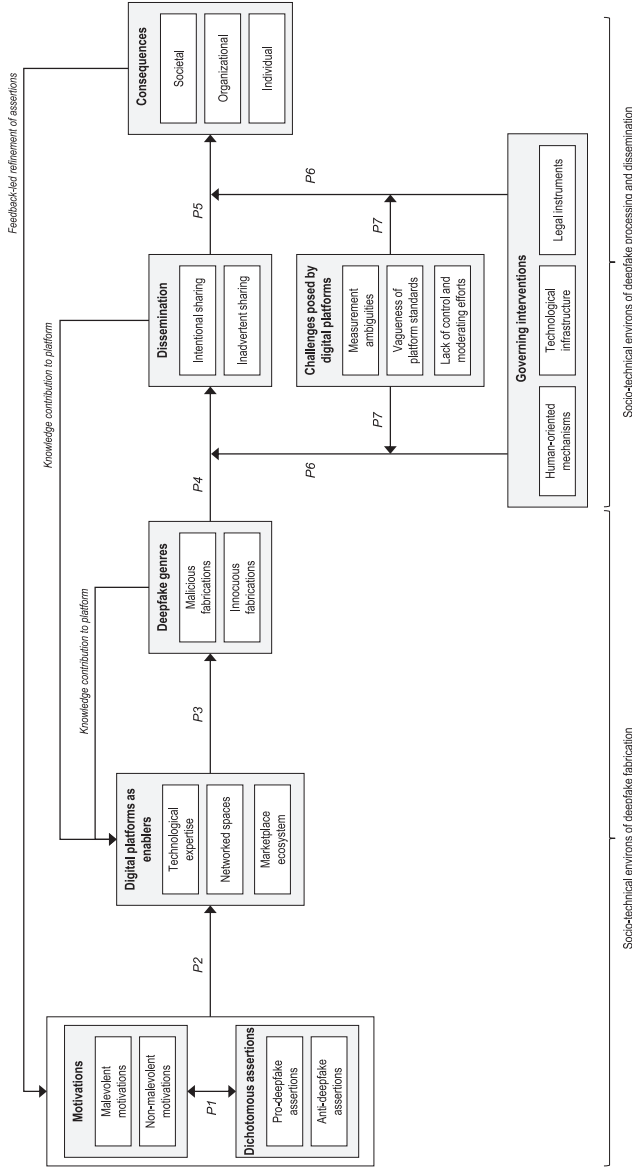
(continued)

Engaging with deepfakes

Table 2.

Table 2.

Aggregate dimension	Second-order themes	First-order categories
Governing interventions	Human directed mechanisms	Increasing media literacy through campaigns, seminars and so on (Ali <i>et al.</i> , 2021; Bode, 2021; Fikse, 2018; Gregory, 2022; Himma-Kadakas and Ojameets, 2022); enhancing critical thinking (Ali <i>et al.</i> , 2021; Gregory, 2022; Himma-Kadakas and Ojameets, 2022; Sabrina, 2022); inoculating through proliferation of deepfakes (Widder <i>et al.</i> , 2022); human-led verification techniques leveraging media features (Thaw <i>et al.</i> , 2020); verification subsidies and collaborative fact-checking mechanisms (Gregory, 2022; Himma-Kadakas and Ojameets, 2022)
	Technological infrastructure	Forensic analysis through image verification tools (Gregory, 2022; Himma-Kadakas and Ojameets, 2022); multi-factor authentication to avoid technology misuse (Ghazi-Tehrani and Pontell, 2021); watermarking (Popova, 2019; de Seta, 2021; Widder <i>et al.</i> , 2022); community efforts to contain material within platforms (Popova, 2019); platform rules and standards including ban on deepfakes (Ali <i>et al.</i> , 2021; Gamage <i>et al.</i> , 2022; de Seta, 2021; Widder <i>et al.</i> , 2022; Winter and Salter, 2020)
Challenges posed by digital platforms	Legal mechanisms	Government regulations (Ali <i>et al.</i> , 2021; Ghazi-Tehrani and Pontell, 2021; Gregory, 2022; Winter and Salter, 2020)
	Measurement ambiguities	Lack of standards to designate deepfakes as art or violation (Eelmaa, 2022); legalistic rather than ethical framing (Newton and Stanfill, 2020); neutrality of technology used to create deepfakes (Widder <i>et al.</i> , 2022); ambiguity over computer-generated explicit content (Eelmaa, 2022)
	Vagueness of platform standards	Disproportionate control over deepfake policies (Gamage <i>et al.</i> , 2022); disregard for upholding community values (Winter and Salter, 2020)
	Lack of control and moderation	Indifference towards harms of deepfakes (Newton and Stanfill, 2020); lack of ethics around seeking consent (Newton and Stanfill, 2020; Winter and Salter, 2020); appreciation for explicit-natured deepfakes and intent towards continuance (Winter and Salter, 2020); use of pseudonymous accounts to engage in deepfake contribution (Winter and Salter, 2020); ineffectiveness of auto-moderation practices in content removal and absence of dedicated moderators (Winter and Salter, 2020); emergence of forking techniques and splinter communities (Widder <i>et al.</i> , 2022)



**Note(s):**

- P1:* Pro- and anti-deepfake assertions interact with deepfake fabrication related motivations.
  - P2:* Interactions between motivations and assertions are backed by digital platforms as enablers.
  - P3:* Digital platforms facilitate interactions between deepfake motivations and assertions, which result in a myriad of deepfake fabrications.
  - P4:* Deepfake genres are either deliberately or inadvertently disseminated.
  - P5:* Deepfake dissemination and consumption leads to a variety of consequences.
  - P6:* A variety of regulatory measures may limit the effects of deepfakes, including their proliferation.
  - P7:* Platform related challenges mitigate the influence of governing interventions.
- Knowledge contribution to platform:** Deepfakes created and distributed contribute to the platform's knowledge base through user feedback, dissemination, and an archive of deepfakes that serves as a repository for future fabrications.
- Feedback-led refinement of assertions:** Consequences of deepfakes could provide feedback for reconsidering prevalent assertions about the phenomenon

**Figure 2.**  
Conceptual view of deepfake engagement

---

this regard, our framework acknowledges the dichotomous structure of arguments for and against deepfakes, while platform-level enablers such as collaboration, space for learning, and so on contribute to the development of malicious and non-malicious deepfakes. This social context comprises a variety of assertions, motivations, and platform-level enablers that influence deepfake technology and its further spread while a range of regulatory interventions may limit the effects of deepfakes, including their proliferation. Lastly, feedback loops not only help enhance platform-level contributions to deepfake creation, but feedback from consequences experienced across various levels could also refine the prevalent opinions and assertions regarding deepfakes in society.

We now elaborate on the preceding discussion by referencing the framework in [Figure 2](#). As indicated by the arrow P1, arguments for and against deepfakes interact with deepfake fabrication related motivations. These interactions are backed by digital platforms as enablers (see arrow P2) and lead to the fabrication of a variety of deepfakes, as indicated by arrow P3. These deepfake genres are disseminated, either deliberately or inadvertently, as indicated by arrow P4, and inflict consequences at various levels, as depicted by arrow P5. However, a range of regulatory measures, such as the improvement of media literacy and governance procedures on technological platforms, may mitigate the effects of the phenomenon, including the spread of deepfakes; we indicate this moderating effect with the arrow P6. However, the persisting ambiguities on these platforms, such as the vagueness of platform standards, may render these regulatory mechanisms ineffectual and are anticipated to mitigate their influence, as indicated by arrow P7. Lastly, the variety of deepfakes created and disseminated contributes knowledge back to the platform through user feedback, dissemination, and an archive of deepfakes that serves as a repository for future fabrications, while the consequences of deepfakes may serve as feedback for refining the prevalent assertions surrounding the phenomenon. In this section, we elaborate on each of the framework's components by employing references from the reviewed literature.

### *5.2 Motivations for deepfake fabrication*

Although the term “deepfake” conveys a negative connotation, a growing number of positive applications are being developed by employing the technology enabling deepfakes ([Somers, 2020](#)). However, a significant portion of deepfake technology manifests online as pornographic content ([Morgan, 2022](#)) and several motivations for deepfake fabrication tend to involve malice. Malicious motives, for instance, may range from a desire to organize political smear campaigns ([Sabrina, 2022](#)) to a desire to gain control or exact revenge on another individual who may become the target of a deepfaked attack ([Fikse, 2018](#); [Flynn et al., 2021](#); [Gamage et al., 2022](#); [Popova, 2019](#); [Winter and Salter, 2020](#)). Rarely, such motivations may also be associated with blackmailing or cyberbullying ([Sabrina, 2022](#)) or the execution of spear-phishing on high-potential targets ([Ghazi-Tehrani and Pontell, 2021](#)). Intriguingly, deepfakes may not always be constructed with malicious intent, but rather as a pastime or leisure tool ([de Seta, 2021](#); [Widder et al., 2022](#)). We also observe instances of online channels creating deepfakes to generate revenue while providing entertainment value ([Bode, 2021](#); [Gamage et al., 2022](#)). Deepfakes have also been reported to have been constructed for demonstration and educational purposes, which may be regarded as a virtuous motive ([Fikse, 2018](#); [Widder et al., 2022](#)).

However, while an individual's motivation to create a deepfake may stem from a variety of reasons, such as seeking revenge on another individual or a more frivolous motive of face-swapping a celebrity in a movie scene, we contend that motivation alone does not result in the fabrication of a deepfake. The reasons are twofold. First, creating a convincing deepfake that is hard to detect is not an easy task and requires sophisticated graphic design and technology expertise ([Bernaciak and Ross, 2022](#)). In this regard, our review suggests that digital platforms act as incubators for aspirants to hone their deepfake creation skills (e.g. [Gamage et al., 2022](#)). We contend that the platforms serve as enablers for the motivation to translate



into an actual deepfake fabrication. Second, deepfake technology has been excessively stigmatized for its malicious uses, and deepfake creators have been branded with a criminal and slanderous charge (Kobielus, 2020). In this context, our review suggests that an individual's sense of ethics may deter them from breaching norms and developing deepfakes that may cause harm to others (e.g. Widder *et al.*, 2022). Additionally, as a result of the sensationalization of deepfakes and their potential ramifications, we contend that individuals will be reluctant to engage in fabricating deepfakes due to the stigma attached to it. However, deepfakes are a dual-use technology, and as more socially redeeming uses emerge, the technology might begin to lose its stigma (Kobielus, 2020), and we contend that they would serve as justification to engage in deepfake fabrications. In a similar vein, if an abundance of illegal acts involving sexualized images ends up desensitizing people and normalizing such acts (e.g. Eelmaa, 2022), this may also dispel the stigma attached to deepfakes and lead to a proliferation of malicious deepfakes. Hence, we argue that the pathway from a specific motivation towards a tangible outcome in a deepfake fabrication is fueled by two essential factors (1) assertions surrounding the phenomenon and (2) digital platforms as enablers. We discuss these further in the following sections.

### 5.3 *Dichotomous assertions*

The literature presents contrasting perspectives on deepfakes as a concept with a split verdict (Gow, 2020). In this regard, there are numerous pro- and anti-deepfake claims in circulation. Arguments against deepfakes indicate grave concerns regarding its use in the production of mature content and the normalizing of such behavior (Eelmaa, 2022; Fikse, 2018; Flynn *et al.*, 2021; Gamage *et al.*, 2022; Newton and Stanfill, 2020; de Seta, 2021; Widder *et al.*, 2022; Winter and Salter, 2020). These anti-deepfake arguments also emphasize the gendered nature of the narrative emerging from the process and the violence connected to it (Eelmaa, 2022; Flynn *et al.*, 2021; Gregory, 2022; Newton and Stanfill, 2020; Widder *et al.*, 2022; Winter and Salter, 2020) including cybercrime (Ali *et al.*, 2021; Eelmaa, 2022; Flynn *et al.*, 2021; Ghazi-Tehrani and Pontell, 2021) and erosion of societal values, ethics and morals in the long run (Ali *et al.*, 2021; Bode, 2021; Eelmaa, 2022; Gamage *et al.*, 2022; Newton and Stanfill, 2020; de Seta, 2021; Widder *et al.*, 2022).

While these anti-deepfake claims are not without merit, it is worth noting that there are other claims that highlight the phenomenon's positive aspects. In this context, pro-deepfake arguments imply that deepfakes provide a platform for creative expression (Ali *et al.*, 2021; Bode, 2021; de Seta, 2021; Widder *et al.*, 2022), a means to engage with the celebrity being deepfaked (Bode, 2021; Popova, 2019) and an opportunity to showcase technological expertise and hone skills (Newton and Stanfill, 2020). Pro-deepfake opinions also indicate that widespread use of deepfake technology is inevitable (Widder *et al.*, 2022). As deepfake technology evolves and more use cases surface, it remains to be seen how the balance shifts in favor of one over the other.

### 5.4 *Digital platforms as enablers*

Our model recognizes the importance of digital platforms in facilitating the generation of deepfakes. In this context, three major forces facilitate the platform-driven creation process. First, platforms provide technical enthusiasts an incredible opportunity to train and master deepfake creation abilities (Gamage *et al.*, 2022; Newton and Stanfill, 2020; Widder *et al.*, 2022). Second, the collaborative nature of these platforms provides a mechanism for developers to participate in discussions that enable them to introduce new tools and techniques to community members (Gamage *et al.*, 2022; de Seta, 2021), who in turn also receive feedback to improve and hone their skills (Gamage *et al.*, 2022; Popova, 2019) while communities offer guidance on navigating legal roadblocks (Gamage *et al.*, 2022) and so on. Third, platforms offer a vast database of deepfakes (Fikse, 2018; Gamage *et al.*, 2022) in addition to the infrastructure for deepfake monetization (Bode, 2021; Gamage *et al.*, 2022).

### 5.5 Genres of deepfakes

The literature contains several subgenres of deepfakes, as delineated by the preceding discussion. On the basis of the reviewed literature, deepfakes can be broadly classified into two main categories: malevolent and innocuous. Among the malevolent deepfake fabrications, political deepfakes (Ali *et al.*, 2021; Gamage *et al.*, 2022; Ghazi-Tehrani and Pontell, 2021; Gregory, 2022; Sabrina, 2022; Widder *et al.*, 2022) and explicit deepfakes (Eelmaa, 2022; Fikse, 2018; Flynn *et al.*, 2021; Gamage *et al.*, 2022; Newton and Stanfill, 2020; Popova, 2019; de Seta, 2021; Widder *et al.*, 2022; Winter and Salter, 2020) have gained the most attention, although concerns are beginning to surface over AI-enabled deepfake reviews as well (Gamage *et al.*, 2022). Among deepfakes of a benign nature, they have been leveraged towards participating in satire (Fikse, 2018; Widder *et al.*, 2022). While celebrities have been the target of explicit deepfakes, material such as a deepfaked video of Keanu Reeves demonstrates a positive use of celebrity involvement through the phenomenon, as well as a method for generating revenue via digital platforms (Bode, 2021). Finally, deepfakes could be used to improve digital media literacy, and online communities have presented themselves as a platform for developing educative deepfakes and discouraging malevolent ones (Ali *et al.*, 2021; Widder *et al.*, 2022; Winter and Salter, 2020).

At this stage, we reiterate that motivations alone do not lead to the creation of deepfakes. While the motivation to fabricate a deepfake may exist, the pathway from this motivation to the actual fabrication of a deepfake is contingent on two factors: the digital platforms as enablers and the assertions surrounding the phenomenon. In the case of the former, the technological sophistication required to create a convincing deepfake (Bernaciak and Ross, 2022) makes it difficult for a layman without the requisite skills to develop a convincing deepfake, whereas in the case of the latter, the stigmatization of deepfakes (Kobielus, 2020) may act as a deterrent against creating a deepfake. As stated previously in this section, digital platforms function as enablers by providing direction through community engagement, curating knowledge through code repositories (e.g. Winter and Salter, 2020), and providing feedback to users (e.g. Gamage *et al.*, 2022). Collectively, these platforms assist aspirants in perfecting their deepfaking skills. Alongside, the inherent stigmatization of deepfakes and an abundance of illegal acts involving sexualized imagery may end up desensitizing individuals and normalizing such acts (e.g. Eelmaa, 2022), thereby dispelling any stigma attached to harmful deepfakes and resulting in their proliferation. In a similar vein, the advent of more socially acceptable use cases may potentially aid in removing the stigma (Kobielus, 2020) and serve as justification for engaging in creating deepfakes.

We illustrate this through an example. For instance, while sexual gratification could be a motivation (Flynn *et al.*, 2021), we contend that this motive alone does not equate to creating a deepfake. First, the technological platforms function as a learning enabler since the developers of these software powering deepfakes, in their pursuit of a “higher purpose,” are continually working to improve the software while overlooking its potential misuse for harmful applications (Widder *et al.*, 2022). Second, perceptions of certain sexualized content as a form of private activity (Eelmaa, 2022) or its framing as an expression of creativity (de Seta, 2021) serve as affirmations that reinforce the motivations. We contend that these digital enablers in the form of technology platforms and the expertise contained therein, along with the normalization of activity around the production of deepfakes, facilitate the transition of motivations towards the fabrication of deepfakes.

### 5.6 Dissemination of deepfakes

Deepfake dissemination may be intentional and linked to foisting political harm and distorting the electoral landscape (Gamage *et al.*, 2022; Sabrina, 2022), or it may be tied to an act of anger and rage leading to the sharing of explicitly natured deepfakes (Flynn *et al.*, 2021),

or it may be aimed at satiating sensual desires (Eelmaa, 2022). Interestingly, deepfake distribution may also be entirely unintentional and unrelated to any goal. The inability to recognize the media as a deepfake may result in its dissemination (Ali *et al.*, 2021; Thaw *et al.*, 2020), while it may also be disseminated for recreational or entertaining purposes (Bode, 2021; de Seta, 2021). In this regard, journalists' fact-checking abilities are of the utmost significance, and this group risks disseminating deepfaked content due to time restrictions and a sense of competition with other broadcaster units (Himma-Kadakas and Ojamets, 2022).

### 5.7 Consequences of deepfakes

Deepfakes pose severe and multifarious repercussions. At the societal level, deepfakes may pose a risk to electoral discourse by swaying public opinion (Fikse, 2018; Gamage *et al.*, 2022). As the pervasiveness of deepfakes increases and pornographic deepfakes abound, there may be an erosion of ethics and morality in the long run (Ali *et al.*, 2021; Eelmaa, 2022; Gamage *et al.*, 2022; Ghazi-Tehrani and Pontell, 2021; Newton and Stanfill, 2020; de Seta, 2021; Widder *et al.*, 2022; Winter and Salter, 2020). At the individual level, the literature highlights the victimization of targets who must deal with deepfaked creations that use their images without consent (Eelmaa, 2022; Flynn *et al.*, 2021) and may also be subject to cyberbullying or blackmail (Sabrina, 2022), while the pervasiveness of deepfakes in the long run ushers in post-truth incredulity and distrust in media (Fikse, 2018; Sabrina, 2022; Widder *et al.*, 2022). However, this distrust and skepticism may also serve to foster a generation of responsible consumers of digital media (Ali *et al.*, 2021). At the organizational level, deepfakes pose firm-level consequences by targeting high-value individuals in spear-phishing attacks (Ghazi-Tehrani and Pontell, 2021), while literature also discusses the possibility of fake reviews (Gamage *et al.*, 2022), which have posed a constant threat to industries (Wu *et al.*, 2020). Moreover, journalists across platforms are exposed to information from numerous groups and readers seeking to advance their agendas and the prevailing information disorder leads them to develop a deep sense of distrust towards information originating from the public such as activists, ordinary citizens, or public relations agents (Himma-Kadakas and Ojamets, 2022). The limited skills of journalists in recognizing digitally manipulated footage such as deepfakes serve as a warning of the risks they are exposed to due to low visual verification skills (Himma-Kadakas and Ojamets, 2022). As a result, journalist organizations are witnessing a decline in the public's trust in journalism (Himma-Kadakas and Ojamets, 2022), making it challenging for them to regain their lost credibility.

### 5.8 Governing interventions

The framework acknowledges the significance of multiple approaches designed to regulate the proliferation of deepfakes and their repercussions. We observe three major interventions in this respect. First, human-directed mechanisms refer to programs designed to improve media literacy and critical thinking abilities (Ali *et al.*, 2021; Bode, 2021; Fikse, 2018; Gregory, 2022; Himma-Kadakas and Ojamets, 2022; Sabrina, 2022). For instance, critical thinking skills among journalists aid in analyzing and verifying the material, which not only aids in spotting erroneous information but also prevents the publication of such reports, which may result in a loss of public faith in the media (e.g. Himma-Kadakas and Ojamets, 2022). Intriguingly, the prevalence of deepfake may act as a form of immunization, as such extensive diffusion will cause individuals to stop trusting videos at face value and instead subject media content to verification (Widder *et al.*, 2022). In this regard, the literature also highlights the efforts of citizen journalism networks and journalist groups' joint fact-checking attempts to combat the problem (Gregory, 2022; Himma-Kadakas and Ojamets, 2022). Second, the framework emphasizes the significance of technology infrastructure and associated interventions as governing mechanisms. For example, forensic techniques can help detect manipulated media

---

(Himma-Kadakas and Ojamets, 2022), whilst multi-factor authentication can thwart phishing attacks (Ghazi-Tehrani and Pontell, 2021). In this scenario, deepfake makers may consider marking content with a watermark that could mitigate effects even if the content is disseminated outside of its original context (Popova, 2019). In other technological infrastructure-related regulations, platform rules and standards on social media sites could help mitigate deepfake-related harm, while such recommendations are also echoed by community members (Gamage *et al.*, 2022) who make efforts to contain deepfaked material within the communities of practice (Popova, 2019). For instance, online platforms have incorporated revisions to their content policies which prohibit involuntary pornography, sexually explicit fake imagery, manipulated media and misleading impersonation via deepfakes (e.g. Bickert, 2020; Binder, 2020; Hawkins, 2018; Reddit, 2020, 2022). While such platform policies typically limit an individual's access to the platform or ban a community of users who engage in such infractions (e.g. Anderson, 2022; Hern, 2018b), the government legislations could result in harsher punishments, which takes us to the third intervention. In this context, our framework emphasizes the need for legislative mechanisms, such as prohibitions on deepfake content, to combat the deepfake menace (Ali *et al.*, 2021; Ghazi-Tehrani and Pontell, 2021; Gregory, 2022; Winter and Salter, 2020). This is consistent with the European Union's recent strengthening of legislation around misinformation, including deepfakes (Collins, 2022), while in other legal reforms, lawmakers are considering criminal penalties and prosecution for individuals who fabricate or share pornographic deepfakes (E&T, 2022). Countries like China are contemplating deepfake legislation that would impose severe penalties on platforms that violate regulations (Elliott and Tobin, 2022), and other nations are expected to follow suit.

### *5.9 Challenges posed by digital platforms*

We contend that platform-level issues impede the regulatory mechanisms from obtaining sufficient momentum, thwarting efforts to regulate the diffusion and adverse consequences of deepfakes. These obstacles can be broadly categorized into three categories. First, with regard to measurement ambiguities, deepfakes of an explicit nature in hentai, a genre of Japanese anime, have frequently been portrayed as an art, with censorship ideas met with strong opposition (Eelmaa, 2022), while legal versus ethical conflicts are noted in the framing of objectionable content on platforms (Newton and Stanfill, 2020). Second, the ambiguity of platform standards presents a problem in that platform-specific timing and nature of policies can afford platforms disproportionate authority over deepfake-related policies (Gamage *et al.*, 2022). Third, the literature demonstrates that a lack of control and moderation on platforms due to indifference towards harms (Newton and Stanfill, 2020), absence of ethics around seeking consent (Newton and Stanfill, 2020; Winter and Salter, 2020) and appreciation for explicit natured deepfake content and intent to continue such practices (Winter and Salter, 2020) clearly hinders the progress of governing mechanisms, such as prohibitions on deepfake content.

## **6. Discussion**

Deepfakes are a relatively new phenomenon that has been extensively debated for both their positive and bad applications. These adverse consequences have also prompted research on deepfake production and detection (e.g. Almars, 2021; Mirsky and Lee, 2021), while it has been argued that deepfake technology poses an epistemic danger and ushers in a post-truth future (Chesney and Citron, 2019; Fallis, 2020). On the downside, deepfakes usher in a profound sense of mistrust and a state of epistemological pluralism in which individuals may not only question reality but strive to reorient towards this plurality through several ways of

---

acquiring knowledge (Porter, 2020). In this regard, deepfakes pose a threat of inflicting epistemic harm wherein our capacities to acquire accurate beliefs may be hindered as a consequence of this technology (Fallis, 2020), while it may also allow us to transition from what existed only in our imaginations to experience it in reality through digitally altered footage (Kerner and Risse, 2021). However, scholars have also argued that such concerns may be unfounded and exaggerated (Harris, 2021). On the bright side, this epistemic ambiguity may in fact prompt viewers to scrutinize even the most legitimate videos if they come from a questionable source (Harris, 2021). Furthermore, we are starting to witness a growing number of positive use cases in instances such as protecting patient privacy in healthcare (e.g. Crystal *et al.*, 2020; Shin *et al.*, 2018), creating educational videos (e.g. Griffin, 2019), promoting fashion through deepfake apparel in the fashion metaverse (e.g. Darko, 2021; McDowell, 2021) and so on. As technology advances and more use cases emerge, it remains to be seen how the balance will shift in favor of one over the other.

This work sheds light on the socio-technical aspects of deepfake engagement through a meta-synthesis of qualitative literature on deepfakes from the theoretical standpoint of SST. The conceptual framework developed through this study helps clarify our understanding of deepfakes through a comprehensive view of the phenomenon. In doing so, we addressed the two research questions raised at the outset of this study. RQ1 was addressed by demonstrating how motivations, contingent on assertions and platforms as enablers, drive the creation and dissemination of deepfakes. RQ2 was addressed by demonstrating how a multitude of governing interventions constrain the dissemination and ramifications of deepfakes, as well as how the influence of these interventions is diminished through challenges posed by the digital platforms. This is one of the first studies to establish a complete, theoretically grounded perspective on how individuals engage with deepfakes, and it has a number of theoretical and practical implications that are discussed below.

### 6.1 Theoretical implications

Meta-synthesis helps in progressing beyond the individual studies and identifying how the categories that emerged from those studies are related to each other across studies (Khan and Krishnan, 2021). As evidenced by our analysis, while the individual studies either uncovered the motivations for deepfake generation, arguments in support of and against deepfakes, or the role of digital technologies as enablers, they rarely discussed all of them in conjunction with one another. Similarly, while the dissemination of deepfakes, related consequences, and measures to curb the spread were discussed in the extant literature, there was rarely a concerted view regarding how the unabated spread may be curbed by a diverse array of governing interventions and why such interventions are generally unsuccessful. As a response to this siloed treatment across studies, the current study systematically synthesizes the varied aspects connected with deepfakes and presents a holistic framework of deepfake engagement. In doing so, the study highlights the role of technology platforms in promoting the fabrication of deepfakes, while ongoing efforts to improve the accuracy of deepfakes serve as feedback to stimulate their continuous adoption. In addition, the absence of clearly defined guardrails on these platforms hinders efforts to curb the unchecked proliferation and devastating effects of deepfakes. However, the narrative has a silver lining in that continual refinement of assertions may strike a balancing act in limiting the malicious use of deepfake technology while fostering its positive use cases. Table 3 provides a summary of the study's key contributions in terms of new insights from the meta-synthesis which were not apparent in the individual studies.

The myriad of attributes that emerge from the meta-synthesis provides a detailed understanding of how deepfakes are produced and consumed, while the attributes prompt researchers investigating the evolutionary nature of deepfakes as a socio-technical

Sr. No	Study contributions	State of the literature	Implications for research
1	Explicates that the pathway from a motivation to a deepfake fabrication is not straightforward but contingent on the digital platforms and assertions. Networked spaces on platforms help aspirants in perfecting deepfake technology skills. However, while assertions that stigmatize deepfakes may dissuade individuals from deepfaking, desensitization caused by an excess of harmful deepfakes may serve as justification for engaging in similar fabrications	Discusses malevolent and non-malevolent motivations largely in isolation and rarely in the context of broader assertions and platforms facilitating deepfake fabrications. Consequently, the extant literature predominantly portrays a picture of motivations directly translating to a deepfake fabrication	Expands the knowledge base for IS researchers contributing towards ethical AI frameworks for responsible development and deployment of synthetic media that powers deepfakes
2	Emphasizes nefarious use of deepfakes and their stigmatization. Nevertheless, strikes a balance with a view on both malicious and innocuous deepfake genres	Focuses minimally on positive use cases and primarily on malicious uses	Opens up new avenues for research on deepfake applications (e.g. preserving patient privacy in healthcare) by emphasizing that AI is intrinsically neither good nor evil
3	Advances our understanding of governing interventions to limit deepfake sharing and its consequences. In this regard, sheds light on platform-level issues such as vague standards and lack of enforcement which dampen efforts	Highlights efforts such as media literacy, image forensics, etc. but rarely discusses the causes for the unabated spread of deepfakes despite these interventions	Raises possibilities for researchers to examine the potential influence on basic liberties such as the right to free speech in light of regulations for synthetic media and the shifting onus on platforms for stricter enforcement
4	Provides an interconnected, multi-layered perspective on deepfake-related consequences at the societal, organizational, and individual levels, as well as how the influence at one level cascades to other levels as a result of both intentional and inadvertent sharing of deepfakes	Predominantly focuses on consequences only at one of the three levels and rarely offers a multi-layered perspective	Creates opportunities for researchers to comprehend how deepfakes influence individuals' affective states and cognitive capacities and potentially implant false memories. The perspective on cascading effects enables researchers to preempt potential ramifications and contribute to the pedagogical development of media literacy and digital awareness initiatives
5	Provides a theoretical foundation for the evolving nature of deepfakes and a perspective on how the underlying technology socially evolves through continual knowledge contribution and feedback-led refinement of assertions	Partially emphasizes knowledge contribution through code repositories and user feedback, but continual refinement of assertions and uses of technology are rarely highlighted in conjunction with deepfake use cases	Opens up possibilities for IS researchers to trace the trajectory as the phenomenon matures and examine the socio-technical linkages embedded in its evolution

**Table 3.**  
Preview of study contributions

phenomenon to pay heed to the contrasting arguments for and against deepfakes. These assertions are likely to influence the evolution of deepfakes in the future. In doing so, the findings from this study have several implications for researchers in the information systems (IS) discipline. First, the meta-synthesis sheds light on how the platforms facilitate the construction of deepfakes. The networked structure of platforms including technological expertise highlights the significance of platform dynamics in deepfake production. The endogenous choices of platforms coevolve with the dynamics of the exogenous environment (Tiwana *et al.*, 2010), and the findings of this study provide an opportunity for IS researchers to examine the evolution of platforms and how they are influenced by deepfakes and the environment. Second, while deepfakes raise several serious concerns related to data ethics and privacy, we also note an emerging trend of positive use cases for synthetic media powering these deepfakes in areas such as healthcare to improve disease diagnosis (e.g. Kearney *et al.*, 2022; Newswise, 2022) or creative pursuits such as reinventing art museums for the digital age (e.g. Snow, 2021). In this regard, the review offers crucial insights for IS researchers to play a pivotal role in developing ethical AI frameworks and shaping the guiding principles for the technology platforms powering these deepfakes, so data ethics is upheld while privacy concerns are alleviated, and privacy by design becomes the norm. Third, the menace of fake news across domains (e.g. Di Domenico and Visentin, 2020; Vasit and Krishnan, 2022b) and its increased diffusion on social media platforms during the pandemic (Di Domenico *et al.*, 2021b) has had these platforms constantly grappling with combating fake news (Collins *et al.*, 2021). Deepfakes threaten to exacerbate this escalating crisis of fake news (Han *et al.*, 2021). In the wake of this crisis, this study offers a balanced view of socio-technical characteristics powering the process of deepfake engagement and offers timely insights into the evolving nature of deepfakes and the conflicting forces which shape this phenomenon. This serves as vital insights for IS researchers to examine the socio-technical linkages embedded in its evolution and inform platform players and policymakers around measures to combat deepfakes and steer clear of ill-conceived measures which may threaten basic liberties such as the freedom of expression. This assumes importance, especially in the context of platforms that are struggling to strike a balance between freedom of expression and limiting the spread of fake news (Di Domenico *et al.*, 2021a). Fourth, the epistemic crisis brought on by deepfakes is anticipated to lead to media mistrust and high levels of skepticism, while the unchecked proliferation of deepfakes risks normalizing the phenomenon's negative use cases. This crisis offers the possibility for IS researchers to comprehend how deepfakes influence the affective states and cognitive abilities of individuals, while an investigation of socio-technical interactions may position them to provide guidance to policymakers and regulatory agencies regarding the dangers of deepfakes and their associated mitigation mechanisms. The findings from the study also serve as useful insights for IS researchers to contribute to the pedagogical development of media literacy and digital awareness initiatives.

### *6.2 Practical implications*

This study's findings have substantial implications for practitioners and policymakers as well. First, in offering an SST perspective on how individuals develop, process, and disseminate deepfakes along with governing interventions to mitigate the consequences of deepfakes, the study offers insights to policymakers who must recognize that deepfake technology will continue to proliferate and its uses will be shaped by the social context in which it evolves. This should serve as useful guidance for policymakers in crafting policy instruments related to the fabrication and dissemination of deepfakes, while the steady evolution underscores the need for constant feedback loops and recalibration of policies to stem the unabated proliferation of deepfakes.

---

Second, the study offers a balanced perspective by highlighting not just the malicious use of deepfake technology but also shedding light on positive use cases that are seldom discussed. This provides guidance for policymakers, who must take a balanced stance to ensure that legislations do not impede the positive applications. As beneficial use cases of synthetic media emerge, the regulations must incorporate necessary exemptions to advance the technology's benefits while curbing its negative use. In the same context, while regulations may help preserve social stability, policymakers must be wary of the potential consequences of such rules on online freedom and creativity, make suitable amendments, and avoid going overboard with radical reforms.

Third, the framework illustrates the influence of platforms on the evolution of deepfakes. In this regard, practitioners may use the study's findings to draft platform rules and standards that promote the appropriate use of platforms and their communities, while combating the relentless proliferation of malevolent deepfakes and their repercussions. This should also encourage practitioners to build more sophisticated deepfake detection mechanisms to detect and flag inappropriate content on their platforms and to control their dissemination.

Fourth, the study highlights the tightly coupled nature of governing interventions and digital platforms, emphasizing that the success of interventions is contingent on the proactive participation of platforms. This implies that governments and policymakers must collaborate with platform leaders to define the guardrails for deepfake engagement, while platforms must assume equal responsibility by supporting regulations and calibrating their response with advanced deepfake detection mechanisms and robust platform standards. The lack of moderating efforts highlighted through the framework stresses the need for platform collaboration to extend beyond the formulation of standards to ensure continual oversight and enforcement of the rules.

Lastly, the study provides a victim's perspective on deepfakes that pro-deepfake groups and deepfake technology developer communities must pay attention to. In doing so, the study aims to stimulate these groups to rethink their future actions to alleviate fears and victimization among deepfaked targets and promote the positive applications of synthetic media. The victimization through deepfakes highlighted in this review should also aid in educating citizens about the debilitating consequences of deepfakes, which are far removed from their frivolous use, and help in mobilizing social welfare organizations and communities of trusted deepfake flaggers who can serve as moderators of deepfake content on social media platforms.

## **7. Limitations and future research directions**

The current study employed a meta-synthesis approach to clarify the links between constructs within the context of deepfake engagement. However, there are limitations to the study. First, the novelty of deepfake as a phenomenon motivated us to include a variety of grey literature in our search strategy and subsequent review. While this inclusion through our search protocol assures comprehensive coverage of the literature, the upsurge in deepfake publications may provide future researchers with an opportunity to refine the findings by concentrating on articles from peer-reviewed journals and conference proceedings. Second, our literature search led us to a few studies that were behind a paywall, limiting our access to the complete text of these articles. In our assessment, we have nonetheless endeavored to capture the substance of such articles. Thirdly, our evaluation was led by a search methodology consisting of a collection of pertinent and contextual keywords and well-known databases to locate relevant publications. On this basis, we position our work as entirely unique, exhaustive, and critical in nature. Nevertheless, as the phenomenon evolves, its scope may extend, and subsequent assessments may include additional works. In this regard, our study may serve as a catalyst for additional research on the emerging topic of deepfakes.



---

## 8. Conclusions

Deepfakes are often viewed as the next generation of fake news, which threatens to undermine faith in online information and precipitate an epistemic crisis (Fallis, 2020; Kalpokas and Kalpokiene, 2022a), while some have claimed that such fears are exaggerated (Harris, 2021). The looming threat of deepfakes and their pervasiveness has attracted the attention of researchers, platform players, and governments around the world, although some have also advocated for its positive applications. This meta-synthesis of 16 qualitative studies comprehensively captures the divergent perspectives and evolving nature of the phenomenon in order to develop a thorough theoretical understanding of the nuances of how individuals engage with deepfakes so that the negative effects can be mitigated, and the positive effects can be fully utilized. We contend that the framework developed through this study will provide a comprehensive understanding of deepfake as a phenomenon, including its merits and demerits, and guide deepfake researchers, administrators, and platform players in charting the future directions for deepfakes by reducing concerns surrounding it and enhancing its positive use cases.

## Note

1. We thank Referee #1 for this thought.

## References

- Adams, R.J., Smart, P. and Huff, A.S. (2017), "Shades of grey: guidelines for working with the grey literature in systematic reviews for management and organizational studies", *International Journal of Management Reviews*, Vol. 19 No. 4, pp. 432-454.
- Afchar, D., Nozick, V., Yamagishi, J. and Echizen, I. (2018), "MesoNet: a compact facial video forgery detection network", *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 1-7.
- Agalianos, A., Whitty, G. and Noss, R. (2006), "The social shaping of logo", *Social Studies of Science*, SAGE Publications, Vol. 36 No. 2, pp. 241-267.
- Ahmed, S. (2021a), "Fooled by the fakes: cognitive differences in perceived claim accuracy and sharing intention of non-political deepfakes", *Personality and Individual Differences*, Vol. 182, p. 111074.
- Ahmed, S. (2021b), "Navigating the maze: deepfakes, cognitive ability, and social media news skepticism", *New Media and Society*, SAGE Publications, Vol. ahead-of-print, p. 14614448211019198.
- Ahmed, S. (2021c), "Who inadvertently shares deepfakes? Analyzing the role of political interest, cognitive ability, and social network size", *Telematics and Informatics*, Vol. 57, p. 101508.
- Ahmed, S. (2022), "Disinformation sharing thrives with fear of missing out among low cognitive news users: a cross-national examination of intentional sharing of deep fakes", *Journal of Broadcasting and Electronic Media*, Routledge, Vol. 66 No. 1, pp. 89-109.
- Al-Sibai, N. (2022), "Cops are now deepfaking murder victims to catch their killers", *Futurism*, available at: <https://futurism.com/the-byte/deepfake-murder-victim> (accessed 5 June 2022).
- Ali, S., DiPaola, D., Lee, I., Sindato, V., Kim, G., Blumofe, R. and Breazeal, C. (2021), "Children as creators, thinkers and citizens in an AI-driven future", *Computers and Education: Artificial Intelligence*, Vol. 2, p. 100040.
- Allen, C., Vassilev, I., Kennedy, A. and Rogers, A. (2016), "Long-term condition self-management support in online communities: a meta-synthesis of qualitative papers", *Journal of Medical Internet Research*, Vol. 18 No. 3, p. e5260.
- Allison, T. (2021), "Race and the digital face: facial (mis)recognition in Gemini Man", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 999-1017.
- Almars, A.M. (2021), "Deepfakes detection techniques using deep learning: a survey", *Journal of Computer and Communications*, Scientific Research Publishing, Vol. 9 No. 5, pp. 20-35.

- 
- Alrasheed, N., Zachariah, A., Prasanna, S., Rao, D. and Rao, P. (2020), "Deepfakes for histopathology images: myth or reality?", *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), presented at the 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pp. 1-7.
- Analytics Insight (2022), "Deepfake is the scariest thing happened to mankind", *Analytics Insight*, available at: <https://www.analyticsinsight.net/deepfake-is-the-scariest-thing-happened-to-mankind/> (accessed 5 June 2022).
- Anderson, M. (2022), "Reddit bans 'SFW' deepfake community", *Unite.AI*, available at: <https://www.unite.ai/reddit-bans-sfw-deepfake-community/> (accessed 21 June 2022).
- Andrejevic, M., Fordyce, R., Li, L. and Trott, V. (2022), "Automated culture: introduction", *Cultural Studies*, Routledge, Vol. ahead-of-print, pp. 1-19.
- Anuar, M. and Ibrahim, F. (2020), "The news gallery: beyond headlines", *BiblioAsia*, available at: <https://biblioasia.nlb.gov.sg/vol-16/issue-1/apr-jun-2020/the-news-gallery> (accessed 12 June 2022).
- Ascott, T. (2020), "Microfake: how small-scale deepfakes can undermine society", *Journal of Digital Media and Policy*, Vol. 11, pp. 215-222.
- Athanasopoulou, A. and Dopson, S. (2018), "A systematic review of executive coaching outcomes: is it the journey or the destination that matters the most?", *The Leadership Quarterly*, Vol. 29 No. 1, pp. 70-88.
- Atkins, S., Lewin, S., Smith, H., Engel, M., Fretheim, A. and Volmink, J. (2008), "Conducting a meta-ethnography of qualitative literature: lessons learnt", *BMC Medical Research Methodology*, Vol. 8 No. 1, p. 21.
- Ayers, D. (2021), "The limits of transactional identity: whiteness and embodiment in digital facial replacement", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 1018-1037.
- Bacchi, U. (2022), "Performing artists push for copyright protection from AI deepfakes | Reuters", *Reuters*, available at: <https://www.reuters.com/legal/litigation/performing-artists-push-copyright-protection-ai-deepfakes-2022-05-18/> (accessed 5 June 2022).
- Bajema, N., Smith, C.S. and Garisto, D. (2022), "AI's Real Worst-Case Scenarios: who needs Terminators when you have precision clickbait and ultra-deepfakes?", *IEEE Spectrum*, Institute of Electrical and Electronics Engineers, Vol. 59 No. 1, pp. 8-14.
- Barber, G. (2019), "Deepfakes are getting better, but they're still easy to spot", *Wired*, available at: <https://www.wired.com/story/deepfakes-getting-better-theyre-easy-spot/> (accessed 21 June 2022).
- Baten, R.A. and Hoque, E. (2021), "Technology-driven alteration of nonverbal cues and its effects on negotiation", *Negotiation Journal*, Vol. 37 No. 1, pp. 35-47.
- Bazarkina, D.Yu. and Pashentsev, Y.N. (2019), "Artificial intelligence and new threats to international psychological security", *Russia in Global Affairs*, available at: <https://eng.globalaffairs.ru/articles/artificial-intelligence-and-new-threats-to-international-psychological-security/> (accessed 24 November 2021).
- Bazarkina, D., Pashentsev, E. and Bazarkina, Y. (2020), "Malicious use of artificial intelligence: new psychological security risks in BRICS countries", *Russia in Global Affairs*, Vol. 18, doi: [10.31278/1810-6374-2020-18-4-154-177](https://doi.org/10.31278/1810-6374-2020-18-4-154-177).
- Bernaciak, C. and Ross, D. (2022), "How easy is it to make and detect a deepfake?", *SEI Blog*, available at: <https://insights.sei.cmu.edu/blog/how-easy-is-it-to-make-and-detect-a-deepfake/> (accessed 14 September 2022).
- Bibri, S.E. (2022), "The social shaping of the metaverse as an alternative to the imaginaries of data-driven smart cities: a study in science, technology, and society", *Smart Cities*, Multidisciplinary Digital Publishing Institute, Vol. 5 No. 3, pp. 832-874.
- Bibri, S.E. and Krogstie, J. (2017), "On the social shaping dimensions of smart sustainable cities: a study in science, technology, and society", *Sustainable Cities and Society*, Vol. 29, pp. 219-246.
- Bickert, M. (2020), "Enforcing against manipulated media", *Meta*, available at: <https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/> (accessed 14 September 2022).

- 
- Binder, M. (2020), "Reddit beats Facebook by rolling out A superior deepfakes policy", *Mashable India*, available at: <https://in.mashable.com/tech/10336/reddit-beats-facebook-by-rolling-out-a-superior-deepfakes-policy> (accessed 14 September 2022).
- Biometric technology today (2020), "iProov helps Singapore to world first for facial recognition", *Biometric Technology Today*, MA Business, pp. 1-2.
- Biometric Technology Today (2021), "Security firm: deepfakes are 'fraud's next Frontier", *Biometric Technology Today*, Vol. 2021 No. 6, pp. 2-3.
- Bode, L. (2021), "Deepfaking Keanu: YouTube deepfakes, platform visual effects, and the complexity of reception", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 919-934.
- Bode, L., Lees, D. and Golding, D. (2021), "The digital face and deepfakes on screen", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 849-854.
- Bodi, M. (2021), "The first amendment implications of regulating political deepfakes", *Rutgers Computer and Technology Law Journal*, Vol. 47 No. 1, pp. 143-172.
- Borges-Tiago, T., Tiago, F., Silva, O., Guaita Martínez, J.M. and Botella-Carrubi, D. (2020), "Online users' attitudes toward fake news: implications for brand management", *Psychology and Marketing*, Vol. 37 No. 9, pp. 1171-1184.
- Boumans, E. (2021), "Animated personalities: cartoon characters and stardom in American theatrical shorts", *Historical Journal of Film, Radio and Television*, Routledge, Vol. 41 No. 4, pp. 870-872.
- Breen, D.C. (2021), "Silent no more: how deepfakes will force courts to reconsider video admission standards", *Journal of High Technology Law*, Vol. 21 No. 1, pp. 122-161.
- Brennen, J.S., Simon, F.M. and Nielsen, R.K. (2021), "Beyond (Mis)Representation: visuals in COVID-19 misinformation", *The International Journal of Press/Politics*, SAGE Publications, Vol. 26 No. 1, pp. 277-299.
- Brown, W. and Fleming, D.H. (2020), "Celebrity headjobs: or oozing squid sex with a framed-up leaky Schar-JØ", *Porn Studies*, Routledge, Vol. 7 No. 4, pp. 357-366.
- Burgstaller, M. and Macpherson, S. (2021), "Deepfakes in international arbitration: how should tribunals treat video evidence and allegations of technological tampering?", *The Journal of World Investment and Trade*, Brill Nijhoff, Vol. 22 Nos 5-6, pp. 860-890.
- Campbell, C., Plangger, K., Sands, S. and Kietzmann, J. (2022), "Preparing for an era of deepfakes and AI-generated ads: a framework for understanding responses to manipulated advertising", *Journal of Advertising*, Routledge, Vol. 51 No. 1, pp. 22-38.
- Carlson, M. (2021), "Skepticism and the digital information environment", *SATS*, De Gruyter, Vol. 22 No. 2, pp. 149-167.
- Carvajal, L. and Iliadis, A. (2020), "Deepfakes: a preliminary systematic review of the literature", *AoIR Selected Papers of Internet Research*, doi: [10.5210/spir.v2020i0.11190](https://doi.org/10.5210/spir.v2020i0.11190).
- Chandler, S. (2020), "Why deepfakes are A net positive for humanity", *Forbes*, available at: <https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/> (accessed 25 June 2022).
- Chang, W.-L. (2019), "Social shaping of robot adoption in organizational contexts", Ph.D., Indiana University, available at: <https://www.proquest.com/docview/2320973964/abstract/206FF60883F94803PQ/1> (accessed 29 August 2022).
- Chaudhary, S., Saifi, R., Chauhan, N. and Agarwal, R. (2021), "A comparative analysis of deep fake techniques", *Presented at the 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, pp. 300-303.
- Chee, F.Y. (2022), "Exclusive: google, Facebook, Twitter to tackle deepfakes or risk EU fines - document | Reuters", *Reuters*, available at: <https://www.reuters.com/technology/google-facebook-twitter-will-have-tackle-deepfakes-or-risk-eu-fines-sources-2022-06-13/> (accessed 14 June 2022).

- 
- Cheng, Y. and Chen, Z.F. (2020), "The influence of presumed fake news influence: examining public support for corporate corrective response, media literacy interventions, and governmental regulation", *Mass Communication and Society*, Routledge, Vol. 23 No. 5, pp. 705-729.
- Cheng, X., Liu, S., Sun, X.-Q., Wang, Z., Zhou, H., Shao, Y. and Shen, H. (2021), "Combating emerging financial risks in the big data era: a perspective review", *Fundamental Research*, Vol. 1 No. 5, pp. 595-606, doi: [10.1016/J.FMRE.2021.08.017](https://doi.org/10.1016/J.FMRE.2021.08.017).
- Chesney, R. and Citron, D.K. (2018), "Deep fakes: a looming challenge for privacy, democracy, and national security", SSRN Scholarly Paper No. ID 3213954, Social Science Research Network, Rochester, NY, doi: [10.2139/ssrn.3213954](https://doi.org/10.2139/ssrn.3213954).
- Chesney, R. and Citron, D. (2019), "Deepfakes and the new disinformation war: the coming age of post-truth geopolitics", *Foreign Affairs*, Vol. 98, p. 147.
- Chou, L.-C. (2022), "The social selection of the third party payment technology in China on the social shaping of technology perspective", *Global Academic Journal of Humanities and Social Sciences*, Vol. 4 No. 1, pp. 29-32.
- Chowdhury, S.M.A.K. and Lubna, J.I. (2020), "Review on deep fake: a looming technological threat", *Presented at the 2020 11th International Conference on Computing, Communication and Networking Technologies ICCCNT*, pp. 1-7.
- Chrystall, A.B. (2021), "Making sense of indigenous ↔ colonial encounters: New Zealand's treaty of waitangi in a digital age", *Laws*, Multidisciplinary Digital Publishing Institute, Vol. 10 No. 2, p. 45.
- Chudinov, A., Koshkarova, N. and Ruzhentseva, N. (2019), "Linguistic interpretation of Russian political agenda through fake, deepfake, post-truth", *Journal of Siberian Federal University. Humanities and Social Sciences*, Vol. 12 No. 10, pp. 1840-1853.
- Citron, D. (2019), "Sexual privacy", *Yale Law Journal*, Vol. 128 No. 7, p. 1870.
- CNN (2022), *Deepfakes Are Now Trying to Change the Course of War*, Fiji Broadcasting Corporation, available at: <https://www.fbcnews.com.fj/business/deepfakes-are-now-trying-to-change-the-course-of-war/> (accessed 5 June 2022).
- Cochran, J.D. and Napshin, S.A. (2021), "Deepfakes: awareness, concerns, and platform accountability", *Cyberpsychology, Behavior and Social Networking*, Vol. 24 No. 3, pp. 164-172.
- Cole, S. (2017), "AI-assisted fake porn is here and we're all fucked", *Vice*, available at: <https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn> (accessed 5 November 2021).
- Collins, K. (2022), "Big tech faces fines for deepfakes, bots, fake accounts under new EU rules", *CNET*, available at: <https://www.cnet.com/news/politics/eu-strengthens-disinformation-rules-to-target-deepfakes-bots-fake-accounts/> (accessed 25 June 2022).
- Collins, B., Hoang, D.T., Nguyen, N.T. and Hwang, D. (2021), "Trends in combating fake news on social media – a survey", *Journal of Information and Telecommunication*, Taylor & Francis, Vol. 5 No. 2, pp. 247-266.
- Conte, P. (2019), "Mockumentality: from hyperfaces to deepfakes", *World Literature Studies*, Institute of world literature Slovak Academy of Sciences, Vol. 11 No. 4, pp. 11-25, available at: <http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-d4691768-ac82-4539-a63e-dce56ff09c27> (accessed 25 November 2021).
- Cross, C. (2022), "Using artificial intelligence (AI) and deepfakes to deceive victims: the need to rethink current romance fraud prevention messaging", *Crime Prevention and Community Safety*, Vol. 24, pp. 1-12.
- Crystal, D.T., Cuccolo, N.G., Ibrahim, A.M.S., Furnas, H. and Lin, S.J. (2020), "Photographic and video deepfakes have arrived: how machine learning may influence plastic surgery", *Plastic and Reconstructive Surgery*, Vol. 145 No. 4, pp. 1079-1086.
- Daly, C. (2022a), "How AI helped Top Gun fans hear Val Kilmer's voice again after throat cancer battle", *Daily Star*, available at: [https://www.dailystar.co.uk/tech/news/how-ai-helped-top-gun-27123647?int\\_source=mantis\\_rec&int\\_medium=web&int\\_campaign=more\\_like\\_this](https://www.dailystar.co.uk/tech/news/how-ai-helped-top-gun-27123647?int_source=mantis_rec&int_medium=web&int_campaign=more_like_this) (accessed 5 June 2022).

- 
- Daly, C. (2022b), "Creepy robot 'replaces' parents and steals their voice - to read kids bedtime stories - daily Star", *Daily Star*, available at: <https://www.dailystar.co.uk/tech/news/creepy-robot-replaces-parents-steals-27129961> (accessed 5 June 2022).
- Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., van der Linden, S. and von Sikorski, C. (2021), "Visual mis- and disinformation, social media, and democracy", *Journalism and Mass Communication Quarterly*, SAGE Publications, Vol. 98 No. 3, pp. 641-664.
- Darko (2021), "The rise of deepfake clothing", *Indie Hackers*, available at: [https://www.indiehackers.com/post/the-rise-of-deepfake-clothing-4519753717?utm\\_source=indie-hackers-emails&utm\\_campaign=ih-newsletter&utm\\_medium=email](https://www.indiehackers.com/post/the-rise-of-deepfake-clothing-4519753717?utm_source=indie-hackers-emails&utm_campaign=ih-newsletter&utm_medium=email) (accessed 23 July 2022).
- Dasilva, J.P., Ayerdi, K.M. and Galdospin, T.M. (2021), "Deepfakes on twitter: which actors control their spread?", *Media and Communication*, Vol. 9 No. 1, pp. 301-312.
- de Araújo e Silva, R., Santa Brígida, A.I., de Freitas Rosa, M., da Silva Neto, R.M., Spinosa, W.A., Benício de Sá Filho, E. and Brito de Figueirêdo, M.C. (2020), "An approach for implementing ecodesign at early research stage: a case study of bacterial cellulose production", *Journal of Cleaner Production*, Vol. 269, p. 122245.
- de Beer, D. and Matthee, M. (2021), "Approaches to identify fake news: a systematic literature review", in Antipova, T. (Ed.), *Integrated Science in Digital Age 2020*, Springer International Publishing, Cham, pp. 13-22.
- de Rancourt-Raymond, A. and Smaili, N. (2022), "The unethical use of deepfakes", *Journal of Financial Crime*, Vol. ahead-of-print, doi: [10.1108/JFC-04-2022-0090](https://doi.org/10.1108/JFC-04-2022-0090).
- de Ruiter, A. (2021), "The distinct wrong of deepfakes", *Philosophy and Technology*, Vol. 34 No. 4, pp. 1311-1332, doi: [10.1007/s13347-021-00459-2](https://doi.org/10.1007/s13347-021-00459-2).
- de Seta, G. (2021), "Huanlian, or changing faces: deepfakes on Chinese digital media platforms", *Convergence*, Vol. 27, doi: [10.1177/13548565211030185](https://doi.org/10.1177/13548565211030185).
- de Vries, K. (2020), "You never fake alone. Creative AI in action", *Information, Communication and Society*, Routledge, Vol. 23 No. 14, pp. 2110-2127.
- Delfino, R. (2019), "Pornographic deepfakes: the case for federal criminalization of revenge porn's next tragic act", *Fordham Law Review*, Vol. 88 No. 3, p. 887.
- Di Domenico, G. and Visentin, M. (2020), "Fake news or true lies? Reflections about problematic contents in marketing", *International Journal of Market Research*, Vol. 62 No. 4, pp. 409-417.
- Di Domenico, G., Nunan, D., Sit, J. and Pitardi, V. (2021a), "Free but fake speech: when giving primacy to the source decreases misinformation sharing on social media", *Psychology and Marketing*, Vol. 38 No. 10, pp. 1700-1711.
- Di Domenico, G., Tuan, A. and Visentin, M. (2021b), "Linguistic drivers of misinformation diffusion on social media during the COVID-19 pandemic", *Italian Journal of Marketing*, Vol. 2021 No. 4, pp. 351-369.
- Ding, J.-J., Hsu, H.-W. and Huang, C.-W. (2021), "Polling mechanism for video deepfake", *Presented at the 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE)*, pp. 320-323.
- Dixon, H.B. Jr (2019), "Deepfakes: more frightening than photoshop on steroids", *American Bar Association*, available at: [https://www.americanbar.org/groups/judicial/publications/judges\\_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/](https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/) (accessed 12 June 2022).
- Dobber, T., Metoui, N., Trilling, D., Helberger, N. and de Vreese, C. (2021), "Do (microtargeted) deepfakes have real effects on political attitudes?", *The International Journal of Press/Politics*, SAGE Publications, Vol. 26 No. 1, pp. 69-91.
- Doffman, Z. (2019), "Chinese deepfake app ZAO goes viral, privacy of millions 'at risk'", *Forbes*, available at: <https://www.forbes.com/sites/zakdoeffman/2019/09/02/chinese-best-ever-deepfake-app-zao-sparks-huge-faceapp-like-privacy-storm/> (accessed 21 June 2022).
- Domenico, G.D., Sit, J., Ishizaka, A. and Nunan, D. (2021), "Fake news, social media and marketing: a systematic review", *Journal of Business Research*, Vol. 124, pp. 329-341.

- 
- Dowdeswell, T.L. and Goltz, N. (2020), "The clash of empires: regulating technological threats to civil society", *Information and Communications Technology Law*, Routledge, Vol. 29 No. 2, pp. 194-217.
- Dutton, W.H., Cheong, P.H. and Park, N. (2004), "The social shaping of a virtual learning environment: the case of a university-wide course management system", *Electronic Journal of E-Learning*, Academic Conferences, Vol. 2 No. 1, pp. 69-80.
- E&T (2022), "Tougher criminal penalties proposed for 'downblousing' and porn deepfakes", available at: <https://eandt.theiet.org/content/articles/2022/07/tougher-criminal-penalties-proposed-for-downblousing-and-porn-deepfakes/> (accessed 14 September 2022).
- Eelmaa, S. (2022), "Sexualization of children in Deepfakes and hentai", *Trames. Journal of the Humanities and Social Sciences*, Vol. 26, p. 229.
- Elliott, V. and Tobin, M. (2022), "China steps up efforts to ban deepfakes. Will it work?", *Rest of World*, available at: <https://restofworld.org/2022/china-steps-up-efforts-to-ban-deepfakes/> (accessed 14 September 2022).
- European Journal of Public Health (2020), "What happened to the Enlightenment?", *European Journal of Public Health*, Vol. 30 Supplement\_5, ckaa164.001.
- Fagni, T., Falchi, F., Gambini, M., Martella, A. and Tesconi, M. (2021), "TweepFake: about detecting deepfake tweets", *PLOS One*, Public Library of Science, Vol. 16 No. 5, p. e0251415.
- Fallis, D. (2020), "The epistemic threat of deepfakes", *Philosophy and Technology*, Vol. 34 No. 4, pp. 623-643, doi: [10.1007/s13347-020-00419-2](https://doi.org/10.1007/s13347-020-00419-2).
- Fan, L., Li, W. and Cui, X. (2021), "Deepfake-image anti-forensics with adversarial examples attacks", *Future Internet*, Multidisciplinary Digital Publishing Institute, Vol. 13 No. 11, p. 288.
- Farish, K. (2020), "Do deepfakes pose a golden opportunity? Considering whether English law should adopt California's publicity right in the age of the deepfake", *Journal of Intellectual Property Law and Practice*, Vol. 15 No. 1, pp. 40-48.
- Farokhmanesh, M. (2018), "Deepfakes are disappearing from parts of the web, but they're not going away", *The Verge*, available at: <https://www.theverge.com/2018/2/9/16986602/deepfakes-banned-reddit-ai-faceswap-porn> (accessed 21 June 2022).
- Fido, D., Rao, J. and Harper, C.A. (2022), "Celebrity status, sex, and variation in psychopathy predicts judgements of and proclivity to generate and distribute deepfake pornography", *Computers in Human Behavior*, Vol. 129, p. 107141.
- Fikse, T.D. (2018), "Imagining Deceptive Deepfakes: an ethnographic exploration of fake videos", available at: <https://www.duo.uio.no/handle/10852/66387> (accessed 12 June 2022).
- Flynn, A., Powell, A., Scott, A.J. and Cama, E. (2021), "Deepfakes and digitally altered imagery abuse: a cross-country exploration of an emerging form of image-based sexual abuse", *The British Journal of Criminology*, Vol. 62 No. 6, pp. 1341-1358, azab111.
- Frith, J. and Wilken, R. (2019), "Social shaping of mobile geomeia services: an analysis of Yelp and Foursquare", *Communication and the Public*, SAGE Publications, Vol. 4 No. 2, pp. 133-149.
- Galyashina, E.I. and Nikishin, V.D. (2022), "The protection of megascience projects from deepfake technologies threats: information law aspects", *Journal of Physics: Conference Series*, IOP Publishing, Vol. 2210 No. 1, p. 012007.
- Gamage, D., Ghasiya, P., Bonagiri, V., Whiting, M.E. and Sasahara, K. (2022), "Are deepfakes concerning? Analyzing conversations of deepfakes on Reddit and exploring societal implications", *CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, pp. 1-19.
- Gandhi, A. (2021), "Deepfakes: fooling humans with artificial intelligence", *Illumin Magazine*, available at: <https://illumin.usc.edu/deepfakes-fooling-humans-with-artificial-intelligence/> (accessed 21 June 2022).
- García Lozano, M., Brynielsson, J., Franke, U., Rosell, M., Tjörnhammar, E., Varga, S. and Vlassov, V. (2020), "Veracity assessment of online data", *Decision Support Systems*, Vol. 129, p. 113132.

- García-Orosa, B. (2021), "Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy", *Profesional de La Información*, Vol. 30 No. 6, doi: [10.3145/epi.2021.nov.03](https://doi.org/10.3145/epi.2021.nov.03).
- García-Ull, F.J. (2021), *Deepfakes: The Next Challenge in Fake News Detection*, Universitat Autònoma de Barcelona, Anàlisi, pp. 103-120.
- Gasser, L. (1986), "The integration of computing and routine work", *ACM Transactions on Information Systems*, Vol. 4 No. 3, pp. 205-225.
- Ghazi-Tehrani, A.K. and Pontell, H.N. (2021), "Phishing evolves: analyzing the enduring cybercrime", *Victims and Offenders*, Routledge, Vol. 16 No. 3, pp. 316-342.
- Gibson, K. (2020), "Deepfakes and involuntary pornography: can our current legal framework address this technology?", *Wayne Law Review*, Vol. 66, p. 259.
- Godulla, A., Hoffmann, C. and Seibert, D. (2021), "Dealing with deepfakes - an interdisciplinary examination of the state of research and implications for communication studies", *Studies in Communication and Media*, Vol. 10, doi: [10.5771/2192-4007-2021-1-72](https://doi.org/10.5771/2192-4007-2021-1-72).
- Goltz, N. and Dowdeswell, T. (2021), "'The law of the dog' — emotion and motivation in teaching children about new media", *Presented at the 2021 IEEE Conference on Norbert Wiener in the 21st Century (21CW)*, pp. 1-5.
- Gosse, C. and Burkell, J. (2020), "Politics and porn: how news media characterizes problems presented by deepfakes", *Critical Studies in Media Communication*, Routledge, Vol. 37 No. 5, pp. 497-511.
- Gow, G. (2020), "Deepfakes are bad—deepfakes are good", *Forbes*, available at: <https://www.forbes.com/sites/glenngow/2020/09/11/deepfakes-are-baddeepfakes-are-good/> (accessed 25 June 2022).
- Greengard, S. (2019), "Will deepfakes do deep damage?", *Communications of the ACM*, Vol. 63 No. 1, pp. 17-19.
- Gregory, S. (2019), "Cameras everywhere revisited: how digital technologies and social media aid and inhibit human rights documentation and advocacy", *Journal of Human Rights Practice*, Vol. 11 No. 2, pp. 373-392.
- Gregory, S. (2022), *Deepfakes, Misinformation and Disinformation and Authenticity Infrastructure Responses: Impacts on Frontline Witnessing, Distant Witnessing, and Civic Journalism*, SAGE Publications, Vol. 23 No. 3, pp. 708-729.
- Griffin, M. (2019), "Edtech company Udacity uses deepfake tech to create educational videos automatically", *Futurist and Virtual Keynote Speaker Matthew Griffin*, available at: <https://www.fanaticalfuturist.com/2019/08/edtech-company-udacity-uses-deepfake-tech-to-create-educational-videos-automatically/> (accessed 15 April 2022).
- Groh, M., Epstein, Z., Obradovich, N., Cebrian, M. and Rahwan, I. (2019), "Human detection of machine manipulated media", No. arXiv:1907.05276, arXiv, doi: [10.48550/arXiv.1907.05276](https://doi.org/10.48550/arXiv.1907.05276).
- Güera, D. and Delp, E.J. (2018), "Deepfake video detection using recurrent neural networks", *Presented at the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1-6.
- Guarnera, L., Giudice, O., Nastasi, C. and Battiato, S. (2020), "Preliminary forensics analysis of DeepFake images", *2020 AEIT International Annual Conference (AEIT)*, pp. 1-6.
- Guo, L. (2020), "China's 'fake news' problem: exploring the spread of online rumors in the government-controlled news media", *Digital Journalism*, Routledge, Vol. 8 No. 8, pp. 992-1010.
- Han, B., Han, X., Zhang, H., Li, J. and Cao, X. (2021), "Fighting fake news: two stream network for deepfake detection via learnable SRM", *IEEE Transactions on Biometrics, Behavior, and Identity Science*, Presented at the IEEE Transactions on Biometrics, Behavior, and Identity Science, Vol. 3 No. 3, pp. 320-331.
- Hancock, J.T. and Bailenson, J.N. (2021), "The social impact of deepfakes", *Cyberpsychology, Behavior, and Social Networking*, Mary Ann Liebert, Vol. 24 No. 3, pp. 149-152.
- Harper, C.A., Fido, D. and Petronzi, D. (2021), "Delineating non-consensual sexual image offending: towards an empirical approach", *Aggression and Violent Behavior*, Vol. 58, p. 101547.

- 
- Harris, K.R. (2021), "Video on demand: what deepfakes do and how they harm", *Synthese*, Vol. 199 Nos 5-6, pp. 13373-13391, doi: [10.1007/s11229-021-03379-y](https://doi.org/10.1007/s11229-021-03379-y).
- Hawkins, D. (2018), "Reddit bans 'deepfakes,' pornography using the faces of celebrities", *AP NEWS*, available at: <https://apnews.com/article/2bc8184c30684830b3779fd3c6883571> (accessed 14 September 2022).
- Hern, A. (2018a), "AI used to face-swap Hollywood stars into pornography films", *The Guardian*, available at: <https://www.theguardian.com/technology/2018/jan/25/ai-face-swap-pornography-emma-watson-scarlett-johansson-taylor-swift-daisy-ridley-sophie-turner-maisie-williams> (accessed 21 June 2022).
- Hern, A. (2018b), "Reddit bans 'deepfakes' face-swap porn community", *The Guardian*, available at: <https://www.theguardian.com/technology/2018/feb/08/reddit-bans-deepfakes-face-swap-porn-community> (accessed 21 June 2022).
- Hight, C. (2022), *Deepfakes and Documentary Practice in an Age of Misinformation*, Routledge, Continuum, Vol. 36 No. 3, pp. 393-410.
- Himma-Kadakas, M. and Ojamets, I. (2022), "Debunking false information: investigating journalists' fact-checking skills", *Digital Journalism*, Routledge, Vol. 10 No. 5, pp. 866-887.
- Hirlekar, V.V. and Kumar, A. (2022), "Tweet credibility detection for COVID-19 tweets using text and user content features", *International Journal of Advanced Computer Science and Applications (IJACSA)*, The Science and Information (SAI) Organization, Vol. 13 No. 4, doi: [10.14569/IJACSA.2022.0130451](https://doi.org/10.14569/IJACSA.2022.0130451).
- Hodge, S. (2021), "Don't always believe what you see: shallowfake and deepfake media has altered the perception of reality", *Hofstra Law Review*, Vol. 50 No. 1, pp. 51-80, available at: <https://scholarlycommons.law.hofstra.edu/hlr/vol50/iss1/4>
- Holliday, C. (2021), "Rewriting the stars: surface tensions and gender troubles in the online media production of digital deepfakes", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 899-918.
- Hoon, C. (2013), "Meta-synthesis of qualitative case studies: an approach to theory building", *Organizational Research Methods*, SAGE Publications, Vol. 16 No. 4, pp. 522-556.
- Horton, L. (2020), "Making qualitative data more visible in policy: a critical appraisal of meta-synthesis", *Qualitative Research*, SAGE Publications, Vol. 20 No. 5, pp. 534-548.
- Howcroft, D. and Light, B. (2010), "The social shaping of packaged software selection", *Journal of the Association for Information Systems*, Vol. 11 No. 3, doi: [10.17705/1jais.00224](https://doi.org/10.17705/1jais.00224).
- Howcroft, D. and Taylor, P. (2022), "Automation and the future of work: a social shaping of technology approach", *New Technology, Work and Employment*, Vol. ahead-of-print, doi: [10.1111/ntwe.12240](https://doi.org/10.1111/ntwe.12240).
- Huang, W. and Yu, Y. (2016), "Is it truly a 5-star movie? Restoring the movie's truthful rating", *Presented at the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 1337-1338.
- Huijstee, M.V., Boheemen, P.V., Das, D., Nierling, L., Jahnel, J., Karaboga, M. and Fatun, M. (2021), "Tackling deepfakes in European policy | Think tank | European parliament", The European Parliament, available at: [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_STU\(2021\)690039](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2021)690039) (accessed 23 August 2022).
- Hunt, M. (1997), *How Science Takes Stock: The Story of Meta-Analysis*, Russell Sage Foundation.
- Hunter, J., Schmidt, F. and Jackson, G. (1982), "Meta-analysis: cumulating research findings across studies", *The Academy of Management Review*, Vol. 9, doi: [10.2307/258247](https://doi.org/10.2307/258247).
- Hwang, Y., Ryu, J.Y. and Jeong, S.-H. (2021), "Effects of disinformation using deepfake: the protective effect of media literacy education", *Cyberpsychology, Behavior, and Social Networking*, Mary Ann Liebert, Vol. 24 No. 3, pp. 188-193.
- Iacobucci, S., De Cicco, R., Michetti, F., Palumbo, R. and Pagliaro, S. (2021), "Deepfakes unmasked: the effects of information priming and bullshit receptivity on deepfake recognition and sharing intention", *Cyberpsychology, Behavior and Social Networking*, Vol. 24 No. 3, pp. 194-202.



- 
- Ice, J. (2019), "Defamatory political deepfakes and the first amendment", *Case Western Reserve Law Review*, Vol. 70 No. 2, p. 417.
- Internet Research (2022), "The social, ethical, economic, and political implications of misinformation", Emerald Publishing | Internet Research, available at: <https://www.emeraldgrouppublishing.com/calls-for-papers/social-ethical-economic-and-political-implications-misinformation> (accessed 26 June 2022).
- Jafar, M.T., Ababneh, M., Al-Zoube, M. and Elhassan, A. (2020), "Forensics and analysis of deepfake videos", *Presented at the 2020 11th International Conference on Information and Communication Systems (ICICS)*, pp. 053-058.
- Jang, S.M. and Kim, J.K. (2018), "Third person effects of fake news: fake news regulation and media literacy interventions", *Computers in Human Behavior*, Vol. 80, pp. 295-302.
- Jaynes, T.L. (2021), "The question of algorithmic personhood and being (or: on the tenuous nature of human status and humanity tests in virtual spaces—why all souls are 'necessarily' equal when considered as energy)", *J-Multidisciplinary Scientific Journal*, Multidisciplinary Digital Publishing Institute, Vol. 4 No. 3, pp. 452-475.
- Jørgensen, M.S., Jørgensen, U. and Clausen, C. (2009), "The social shaping approach to technology foresight", *Futures*, Vol. 41 No. 2, pp. 80-86.
- Johnson, J. (2022), "Delegating strategic decision-making to machines: Dr. Strangelove Redux?", *Journal of Strategic Studies*, Routledge, Vol. 45 No. 3, pp. 439-477.
- Johnson, D.G. and Diakopoulos, N. (2021), "What to do about deepfakes", *Communications of the ACM*, Vol. 64 No. 3, pp. 33-35.
- Jones-Jang, S.M., Mortensen, T. and Liu, J. (2021), "Does media literacy help identification of fake news? Information literacy helps, but other literacies don't", *American Behavioral Scientist*, SAGE Publications, Vol. 65 No. 2, pp. 371-388.
- Joseph, R. (2019), "Fakebusters strike back: how to spot deep fakes, the manipulated videos that are the newest form of 'fake news' to hit the internet", *Index on Censorship*, SAGE Publications, Vol. 48 No. 1, pp. 76-79.
- Juefei-Xu, F., Wang, R., Huang, Y., Guo, Q., Ma, L. and Liu, Y. (2022), "Countering malicious DeepFakes: survey, battleground, and horizon", *International Journal of Computer Vision*, Vol. 130 No. 7, pp. 1678-1734, doi: [10.1007/s11263-022-01606-8](https://doi.org/10.1007/s11263-022-01606-8).
- Kabelka, L. (2022), "Austria to combat deep fakes amid increasing use of the technology", available at: <https://www.euractiv.com/section/disinformation/news/austria-to-combat-deep-fakes-amid-increasing-use-of-the-technology/> (accessed 5 June 2022).
- Kalpokas, I. and Kalpokiene, J. (2022a), "On alarmism: between infodemic and epistemic anarchy", in Kalpokas, I. and Kalpokiene, J. (Eds), *Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation*, Springer International Publishing, Cham, pp. 41-53.
- Kalpokas, I. and Kalpokiene, J. (2022b), "Fake news: exploring the backdrop", in Kalpokas, I. and Kalpokiene, J. (Eds), *Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation*, Springer International Publishing, Cham, pp. 7-17.
- Kalpokas, I. and Kalpokiene, J. (2022c), "From GANs to deepfakes: getting the characteristics right", in Kalpokas, I. and Kalpokiene, J. (Eds), *Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation*, Springer International Publishing, Cham, pp. 29-39.
- Karasavva, V. and Noorbhai, A. (2021), "The real threat of deepfake pornography: a review of Canadian policy", *Cyberpsychology, Behavior and Social Networking*, Vol. 24 No. 3, pp. 203-209.
- Karnouskos, S. (2020), "Artificial intelligence in digital media: the era of deepfakes", *Presented at the IEEE Transactions on Technology and Society*, Vol. 1 No. 3, pp. 138-147.
- Kaur, P., Dhir, A., Alkhalifa, A.K. and Tandon, A. (2021), "Social media platforms and sleep problems: a systematic literature review, synthesis and framework for future research", *Internet Research*, Emerald Publishing, Vol. 31 No. 4, pp. 1121-1152.

- 
- Kawa, P. and Syga, P. (2020), "A note on deepfake detection with low-resources", *arXiv*. doi: [10.48550/arXiv.2006.05183](https://doi.org/10.48550/arXiv.2006.05183).
- Kearney, V.P., Yansane, A.-I.M., Brandon, R.G., Vaderhobli, R., Lin, G.-H., Hekmatian, H., Deng, W., Joshi, N., Bhandari, H., Sadat, A.S. and White, J.M. (2022), "A generative adversarial inpainting network to enhance prediction of periodontal clinical attachment level", *Journal of Dentistry*, Vol. 123, 104211.
- Kelly, G. (2021), "Starring tom cruise", *Historical Journal of Film, Radio and Television*, Routledge, Vol. 41 No. 4, pp. 872-874.
- Kerner, C. and Risse, M. (2021), "Beyond porn and discreditation: epistemic promises and perils of deepfake technology in digital lifeworlds", *Moral Philosophy and Politics*, De Gruyter, Vol. 8 No. 1, pp. 81-108.
- Khan, A. and Krishnan, S. (2021), "Citizen engagement in co-creation of e-government services: a process theory view from a meta-synthesis approach", *Internet Research*, Emerald Publishing, Vol. 31 No. 4, pp. 1318-1375.
- Khichi, M. and Kumar Yadav, R. (2021), "A threat of deepfakes as a weapon on digital platform and their detection methods", *Presented at the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-8.
- Kietzmann, J., Lee, L.W., McCarthy, I.P. and Kietzmann, T.C. (2020), "Deepfakes: trick or treat?", *Business Horizons*, Vol. 63 No. 2, pp. 135-146.
- Kietzmann, J., Mills, A.J. and Plangger, K. (2021), "Deepfakes: perspectives on the future 'reality' of advertising and branding", *International Journal of Advertising*, Routledge, Vol. 40 No. 3, pp. 473-485.
- Kikerpill, K. (2020), "Choose your stars and studs: the rise of deepfake designer porn", *Porn Studies*, Vol. 7, doi: [10.1080/23268743.2020.1765851](https://doi.org/10.1080/23268743.2020.1765851).
- Kim, J., Han, S. and S.Woo, S. (2019), "Classifying genuine face images from disguised face images", *Presented at the 2019 IEEE International Conference on Big Data (Big Data)*, pp. 6248-6250.
- Kirchengast, T. (2020), "Deepfakes and image manipulation: criminalisation and control", *Information and Communications Technology Law*, Routledge, Vol. 29 No. 3, pp. 308-323.
- Köbis, N.C., Doležalová, B. and Soraperra, I. (2021), "Fooled twice: people cannot detect deepfakes but think they can", *IScience*, Vol. 24 No. 11, p. 103364.
- Kobielus, J. (2020), "Deepfake technology loses its stigma amid socially redeeming uses", available at: <https://www.linkedin.com/pulse/deepfake-technology-loses-its-stigma-amid-socially-uses-kobielus> (accessed 14 September 2022).
- Korshunov, P. and Marcel, S. (2019), "Vulnerability assessment and detection of Deepfake videos", *Presented at the 2019 International Conference on Biometrics (ICB)*, pp. 1-6.
- Krishna, D. (2021), "Deepfakes, online platforms, and a novel proposal for transparency, collaboration, and education", *Richmond Journal of Law and Technology*, Vol. 27 No. 3, p. 74.
- Kugler, M.B. and Pace, C. (2021), "Deepfake privacy: attitudes and regulation", SSRN Scholarly Paper No. 3781968, Social Science Research Network, Rochester, NY, doi: [10.2139/ssrn.3781968](https://doi.org/10.2139/ssrn.3781968).
- Kwok, A.O.J. and Koh, S.G.M. (2021), "Deepfake: a social construction of technology perspective", *Current Issues in Tourism*, Routledge, Vol. 24 No. 13, pp. 1798-1802.
- Langa, J. (2021), "Deepfakes, real consequences: crafting legislation to combat threats posed by Deepfakes", *Boston University Law Review*, Vol. 101, p. 41.
- Langguth, J., Pogorelov, K., Brenner, S., Filkkuková, P. and Schroeder, D.T. (2021), "Don't trust your eyes: image manipulation in the age of DeepFakes", *Frontiers in Communication*, Vol. 6, p. 26.
- Lazazzara, A., Tims, M. and de Gennaro, D. (2020), "The process of reinventing a job: a meta-synthesis of qualitative job crafting research", *Journal of Vocational Behavior*, Vol. 116, p. 103267.

- 
- Lee, Y., Huang, K.-T.T., Blom, R., Schriener, R. and 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*, Vol. 24 No. 3, pp. 153-158.
- Lee, J. and Shin, S.Y. (2022), "Something that they never said: multimodal disinformation and source vividness in understanding the power of AI-enabled deepfake news", *Media Psychology*, Routledge, Vol. 25 No. 4, pp. 531-546.
- Lees, D., Bashford-Rogers, T. and Keppel-Palmer, M. (2021), "The digital resurrection of Margaret Thatcher: creative, technological and legal dilemmas in the use of deepfakes in screen drama", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 954-973.
- Leibowicz, C.R., McGregor, S. and Ovadya, A. (2021), "The deepfake detection dilemma: a multistakeholder exploration of adversarial dynamics in synthetic media", *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery, New York, NY, pp. 736-744.
- Lemley, M.A. and Casey, B. (2020), "Fair learning", SSRN Scholarly Paper, Rochester, NY, doi: [10.2139/ssrn.3528447](https://doi.org/10.2139/ssrn.3528447).
- Lin, J., Tremblay-Taylor, G., Mou, G., You, D. and Lee, K. (2019), "Detecting fake news articles", *Presented at the 2019 IEEE International Conference on Big Data (Big Data)*, pp. 3021-3025.
- López, J.F.R. (2022), "Tragic realism: how to regulate deepfakes in Colombia?\*", *Latin American Law Review*, Universidad de los Andes, No. 8, pp. 125-145, doi: [10.29263/lar08.2022.08](https://doi.org/10.29263/lar08.2022.08).
- Locke, K. (2001), *Grounded Theory in Management Research*, SAGE.
- MacKenzie, D. and Wajcman, J. (1999), *The Social Shaping of Technology*, Open University Press, Buckingham, available at: <http://mcgraw-hill.co.uk/openup/> (accessed 19 June 2022).
- Mackey, A. and Gass, S.M. (2015), *Second Language Research: Methodology and Design*, 2nd ed., Routledge, New York, doi: [10.4324/9781315750606](https://doi.org/10.4324/9781315750606).
- Maddalena, G. and Gili, G. (2020), *The History and Theory of Post-truth Communication*, Palgrave Macmillan.
- Maddocks, S. (2020), "'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political' deep fakes", *Porn Studies*, Routledge, Vol. 7 No. 4, pp. 415-423.
- Major, C. and Savin-Baden, M. (2012), *An Introduction to Qualitative Research Synthesis: Managing the Information Explosion in Social Science Research*, doi: [10.4324/9780203497555](https://doi.org/10.4324/9780203497555).
- Malik, Y.S., Sabahat, N. and Moazzam, M.O. (2020), "Image animations on driving videos with DeepFakes and detecting DeepFakes generated animations", *Presented at the 2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 1-6.
- Mangaokar, N. and Prakash, A. (2022), "Dispelling misconceptions and characterizing the failings of deepfake detection", *Presented at the IEEE Security and Privacy*, Vol. 20 No. 2, pp. 61-67.
- Mbinjama-Gamatham, A. and Olivier, B. (2020), "'Dark technology', aggressiveness and the question of cyber-ethics", *Acta Academica: Critical Views on Society, Culture and Politics*, Vol. 52 No. 1, pp. 99-120.
- McCosker, A. (2022), "Making sense of deepfakes: socializing AI and building data literacy on GitHub and YouTube", *New Media and Society*, SAGE Publications, Vol. ahead-of-print, p. 14614448221093944.
- McDowell, M. (2021), "Influencers are wearing digital versions of physical clothes now", *Vogue Business*, available at: <https://www.voguebusiness.com/technology/influencers-are-wearing-digital-versions-of-physical-clothes-now> (accessed 23 July 2022).
- McPeak, A. (2021), "The threat of deepfakes in litigation: raising the authentication bar to combat falsehood", *Vanderbilt Journal of Entertainment and Technology Law*, Vol. 23 No. 2, pp. 433-450, available at: <https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss2/5>
- Meskys, E., Liaudanskas, A., Kalpokiene, J. and Jurcys, P. (2020), "Regulating deep fakes: legal and ethical considerations", *Journal of Intellectual Property Law and Practice*, Vol. 15 No. 1, pp. 24-31.

- 
- Mihailidis, P. and Viotty, S. (2017), "Spreadable spectacle in digital culture: civic expression, fake news, and the role of media literacies in 'post-fact' society", *American Behavioral Scientist*, SAGE Publications, Vol. 61 No. 4, pp. 441-454.
- Mihailova, M. (2021), "To dally with dalí: deepfake (Inter)faces in the art museum", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 882-898.
- Miles, M.B. and Huberman, A.M. (1994), *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed., Sage Publications, Thousand Oaks, CA, pp. xiv-338.
- Millière, R. (2022), "Deep learning and synthetic media", *Synthese*, Springer Verlag, Vol. 200 No. 3, pp. 1-27.
- Mirsky, Y. and Lee, W. (2021), "The creation and detection of deepfakes: a survey", *ACM Computing Surveys*, Vol. 54 No. 1, pp. 7:1-7:41.
- Modern Diplomacy (2022), "Triumph of simulacra – how deepfakes aim to rule our minds", *Modern Diplomacy*, available at: <https://moderndiplomacy.eu/2022/05/19/triumph-of-simulacra-how-deepfakes-aim-to-rule-our-minds/> (accessed 5 June 2022).
- Mone, G. (2022), "Raising robovoices", *Communications of the ACM*, Vol. 65 No. 5, pp. 30-31.
- Moon, S. and Iacobucci, D. (2022), "Social media analytics and its applications in marketing", *Foundations and Trends® in Marketing*, Now Publishers, Vol. 15 No. 4, pp. 213-292.
- Morgan, R. (2022), "Can anything stop deepfake porn?", *Morning Brew*, available at: <https://www.morningbrew.com/daily/stories/2022/07/29/can-anything-stop-deepfake-porn> (accessed 23 August 2022).
- Moss-Wellington, W., Lam, C. and Gilardi, F. (2022), "Introduction: media and fakery", *Continuum*, Routledge, Vol. 36 No. 3, pp. 333-338.
- Mullen, M. (2022), "A new reality: deepfake technology and the world around us", *Mitchell Hamline Law Review*, Vol. 48 No. 1, pp. 210-234, available at: <https://open.mitchellhamline.edu/mhlr/vol48/iss1/5>
- Murphy, G. and Flynn, E. (2022), "Deepfake false memories", *Memory*, Routledge, Vol. 30 No. 4, pp. 480-492.
- Murray, S. (2020), "Postdigital cultural studies", *International Journal of Cultural Studies*, SAGE Publications, Vol. 23 No. 4, pp. 441-450.
- Muza, O. and Debnath, R. (2021), "Disruptive innovation for inclusive renewable policy in sub-Saharan Africa: a social shaping of technology analysis of appliance uptake in Rwanda", *Renewable Energy*, Vol. 168, pp. 896-912.
- Nazar, S. and Bustam, M.R. (2020), "Artificial intelligence and new level of fake news", *IOP Conference Series: Materials Science and Engineering*, Vol. 879 No. 1, p. 012006.
- Newswise (2022), "Synthetic medical imaging: how deepfakes could improve healthcare", available at: <https://www.newswise.com/articles/synthetic-medical-imaging-how-deepfakes-could-improve-healthcare> (accessed 27 August 2022).
- Newton, O.B. and Stanfill, M. (2020), "My NSFW video has partial occlusion: deepfakes and the technological production of non-consensual pornography", *Porn Studies*, Routledge, Vol. 7 No. 4, pp. 398-414.
- Nguyen, T.T., Nguyen, Q.V.H., Nguyen, D.T., Nguyen, D.T., Huynh-The, T., Nahavandi, S., Nguyen, T.T., Pham, Q.V. and Nguyen, C.M. (2022), "Deep learning for deepfakes creation and detection: a survey", *Computer Vision and Image Understanding*, Vol. 223, 103525.
- Noblit, G.W. and Hare, R.D. (1988), *Meta-Ethnography: Synthesizing Qualitative Studies*, Sage Publications, Newbury Park, CA.
- Nour, N. and Gelfand, J. (2022), "Deepfakes: a digital transformation leads to misinformation", *GL- Conference Series: Conference Proceedings*, p. 11.
- Öhman, C. (2020), "Introducing the pervert's dilemma: a contribution to the critique of Deepfake Pornography", *Ethics and Information Technology*, Vol. 22 No. 2, pp. 133-140.

- Orlikowski, W.J. and Gash, D.C. (1994), "Technological frames: making sense of information technology in organizations", *ACM Transactions on Information Systems*, Vol. 12 No. 2, pp. 174-207.
- Outlook Money (2022), "Criminals use elon musk's deepfake video to dupe crypto investors, crypto market rises", available at: <https://www.Outlookindia.Com/https://www.outlookindia.com/business/criminals-use-elon-musk-s-deepfake-video-to-dupe-crypto-investors-crypto-market-rises-news-198403> (accessed 5 June 2022).
- O'Connell, A. (2020), "Image rights and image wrongs: image-based sexual abuse and online takedown", *Journal of Intellectual Property Law and Practice*, Vol. 15 No. 1, pp. 55-65.
- O'Donnell, N. (2021), "Have we no decency? Section 230 and the liability of social media companies for DEEPFAKE videos", *University of Illinois Law Review*, Vol. 2021 No. 2, p. 40.
- Pandey, C.K., Mishra, V.K. and Tiwari, N.K. (2021), "Deepfakes: when to use it", *Presented at the 2021 10th International Conference on System Modeling and Advancement in Research Trends (SMART)*, pp. 80-84.
- Park, C. and McDowell, S.D. (2005), "Direct broadcast satellites and the social shaping of technology: comparing South Korea and Canada", *Canadian Journal of Communication*, University of Toronto Press, Vol. 30 No. 1, pp. 111-138.
- Partadiredja, R.A., Serrano, C.E. and Ljubenkov, D. (2020), "AI or human: the socio-ethical implications of AI-generated media content", *Presented at the 2020 13th CMI Conference on Cybersecurity and Privacy (CMI) - Digital Transformation - Potentials and Challenges (51275)*, pp. 1-6.
- Paterson, T. and Hanley, L. (2020), "Political warfare in the digital age: cyber subversion, information operations and 'deep fakes'", *Australian Journal of International Affairs*, Routledge, Vol. 74 No. 4, pp. 439-454.
- Pavis, M. (2021), "Rebalancing our regulatory response to Deepfakes with performers' rights", *Convergence*, SAGE Publications, Vol. 27 No. 4, pp. 974-998.
- Pechenik Gieseke, A. (2020), "The new weapon of choice': law's current inability to properly address deepfake pornography", *Vanderbilt Law Review*, Vol. 73 No. 5, p. 1479.
- Perot, E. and Mostert, F. (2020), "Fake it till you make it: an examination of the US and English approaches to persona protection as applied to deepfakes on social media", *Journal of Intellectual Property Law and Practice*, Vol. 15 No. 1, pp. 32-39.
- Pesetski, A. (2021), "Deepfakes: a new content category for a digital age", *William and Mary Bill of Rights Journal*, Vol. 29 No. 2, p. 503.
- Pfefferkorn, R. (2020), "DEEPFAKES' in the courtroom", *Public Interest Law Journal*, Vol. 29, p. 32.
- Popescu, M. (2021), "Disinformation deconstructed- cognition security and digital control", *Bulletin of the Transilvania University of Brasov*, Vol. 14 No. 63, pp. 123-134.
- Popova, M. (2019), "Reading out of context: pornographic deepfakes, celebrity and intimacy", *Porn Studies*, Taylor & Francis (Routledge), Vol. 7 No. 4, doi: [10.1080/23268743.2019.1675090](https://doi.org/10.1080/23268743.2019.1675090).
- Popova, M. (2020), "Deepfakes: an introduction", *Porn Studies*, Routledge, Vol. 7 No. 4, pp. 350-351.
- Porter, A. (2020), "Bioethics in the ruins", *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, Vol. 45 No. 3, pp. 259-276.
- Rao, S., Verma, A.K. and Bhatia, T. (2021), "A review on social spam detection: challenges, open issues, and future directions", *Expert Systems with Applications*, Vol. 186, p. 115742.
- Rasmussen, P.K.B. (2021), "Researching young people's sexualized digital practices involving imagery: a transmethodological approach", *Outlines*, Critical Practice Studies, Vol. 22, pp. 89-124.
- Ratner, C. (2021), "When 'sweetie' is not so sweet: artificial intelligence and its implications for child pornography", *Family Court Review*, Vol. 59 No. 2, pp. 386-401.
- Ray, A. (2021), "Disinformation, deepfakes and democracies: the need for legislative reform", *UNSW Law Journal*, Vol. 44 No. 3, pp. 983-1013, available at: <https://www.unswlawjournal.unsw.edu>.

- 
- [au/article/disinformation-deepfakes-and-democracies-the-need-for-legislative-reform/](#) (accessed 26 November 2021).
- Reddit (2020), “Do not impersonate an individual or entity”, *Reddit Help*, available at: <https://reddit.zendesk.com/hc/en-us/articles/360043075032-Do-not-impersonate-an-individual-or-entity> (accessed 14 September 2022).
- Reddit (2022), “Never post intimate or sexually explicit media of someone without their consent”, *Reddit Help*, available at: <https://reddit.zendesk.com/hc/en-us/articles/360043513411-Never-Post-Intimate-or-Sexually-Explicit-Media-of-Someone-Without-Their-Consent> (accessed 14 September 2022).
- Reid, S. (2021), “The deepfake dilemma: reconciling privacy and first amendment protections”, *University of Pennsylvania Journal of Constitutional Law*, Vol. 23 No. 1, p. 209.
- Reuters (2019), “China seeks to root out fake news and deepfakes with new online content rules”, *Reuters*, available at: <https://www.reuters.com/article/us-china-technology-idUSKBN1Y30VU> (accessed 21 June 2022).
- Riar, M., Xi, N., Korbel, J.J., Zarnekow, R. and Hamari, J. (2022), “Using augmented reality for shopping: a framework for AR induced consumer behavior, literature review and future agenda”, *Internet Research*, Vol. ahead-of-print, doi: 10.1108/INTR-08-2021-0611.
- Ridden, P. (2022), “Deepfake tool gives animated life stories of dead relatives a voice”, *New Atlas*, available at: <https://newatlas.com/computers/myheritage-livestory/> (accessed 5 June 2022).
- Rini, R. (2020), “Deepfakes and the epistemic backstop”, *Philosophers’ Imprint*, Vol. 20 No. 24, pp. 1-16.
- Ripoll, L., Matos, J.C., Ripoll, L. and Matos, J.C. (2020), “Information reliability: criteria to identify misinformation in the digital environment”, *Investigación Bibliotecológica, Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliotecológicas y de la Información*, Vol. 34 No. 84, pp. 79-101.
- Rohracher, H. (2003), “The role of users in the social shaping of environmental technologies”, *Innovation: The European Journal of Social Science Research*, Routledge, Vol. 16 No. 2, pp. 177-192.
- Rousseau, D.M., Manning, J. and Denyer, D. (2008), “11 evidence in management and organizational science: assembling the field’s full weight of scientific knowledge through syntheses”, *Academy of Management Annals*, Academy of Management, Vol. 2 No. 1, pp. 475-515.
- Russell, J.E., Singer, B.W., Perry, J.J. and Bacon, A. (2012), “The materials and techniques used in the paintings of Francis Bacon (1909-1992)”, *Studies in Conservation*, Routledge, Vol. 57 No. 4, pp. 207-217.
- Sabrina, B. (2022), “How do students engage with fake news and deepfakes in Korea?”, *부경대학교*, available at: <https://repository.pknu.ac.kr:8443/handle/2021.oak/24122> (accessed 13 June 2022).
- Saif, S. and Tehseen, S. (2022), “Deepfake videos: synthesis and detection techniques – a survey”, *Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology*, Vol. 42 No. 4, pp. 2989-3009.
- Sample, I. (2020), “What are deepfakes – and how can you spot them?”, *The Guardian*, available at: <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them> (accessed 21 June 2022).
- Sandelowski, M. and Barroso, J. (2007), *Handbook for Synthesizing Qualitative Research*, Springer Publishing Company.
- Sample, C., Jensen, M.J., Scott, K., McAlaney, J., Fitchpatrick, S., Brockinton, A., Ormrod, D. and Ormrod, A. (2020), “Interdisciplinary lessons learned while researching fake news”, *Frontiers in Psychology*, Vol. 11, p. 2947.
- Saxena, R. and Gayathri, E. (2022), “Cyber threat intelligence challenges: leveraging blockchain intelligence with possible solution”, *Materials Today: Proceedings*, Vol. 51, pp. 682-689.
- Schick, N. (2020), “Deep fakes and the infocalypse: what you urgently need to know”, *Monoray*, p. 224.
- Sensity (2019), “The state of deepfakes 2019 landscape, threats, and impact”, available at: [https://share.hsforms.com/1cg\\_h2aPnRrufZeN8HDjWPw3hq83](https://share.hsforms.com/1cg_h2aPnRrufZeN8HDjWPw3hq83)

- 
- Šepec, M. and Lango, M. (2020), "Virtual revenge pornography as a new online threat to sexual integrity", *Balkan Social Science Review*, Vol. 15, pp. 117-135.
- Sethi, L., Dave, A., Bhagwani, R. and Biwalkar, A. (2020), "Video security against deepfakes and other forgeries", *Journal of Discrete Mathematical Sciences and Cryptography*, Taylor & Francis, Vol. 23 No. 2, pp. 349-363.
- Sharma, A. (2020), "'We do not want fake energy': the social shaping of a solar micro-grid in rural India", *Science, Technology and Society*, SAGE Publications, Vol. 25 No. 2, pp. 308-324.
- Shelke, S. and Attar, V. (2022), "Rumor detection in social network based on user, content and lexical features", *Multimedia Tools and Applications*, Vol. 81 No. 12, pp. 17347-17368.
- Shin, S.Y. and Lee, J. (2022), "The effect of deepfake video on news credibility and corrective influence of cost-based knowledge about deepfakes", *Digital Journalism*, Routledge, Vol. 10 No. 3, pp. 412-432.
- Shin, H.-C., Tenenholtz, N.A., Rogers, J.K., Schwarz, C.G., Senjem, M.L., Gunter, J.L., Andriole, K.P. and Michalski, M. (2018), "Medical image synthesis for data augmentation and anonymization using generative adversarial networks", in Gooya, A., Goksel, O., Oguz, I. and Burgos, N. (Eds), *Simulation and Synthesis in Medical Imaging*, Springer International Publishing, Cham, Vol. 11037, pp. 1-11.
- Snow, J. (2021), "These historical artefacts are totally faked", *Wired UK*, available at: <https://www.wired.co.uk/article/fake-artefacts-ai> (accessed 27 August 2022).
- Sobreperez, P. (2012), "Hubble bubble toil and trouble: the special case of emergency services", in Dwivedi, Y.K., Wade, M.R. and Schneberger, S.L. (Eds), *Information Systems Theory: Explaining and Predicting Our Digital Society*, Springer, New York, NY, Vol. 2, pp. 143-157.
- Solo, A.M.G. (2019), "Combating online impersonation in the United States", *Presented at the 2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1552-1553.
- Somers, M. (2020), "Deepfakes, explained", *MIT Sloan*, available at: <https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained> (accessed 8 December 2021).
- Soral, P., Pati, S.P., Singh, S.K. and Cooke, F.L. (2022), "Coping with dirty work: a meta-synthesis from a resource perspective", *Human Resource Management Review*, Vol. 32 No. 4, pp. 1-21, 100861.
- Stadler, J. (2019), "Synthetic beings and synthespian ethics: embodiment technologies in science/fiction", *Projections, Berghahn Journals*, Vol. 13 No. 2, pp. 123-141.
- Stover, D. (2018), "Garlin Gilchrist: fighting fake news and the information apocalypse", *Bulletin of the Atomic Scientists*, Routledge, Vol. 74 No. 4, pp. 283-288.
- Strickland, E. (2020), "Facebook takes on deepfakes", *Presented at the IEEE Spectrum*, Vol. 57 No. 1, pp. 40-57.
- Stroud, D., Timperley, V. and Weinel, M. (2020), "Digitalized drones in the steel industry: the social shaping of technology", *Relations Industrielles/Industrial Relations*, Département des relations industrielles de l'Université Laval, Vol. 75 No. 4, pp. 730-750.
- Stupp, C. (2019), "Fraudsters used AI to mimic CEO's voice in unusual cybercrime case", *The Wall Street Journal*, available at: <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402> (accessed 5 June 2022).
- Su, Z., Cheshmehzangi, A., McDonnell, D., Bentley, B.L., da Veiga, C.P. and Xiang, Y.-T. (2022), "Facial recognition law in China", *Journal of Medical Ethics*, medethics-2022-108130.
- Syarif Hartawan, M., Maharani, M.D.D. and Krisnanik Information System, E. (2020), "Structural model of system information for management innovation ruminant-slaughterhouse", *Presented at the 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, pp. 319-323.
- Sylvester, S. (2021), "Don't let them fake you out: how artificially mastered videos are becoming the newest threat in the disinformation war and what social media platforms should do about it", *Federal Communications Law Journal*, George Washington University, National Law Center, Vol. 73 No. 3, pp. 369-392.

- 
- Tandon, A., Dhir, A., Almgren, I., AlNemer, G.N. and Mäntymäki, M. (2021), "Fear of missing out (FoMO) among social media users: a systematic literature review, synthesis and framework for future research", *Internet Research*, Emerald Publishing, Vol. 31 No. 3, pp. 782-821.
- Tang, H. and Cao, H. (2020), "A review of research on detection of fake commodity reviews", *Journal of Physics: Conference Series*, IOP Publishing, Vol. 1651 No. 1, p. 012055.
- Tashman, A. (2021), "Malicious deepfakes' - how California's A.B. 730 tries (and fails) to address the internet's burgeoning political crisis", *Loyola of Los Angeles Law Review*, Vol. 54 No. 4, p. 1391.
- Taylor, B.C. (2021), "Defending the state from digital Deceit: the reflexive securitization of deepfake", *Critical Studies in Media Communication*, Routledge, Vol. 38 No. 1, pp. 1-17.
- Templeton-Knight, K. (2021), "Deepfakes, disinformation and detection: how can journalists know what is real?", *Journalism.Co.Uk*, available at: <https://www.journalism.co.uk/news/deep-fakes-dis-information-and-detection-how-can-journalists-know-what-s-real-/s2/a831519/> (accessed 21 June 2022).
- Thaipisutikul, T., Shih, T.K., Enkhbat, A., Aditya, W., Shih, H.-C. and Mongkolwat, P. (2022), "Beyond fear go viral: a machine learning study on infodemic detection during covid-19 pandemic", *Presented at the 2022 14th International Conference on Knowledge and Smart Technology (KST)*, pp. 1-6.
- Thaw, N.N., July, T., Wai, A.N., Goh, D.H.-L. and Chua, A.Y.K. (2020), "Is it real? A study on detecting deepfake videos", *Proceedings of the Association for Information Science and Technology*, Vol. 57 No. 1, p. e366.
- Ticau, I. and Hadad, S. (2022), "Technological Determinism vs. Social Shaping of Technology. The influence of activity trackers on user's attitudes", *Management Dynamics in the Knowledge Economy*, Vol. 9, pp. 147-163.
- Tiwana, A., Konsynski, B. and Bush, A.A. (2010), "Research commentary—platform evolution: coevolution of platform architecture, governance, and environmental dynamics", *Information Systems Research*, INFORMS, Vol. 21 No. 4, pp. 675-687.
- Tolosana, R., Romero-Tapiador, S., Vera-Rodriguez, R., Gonzalez-Sosa, E. and Fierrez, J. (2022), "DeepFakes detection across generations: analysis of facial regions, fusion, and performance evaluation", *Engineering Applications of Artificial Intelligence*, Vol. 110, p. 104673.
- Tranfield, D., Denyer, D. and Smart, P. (2003), "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", *British Journal of Management*, Vol. 14 No. 3, pp. 207-222.
- Tuomi, A. (2021), "Deepfake consumer reviews in tourism: preliminary findings", *Annals of Tourism Research Empirical Insights*, Vol. 2 No. 2, p. 100027.
- Uçan, A.S., Buçak, F.M., Tutuk, M.A.H., Aydin, H.İ., Semiz, E. and Bahtiyar, Ş. (2021), "Deepfake and security of video conferences", *Presented at the 2021 6th International Conference on Computer Science and Engineering (UBMK)*, pp. 36-41.
- Ullrich, Q.J. (2021), "Is this video real? The principal mischief of deepfakes and how the lanham act can address it - ProQuest", *Columbia Journal of Law and Social Problems*, Vol. 55 No. 1, pp. 1-56, available at: <https://www.proquest.com/openview/68d52bc25ea8d652e7e1ac5be7ed921c/1?pq-origsite=gscholar&cbl=38183> (accessed 18 June 2022).
- Ürmösné Simon, G. and Nyitrai, E. (2021), "The phenomena of epidemic crime, deepfakes, fake news, and the role of forensic linguistics", *Információs Társadalom*, Vol. 21 No. 4, p. 86.
- Vaccari, C. and Chadwick, A. (2020), "Deepfakes and disinformation: exploring the impact of synthetic political video on deception, uncertainty, and trust in news", *Social Media + Society*, SAGE Publications, Vol. 6 No. 1, p. 2056305120903408.
- van der Nagel, E. (2020), "Verifying images: deepfakes, control, and consent", *Porn Studies*, Vol. 7, pp. 1-6.
- Vasist, P.N. and Krishnan, S. (2022a), "Deepfakes: an integrative review of the literature and an agenda for future research", *Communications of the Association for Information Systems*, Vol. 51 No. 1, available at: <https://aisel.aisnet.org/cais/vol51/iss1/14>



- 
- Vasist, P.N. and Krishnan, S. (2022b), "Demystifying fake news in the hospitality industry: a systematic literature review, framework, and an agenda for future research", *International Journal of Hospitality Management*, Vol. 106, p. 103277.
- Velasco, C. (2022), "Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments", *ERA Forum*, Vol. 23 No. 1, pp. 109-126.
- Verdoliva, L. (2020), "Media forensics and DeepFakes: an overview", ArXiv:2001.06564 [Cs], available at: <http://arxiv.org/abs/2001.06564> (accessed 30 December 2021).
- Viglia, G., Kumar, S., Pandey, N. and Joshi, Y. (2022), "Forty years of the service industries journal: a bibliometric review", *The Service Industries Journal*, Routledge, Vol. 42 Nos 1-2, pp. 1-20.
- Vizoso, Á., Vaz-Álvarez, M. and López-García, X. (2021), "Fighting deepfakes: media and internet giants' converging and diverging strategies against hi-tech misinformation", *Media and Communication*, Vol. 9 No. 1, pp. 291-300.
- Wæver, O. and Buzan, B. (2020), "Racism and responsibility – the critical limits of deepfake methodology in security studies: a reply to Howell and Richter-Montpetit", *Security Dialogue*, SAGE Publications, Vol. 51 No. 4, pp. 386-394.
- Wagner, T.L. and Blewer, A. (2019), "'The word real is No longer real': deepfakes, gender, and the challenges of AI-altered video", *Open Information Science*, De Gruyter Open Access, Vol. 3 No. 1, pp. 32-46.
- Wahl-Jorgensen, K. and Carlson, M. (2021), "Conjecturing fearful futures: journalistic discourses on deepfakes", *Journalism Practice*, Routledge, Vol. 15 No. 6, pp. 803-820.
- Wallach, O. (2021), "How to spot fake news", *Geography Bulletin*, Geography Teachers' Association of New South Wales, Vol. 53 No. 1, pp. 46-49.
- Walsh, D. and Downe, S. (2005), "Meta-synthesis method for qualitative research: a literature review", *Journal of Advanced Nursing*, Vol. 50 No. 2, pp. 204-211.
- Wang, C. (2019), "Deepfakes, revenge porn, and the impact on women", *Forbes*, available at: <https://www.forbes.com/sites/chenxiwang/2019/11/01/deepfakes-revenge-porn-and-the-impact-on-women/> (accessed 21 June 2022).
- Wang, S. and Kim, S. (2022), "Users' emotional and behavioral responses to deepfake videos of K-pop idols", *Computers in Human Behavior*, Vol. 134, p. 107305.
- Wang, L., Zhou, L., Yang, W. and Yu, R. (2022), "Deepfakes: a new threat to image fabrication in scientific publications?", *Patterns*, Vol. 3 No. 5, p. 100509.
- Weerawardana, M. and Fernando, T. (2021), "Deepfakes detection methods: a literature survey", *Presented at the 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, pp. 76-81.
- Westerlund, M. (2019), "The emergence of deepfake technology: a review", *Technology Innovation Management Review*, Talent First Network, Vol. 9 No. 11, pp. 40-53.
- Whitaker, B. (2021), "Synthetic Media: how deepfakes could soon change our world", *CBS News*, available at: <https://www.cbsnews.com/news/deepfake-artificial-intelligence-60-minutes-2021-10-10/> (accessed 7 June 2022).
- Whittaker, L., Kietzmann, T.C., Kietzmann, J. and Dabirian, A. (2020), "All around me are synthetic faces': the mad world of AI-generated media", *Presented at the IT Professional*, Vol. 22 No. 5, pp. 90-99.
- Whittaker, L., Letheren, K. and Mulcahy, R. (2021), "The rise of deepfakes: a conceptual framework and research agenda for marketing", *Australasian Marketing Journal*, SAGE Publications, Vol. 29 No. 3, pp. 204-214.
- Widder, D.G., Nafus, D., Dabbish, L. and Herbsleb, J. (2022), "Limits and possibilities for 'ethical AI' in open source: a study of deepfakes", *Presented at the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*, ACM, New York, NY, USA, Seoul, Republic of Korea, p. 12.
-

- 
- Wiederhold, B.K. (2021), "Can deepfakes improve therapy?", *Cyberpsychology, Behavior, and Social Networking*, Mary Ann Liebert, Vol. 24 No. 3, pp. 147-148.
- Wiggers, K. (2022), "Google bans deepfake-generating AI from colab | TechCrunch", *TechCrunch*, available at: <https://techcrunch.com/2022/06/01/2328459/> (accessed 5 June 2022).
- Wilkerson, L. (2021), "Still waters run deep(fakes): the rising concerns of 'deepfake' technology and its influence on democracy and the first amendment", *Missouri Law Review*, Vol. 86 No. 1, pp. 407-432, available at: <https://scholarship.law.missouri.edu/mlr/vol86/iss1/12>
- 
- Williams, R. and Edge, D. (1996), "The social shaping of technology", *Research Policy*, Vol. 25 No. 6, pp. 865-899.
- Wilson, M. and Howcroft, D. (2005), "Power, politics and persuasion in IS evaluation: a focus on 'relevant social groups'", *The Journal of Strategic Information Systems*, Vol. 14 No. 1, pp. 17-43.
- Winner, L. (1980), "Do artifacts have politics?", *Daedalus*, The MIT Press, Vol. 109 No. 1, pp. 121-136.
- Winter, R. and Salter, A. (2020), "DeepFakes: uncovering hardcore open source on GitHub", *Porn Studies*, Routledge, Vol. 7 No. 4, pp. 382-397.
- Wood, J. and Sanders, N. (2020), "Dealing with 'deepfakes': how synthetic media will distort reality, corrupt data, and impact forecasts", *Foresight: The International Journal of Applied Forecasting*, International Institute of Forecasters, Vol. 59, pp. 32-37.
- Wu, Y., Ngai, E.W.T., Wu, P. and Wu, C. (2020), "Fake online reviews: literature review, synthesis, and directions for future research", *Decision Support Systems*, Vol. 132, p. 113280.
- Wu, F., Ma, Y. and Zhang, Z. (2021), "I found a more attractive deepfaked self: the self-enhancement effect in deepfake video exposure", *Cyberpsychology, Behavior, and Social Networking*, Mary Ann Liebert, Vol. 24 No. 3, pp. 173-181.
- Wu, Y., Ngai, E.W.T., Wu, P. and Wu, C. (2022), "Fake news on the internet: a literature review, synthesis and directions for future research", *Internet Research*, Vol. 32 No. 5, pp. 1662-1699, doi: [10.1108/INTR-05-2021-0294](https://doi.org/10.1108/INTR-05-2021-0294).
- Xie, J., He, Z., Burnett, G. and Cheng, Y. (2021), "How do mothers exchange parenting-related information in online communities? A meta-synthesis", *Computers in Human Behavior*, Vol. 115, p. 106631.
- Yadlin-Segal, A. and Oppenheim, Y. (2021), "Whose dystopia is it anyway? Deepfakes and social media regulation", *Convergence*, SAGE Publications, Vol. 27 No. 1, pp. 36-51.
- Yamaoka-Enkerlin, A. (2019), "Disrupting disinformation: deepfakes and the law", *New York University Journal of Legislation and Public Policy*, Vol. 22, p. 725.
- Yang, X., Li, Y. and Lyu, S. (2018), "Exposing deep fakes using inconsistent head poses", ArXiv: 1811.00661 [Cs], available at: <http://arxiv.org/abs/1811.00661> (accessed 5 November 2021).
- Yang, H.-C., Rahmanti, A.R., Huang, C.-W. and Li, Y.-C.J. (2022), "How can research on artificial empathy Be enhanced by applying deepfakes?", *Journal of Medical Internet Research*, Vol. 24 No. 3, p. e29506.
- Yao, Y., Liu, S., Wang, H., Shen, Z. and Ni, X. (2022), "An improved self-training model with fine-tuning teacher/student exchange strategy to detect computer-generated images", *Security and Communication Networks*, Hindawi, Vol. 2022, p. e3493336.
- Yu, P., Xia, Z., Fei, J. and Lu, Y. (2021a), "A survey on deepfake video detection", *IET Biometrics*, Vol. 10 No. 6, pp. 607-624.
- Yu, X., Wojcieszak, M., Lee, S., Casas, A., Azrout, R. and Gackowski, T. (2021b), "The (null) effects of happiness on affective polarization, conspiracy endorsement, and deep fake recognition: evidence from five survey experiments in three countries", *Political Behavior*, Vol. 43 No. 3, pp. 1265-1287.
- Zachary, G.P. (2020), "Digital manipulation and the future of electoral democracy in the US", *Presented at the IEEE Transactions on Technology and Society*, Vol. 1 No. 2, pp. 104-112.

---

Zhang, T. (2022), "Deepfake generation and detection, a survey", *Multimedia Tools and Applications*, Vol. 81 No. 5, pp. 6259-6276.

Zhao, B., Zhang, S., Xu, C., Sun, Y. and Deng, C. (2021), "Deep fake geography? When geospatial data encounter Artificial Intelligence", *Cartography and Geographic Information Science*, Taylor & Francis, Vol. 48 No. 4, pp. 338-352.

Engaging with  
deepfakes

**Corresponding author**

Satish Krishnan can be contacted at: [satishk@iimk.ac.in](mailto:satishk@iimk.ac.in)

---

**Table A1.**  
Evaluation of abstract  
and full text of 211  
studies<sup>a</sup> based on  
search protocol

Sr. No	Title	Assessment based on abstract			Full text based assessment
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	
1	Russell <i>et al.</i> (2012)	No	No	Excluded	Full text based assessment Any exclusion criteria (EC) satisfied?  Included as part of final sample?
2	Huang and Yu (2016)	No	No	Excluded	
3	Chesney and Citron (2018)	No	Yes	Excluded	
4	Stover (2018)	No	Yes	Excluded	
5	Bazarkina and Pashentsev (2019)	No	No	Excluded	
6	Chesney and Citron (2019)	Could not be determined	Yes	Included	
7	Chudinov <i>et al.</i> (2019)	Yes	Yes	Included	
8	Citron (2019)	No	Yes	Excluded	
9	Conte (2019)	Yes	Yes	Included	
10	Delfino (2019)	No	Yes	Excluded	
11	Dixon (2019)	Could not be determined	Yes	Included	
12	Gregory (2019)	No	No	Excluded	
13	Greengard (2019)	No	Yes	Excluded	
14	Groh <i>et al.</i> (2019)	No	Yes	Excluded	
15	Ice (2019)	No	Yes	Excluded	
16	Joseph (2019)	Could not be determined	Yes	Included	
17	Kim <i>et al.</i> (2019)	No	Yes	Excluded	
18	Korshunov and Marcel (2019)	No	Yes	Excluded	
19	Lin <i>et al.</i> (2019)	No	No	Excluded	
20	Popova (2019)	Yes	Yes	Included	
21	Solo (2019)	No	No	Excluded	
22	Stadler (2019)	No	No	Excluded	
23	Wagner and Blewer (2019)	No	Yes	Excluded	
24	Yamaoka-Enkerlin (2019)	No	Yes	Excluded	

(continued)

Sr. No	Title	Assessment based on abstract			Full text based assessment		
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Any exclusion criteria (EC) satisfied?	Included as part of final sample?
25	<a href="#">Paterson and Hanley (2020)</a>	No	Yes	Excluded			
26	<a href="#">Alrasheed et al. (2020)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-b	
27	<a href="#">Anuar and Ibrahim (2020)</a>	Could not be determined	Could not be determined	Included	IC-a, b, c, d not satisfied		
28	<a href="#">Ascott (2020)</a>	No	Yes	Excluded			
29	<a href="#">Bazarkina et al. (2020)</a>	Yes	Yes	Included	IC-b not satisfied		
30	<a href="#">Brown and Fleming (2020)</a>	Yes	Yes	Included	IC-b not satisfied		
31	<a href="#">Chowdhury and Lubna (2020)</a>	No	Yes	Excluded			
32	<a href="#">de Araujo e Silva et al. (2020)</a>	Yes	No	Excluded			
33	<a href="#">de Vries (2020)</a>	No	Yes	Excluded			
34	<a href="#">Dowdeswell and Goltz (2020)</a>	No	No	Excluded			
35	<a href="#">Fallis (2020)</a>	No	Yes	Excluded			
36	<a href="#">Farrish (2020)</a>	No	Yes	Excluded			
37	<a href="#">Garcia Lozano et al. (2020)</a>	No	No	Excluded			
38	<a href="#">Gibson (2020)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied		
39	<a href="#">Pechenik Gieseke (2020)</a>	No	Yes	Excluded			
40	<a href="#">Gosse and Burkell (2020)</a>	Yes	Yes	Included	IC-b not satisfied		
41	<a href="#">Guamera et al. (2020)</a>	No	Yes	Excluded			
42	<a href="#">Jafar et al. (2020)</a>	No	Yes	Excluded			
43	<a href="#">Karnouskos (2020)</a>	No	Yes	Excluded			
44	<a href="#">Kawa and Syga (2020)</a>	No	Yes	Excluded			
45	<a href="#">Kietzmann et al. (2020)</a>	No	Yes	Excluded			

(continued)

Table A1.

Table A1.

Sr. No	Title	Assessment based on abstract			Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Full text based assessment	
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?			Any exclusion criteria (EC) satisfied?	Included as part of final sample?
46	<a href="#">Kikerpill (2020)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied			
47	<a href="#">Kirchengast (2020)</a>	No	Yes	Excluded				
48	<a href="#">Lemley and Casey (2020)</a>	No	No	Excluded				
49	<a href="#">Maddocks (2020)</a>	Yes	Yes	Included	IC-b not satisfied			
50	<a href="#">Mailik et al. (2020)</a>	No	Yes	Excluded				
51	<a href="#">Mbinjama-Gamatham and Olivier (2020)</a>	No	No	Excluded				
52	<a href="#">Meskys et al. (2020)</a>	No	Yes	Excluded				
53	<a href="#">Murray (2020)</a>	No	No	Excluded				
54	<a href="#">Nazar and Bustam (2020)</a>	Could not be determined	Yes	Included	IC-d not satisfied			
55	<a href="#">Newton and Starfill (2020)</a>	Yes	Yes	Included	All satisfied		Yes	
56	<a href="#">O'Connell (2020)</a>	Could not be determined	Could not be determined	Included	IC-a, b, d not satisfied			
57	<a href="#">Öhman (2020)</a>	No	Yes	Excluded				
58	<a href="#">Partadiredja et al. (2020)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a		
59	<a href="#">Paterson and Hanley (2020)</a>	No	Yes	Excluded				
60	<a href="#">Perot and Mostert (2020)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied			
61	<a href="#">Pfefferkorn (2020)</a>	No	Yes	Excluded				
62	<a href="#">Popova (2020)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied			
63	<a href="#">Porter (2020)</a>	No	No	Excluded				
64	<a href="#">Rimi (2020)</a>	No	Yes	Excluded				
65	<a href="#">Ripoll et al. (2020)</a>	Yes	No	Excluded				
66	<a href="#">Sample et al. (2020)</a>	No	No	Excluded				

(continued)

Sr. No	Title	Assessment based on abstract			Full text based assessment		
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Any exclusion criteria (EC) satisfied?	Included as part of final sample?
67	<a href="#">Sepec and Lango (2020)</a>	No	Yes	Excluded			
68	<a href="#">Sethi <i>et al.</i> (2020)</a>	No	Yes	Excluded			
69	<a href="#">Strickland (2020)</a>	No	Yes	Excluded			
70	<a href="#">Syarif Hartawan <i>et al.</i> (2020)</a>	Yes	No	Excluded			
71	<a href="#">Tang and Cao (2020)</a>	No	No	Excluded			
72	<a href="#">Vaccari and Chadwick (2020)</a>	No	Yes	Excluded			
73	<a href="#">van der Nagel (2020)</a>	Yes	Yes	Included	IC-b not satisfied		
74	<a href="#">Verdoliva (2020)</a>	No	Yes	Excluded			
75	<a href="#">Waeber and Buzan (2020)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied	EC-c	
76	<a href="#">Whittaker <i>et al.</i> (2020)</a>	No	Yes	Excluded			Yes
77	<a href="#">Winter and Salter (2020)</a>	Yes	Yes	Included	All satisfied		
78	<a href="#">Wood and Sanders (2020)</a>	No	Yes	Excluded			
79	<a href="#">Zachary (2020)</a>	No	Yes	Excluded			
80	<a href="#">Biometric technology today (2020)</a>	Could not be determined	Could not be determined	Included	IC-a, b, c, d not satisfied		
81	<a href="#">European Journal of Public Health (2020)</a>	Could not be determined	Yes	Included	IC-a, b, c, d not satisfied		
82	<a href="#">Ahmed (2021a)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a	
83	<a href="#">Ahmed (2021b)</a>	No	Yes	Excluded			
84	<a href="#">Ahmed (2021c)</a>	No	Yes	Excluded			
85	<a href="#">Ali <i>et al.</i> (2021)</a>	Yes	Yes	Included	All satisfied		Yes
86	<a href="#">Allison (2021)</a>	Yes	Yes	Included	IC-b not satisfied		
87	<a href="#">Ayers (2021)</a>	Yes	Yes	Included	IC-b not satisfied		
88	<a href="#">Baten and Hoque (2021)</a>	No	Yes	Excluded			
89	<a href="#">Bode (2021)</a>	Yes	Yes	Included	All satisfied		Yes

(continued)

Table A1.

Table A1.

Sr. No	Title	Assessment based on abstract			Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Full text based assessment	
		Qualitative research design?	Theme of deepfakes?	Included			Any exclusion criteria (EC) satisfied?	Included as part of final sample?
90	<a href="#">Bode et al. (2021)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied			
91	<a href="#">Bodi (2021)</a>	No	Yes	Excluded				
92	<a href="#">Boumans (2021)</a>	Could not be determined	Could not be determined	Included	IC-a, b, c, d not satisfied			
93	<a href="#">Brennen et al. (2021)</a>	Yes	No	Excluded				
94	<a href="#">Burgstaller and Macpherson (2021)</a>	No	Yes	Excluded				
95	<a href="#">Campbell et al. (2022)</a>	No	Yes	Excluded				
96	<a href="#">Carlson (2021)</a>	No	Yes	Excluded				
97	<a href="#">Chaudhary et al. (2021)</a>	No	Yes	Excluded				
98	<a href="#">Cheng et al. (2021)</a>	No	No	Excluded				
99	<a href="#">Chrystall (2021)</a>	Yes	No	Excluded				
100	<a href="#">Cochran and Napshin (2021)</a>	No	Yes	Excluded				
101	<a href="#">Dan et al. (2021)</a>	No	Yes	Excluded				
102	<a href="#">Dasilva et al. (2021)</a>	No	Yes	Excluded				
103	<a href="#">de Ruiter (2021)</a>	No	Yes	Excluded				
104	<a href="#">de Seta (2021)</a>	Yes	Yes	Included	All satisfied		Yes	
105	<a href="#">Ding et al. (2021)</a>	No	Yes	Excluded				
106	<a href="#">Dobber et al. (2021)</a>	No	Yes	Excluded				
107	<a href="#">Fagni et al. (2021)</a>	No	Yes	Excluded				
108	<a href="#">Flynn et al. (2021)</a>	Could not be determined	Yes	Included	All satisfied		Yes	
109	<a href="#">García-Orosa (2021)</a>	No	No	Excluded				
110	<a href="#">Ghazi-Tehrani and Pontell (2021)</a>	Yes	Yes	Included	All satisfied		Yes	
111	<a href="#">Godulla et al. (2021)</a>	No	Yes	Excluded				
112	<a href="#">Goltz and Dowdeswell (2021)</a>	No	No	Excluded				

(continued)



Sr. No	Title	Assessment based on abstract			Full text based assessment		
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Any exclusion criteria (EC) satisfied?	Included as part of final sample?
113	<a href="#">Hancock and Bailenson (2021)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied		
114	<a href="#">Harper et al. (2021)</a>	Could not be determined	Yes	Included	IC-a, b, c, d not satisfied		
115	<a href="#">Harris (2021)</a>	No	Yes	Excluded			
116	<a href="#">Hodge (2021)</a>	No	Yes	Excluded			
117	<a href="#">Holliday (2021)</a>	Yes	Yes	Included	IC-b not satisfied		
118	<a href="#">Hwang et al. (2021)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a	
119	<a href="#">Iacobucci et al. (2021)</a>	No	Yes	Excluded			
120	<a href="#">Jaynes (2021)</a>	No	No	Excluded			
121	<a href="#">Johnson and Diakopoulos (2021)</a>	No	Yes	Excluded			
122	<a href="#">García-Ull (2021)</a>	No	Yes	Excluded			
123	<a href="#">Gregory (2022)</a>	Yes	Yes	Included	All satisfied		Yes
124	<a href="#">Hight (2022)</a>	No	Yes	Excluded			
125	<a href="#">Karasavva and Noorbhai (2021)</a>	No	Yes	Excluded			
126	<a href="#">Kelly (2021)</a>	Could not be determined	Could not be determined	Included	IC-a, b, c, d not satisfied	EC-c	
127	<a href="#">Kerner and Risse (2021)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied	EC-d	
128	<a href="#">Khichi and Kumar Yadav (2021)</a>	No	Yes	Excluded			
129	<a href="#">Kietzmann et al. (2021)</a>	No	Yes	Excluded			
130	<a href="#">Köbis et al. (2021)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a	
131	<a href="#">Krishna (2021)</a>	No	Yes	Excluded			
132	<a href="#">Kugler and Pace (2021)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a	

(continued)

Table A1.

Table A1.

Sr. No	Title	Assessment based on abstract			Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Full text based assessment	
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?			Any exclusion criteria (EC) satisfied?	Included as part of final sample?
133	<a href="#">Kwok and Koh (2021)</a>	No	Yes	Excluded				
134	<a href="#">Langa (2021)</a>	No	Yes	Excluded				
135	<a href="#">Langguth et al. (2021)</a>	No	Yes	Excluded				
136	<a href="#">Lee and Shin (2022)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a		
137	<a href="#">Lee et al. (2021)</a>	Yes	Yes	Included				
138	<a href="#">Lees et al. (2021)</a>	Could not be determined	Yes	Included				
139	<a href="#">Fan et al. (2021)</a>	No	Yes	Excluded				
140	<a href="#">McPeak (2021)</a>	No	Yes	Excluded				
141	<a href="#">Mihailova (2021)</a>	Yes	Yes	Included				
142	<a href="#">Mirsky and Lee (2021)</a>	No	Yes	Excluded	IC-b not satisfied			
143	<a href="#">O'Donnell (2021)</a>	No	Yes	Excluded				
144	<a href="#">Pandey et al. (2021)</a>	No	Yes	Excluded				
145	<a href="#">Pavis (2021)</a>	No	Yes	Excluded				
146	<a href="#">Pesetski (2021)</a>	Could not be determined	Yes	Included				
147	<a href="#">Popescu (2021)</a>	No	Yes	Excluded				
148	<a href="#">Rao et al. (2021)</a>	No	No	Excluded				
149	<a href="#">Rasmussen (2021)</a>	Could not be determined	Yes	Included	IC-a, b, c, d not satisfied	EC-c		
150	<a href="#">Ratner (2021)</a>	No	Yes	Excluded				
151	<a href="#">Ray (2021)</a>	No	Yes	Excluded				
152	<a href="#">Reid (2021)</a>	Could not be determined	Yes	Included				
153	<a href="#">Ürmösné Simon and Nyitrai (2021)</a>	No	Yes	Excluded				
154	<a href="#">Sylvester (2021)</a>	No	Yes	Excluded				
155	<a href="#">Tashman (2021)</a>	No	Yes	Excluded				

(continued)

Sr. No	Title	Assessment based on abstract			Full text based assessment		
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Any exclusion criteria (EC) satisfied?	Included as part of final sample?
156	<a href="#">Taylor (2021)</a>	Yes	Yes	Included	IC-a, b, d not satisfied		
157	<a href="#">Tuomi (2021)</a>	Could not be determined	Yes	Included	IC-a, b not satisfied	EC-a	
158	<a href="#">Uçan et al. (2021)</a>	No	Yes	Excluded			
159	<a href="#">Ulrich (2021)</a>	No	Yes	Excluded			
160	<a href="#">Vizoso et al. (2021)</a>	Yes	Yes	Included	IC-b not satisfied		
161	<a href="#">Wahl-Jørgensen and Carlson (2021)</a>	Yes	Yes	Included	IC-b not satisfied		
162	<a href="#">Weerawardana and Fernando (2021)</a>	No	Yes	Excluded			
163	<a href="#">Whittaker et al. (2021)</a>	No	Yes	Excluded			
164	<a href="#">Wiederhold (2021)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied		
165	<a href="#">Wilkerson (2021)</a>	No	Yes	Excluded			
166	<a href="#">Wu et al. (2021)</a>	No	Yes	Excluded			
167	<a href="#">Yadlin-Segal and Oppenheim (2021)</a>	Could not be determined	Yes	Included	IC-b not satisfied		
168	<a href="#">Yu et al. (2021b)</a>	No	Yes	Excluded			
169	<a href="#">Yu et al. (2021a)</a>	No	Yes	Excluded			
170	<a href="#">Zhao et al. (2021)</a>	No	Yes	Excluded			
171	<a href="#">Biometric Technology Today (2021)</a>	No	Yes	Excluded			
172	<a href="#">Wallach (2021)</a>	Could not be determined	Could not be determined	Included	IC-a, b, c, d not satisfied		
173	<a href="#">Ahmed (2022)</a>	No	Yes	Excluded			
174	<a href="#">Andrejevic et al. (2022)</a>	No	No	Excluded			
175	<a href="#">Cross (2022)</a>	No	Yes	Excluded			
176	<a href="#">de Rancourt-Raymond and Smaili (2022)</a>	Yes	Yes	Included	IC-b not satisfied		

(continued)

Table A1.

Table A1.

Sr. No	Title	Assessment based on abstract			Full text based assessment		
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Any exclusion criteria (EC) satisfied?	Included as part of final sample?
177	<a href="#">Belmaa (2022)</a>	Yes	Yes	Included	All satisfied		Yes
178	<a href="#">Fido <i>et al.</i> (2022)</a>	No	Yes	Excluded			
179	<a href="#">Galyashina and Nikishin (2022)</a>	No	Yes	Excluded			
180	<a href="#">Himma-Kadakas and Ojiamets (2022)</a>	Yes	Yes	Included	All satisfied		Yes
181	<a href="#">Hirlekar and Kumar (2022)</a>	No	No	Excluded			
182	<a href="#">Yang <i>et al.</i> (2022)</a>	No	Yes	Excluded			
183	<a href="#">Johnson (2022)</a>	No	No	Excluded			
184	<a href="#">Lopez (2022)</a>	No	Yes	Excluded			
185	<a href="#">Juefei-Xu <i>et al.</i> (2022)</a>	No	Yes	Excluded			
186	<a href="#">Mangaokar and Prakash (2022)</a>	No	Yes	Excluded			
187	<a href="#">McCosker (2022)</a>	Yes	Yes	Included	IC-b not satisfied		
188	<a href="#">Milliere (2022)</a>	No	Yes	Excluded			
189	<a href="#">Mone (2022)</a>	Could not be determined	Yes	Included	IC-a, b, d not satisfied		
190	<a href="#">Moon and Iacobucci (2022)</a>	No	No	Excluded			
191	<a href="#">Moss-Wellington <i>et al.</i> (2022)</a>	No	No	Excluded			
192	<a href="#">Mullen (2022)</a>	No	Yes	Excluded			
193	<a href="#">Murphy and Flynn (2022)</a>	No	Yes	Excluded			
194	<a href="#">Nour and Gelfand (2022)</a>	No	Yes	Excluded			
195	<a href="#">Saif and Tehseen (2022)</a>	No	Yes	Excluded			
196	<a href="#">Saxena and Gayathri (2022)</a>	No	No	Excluded			
197	<a href="#">Shelke and Attrar (2022)</a>	No	No	Excluded			
198	<a href="#">Shin and Lee (2022)</a>	No	Yes	Excluded			

(continued)

Sr. No	Title	Assessment based on abstract			Full-text evaluation (included/Excluded)?	All inclusion criteria (IC) satisfied?	Full text based assessment	
		Qualitative research design?	Theme of deepfakes?	Full-text evaluation (included/Excluded)?			Any exclusion criteria (EC) satisfied?	Included as part of final sample?
199	<i>Bajema et al. (2022)</i>	No	Yes	Excluded				
200	<i>Su et al. (2022)</i>	Could not be determined	No	Excluded				
201	<i>Thaipisutikul et al. (2022)</i>	No	No	Excluded				
202	<i>Velasco (2022)</i>	No	No	Excluded				
203	<i>Wang et al. (2022)</i>	No	Yes	Excluded				
204	<i>Wang and Kim (2022)</i>	No	Yes	Excluded				
205	<i>Yao et al. (2022)</i>	No	Yes	Excluded				
206	<i>Zhang (2022)</i>	No	Yes	Excluded				
207	<i>Fikse (2018)</i>	Additional studies sourced in final step of literature search. Assessed directly for full text				All satisfied	Yes	Yes
208	<i>Thaw et al. (2020)</i>					All satisfied	Yes	Yes
209	<i>Gamage et al. (2022)</i>					All satisfied	Yes	Yes
210	<i>Sabrina (2022)</i>					All satisfied	Yes	Yes
211	<i>Widder et al. (2022)</i>					All satisfied	Yes	Yes

**Note(s):** <sup>a</sup>Includes 206 studies sourced through the search protocol and 5 studies sourced in the final step of literature search through hand searching and cross-referencing mechanisms

**Table A2.**  
A synopsis of studies  
in focus

Sr. No	Study	Description of the study	Details on method(s) and sample(s) utilized for the study	Country of the research project	Key finding(s)
1	<a href="#">Fikse (2018)</a>	Recognized the concerns around misuse of face-swapping technologies and socio-technical consequences of deepfakes	Digital ethnographic methods through exploration and participation in discussion forums, code sharing services, video sharing sites, and Internet archives. Autoethnographic account through interaction with deepfake technology and tools in own deepfakes lab	Norway	Found that deceptive deepfakes, although increasingly difficult to create, reinforce the central tenets of post-truth while the fears and interests surrounding the phenomenon are reinforced by the notion of post-truth
2	<a href="#">Popova (2019)</a>	Analyzed deepfakes from the lens of how the audience engages with celebrity deepfakes and the purposes behind creation of such fakery	Digital ethnographic approach with primary data from two deepfake sites: <a href="#">mrdeepfakes.com</a> and the <a href="#">voat.co</a> "subverse" v/DeepFake	UK	Noted low levels of concerns among deepfake communities for the private individuals behind the deepfake imagery and highlights the attempts of these communities to restrict such content within the communities as opposed to distributing it to a larger audience
3	<a href="#">Newton and Stanfill (2020)</a>	Examined the community surrounding software related to deepfakes to understand the development and distribution through collaborative efforts	Thematic analysis of deepfake-related discussions and comments on two GitHub repositories: <a href="#">faceswap</a> and <a href="#">faceswap-playground</a>	USA	Found that while open-source software helps individuals to develop technological tools, it also evolves into a site of toxic greek masculinity and deepfakes enable this particular enactment while hiding behind the technocratic rationale to deny compassion
4	<a href="#">Thaw et al. (2020)</a>	Aimed to understand features that individuals use to characterize videos as deepfake, and strategies involved in the identification process	Semi-structured interviews with 20 individuals who were either graduate students or working professionals	Singapore	Few participants were able to correctly identify deepfake content, and the study identified a few additional factors, such as the naturalness of speech and emotional expressions, which may be integrated into detection algorithms

(continued)

Sr. No	Study	Description of the study	Details on method(s) and sample(s) utilized for the study	Country of the research project	Key finding(s)
5	Winter and Salter (2020)	Investigated roles of two platforms, GitHub and Reddit in the propagation of deepfakes and their ethical stance including the approach to regulation of content	Case study with analysis of community interactions on a deepfake Reddit forum and GitHub's faceswap code repository	USA	The spread of the DeepFakes project to settings with minimal regulation was found to be intentional and that in turn has led to the unabated development of non-consensual content of an explicit nature Demonstrated the efficacy of awareness-related activities that assisted children in recognizing the dangers of disinformation and gaining knowledge of media that may be convincing but are not always authentic
6	Ali <i>et al.</i> (2021)	Analyzed the effectiveness of specific interventions aimed at digital media literacy among children in the evolving context of AI systems and their societal and ethical implications	Activities with 38 middle and high school students from five states in the US	USA	Demonstrated how content is designed to attract viewership and how the platform features shape the discourse and the role of networked publics in this context
7	Bode (2021)	Evaluated how the audience reacts and engages with deepfaked videos and the role of framing contexts in this regard	Case study of a deepfaked video on a YouTube visual effects channel, Corridor and discussions surrounding this video on the YouTube channel	Australia	Highlighted the role of the Chinese tech industry in shaping deepfakes, the practical focus on ramifications such as ethical imbalances, image rights, and so on, and the co-linkages between deepfakes and other technologies such as biometric systems
8	de Seta (2021)	Explored how deepfakes are evolving in the local context of the market in China and its various local interpretations	Case study on deepfake domestication in China through Zao app, commercialization of deepfakes, and communities of practice	China	Emphasized the pervasive nature of deepfakes and the harms they inflict, as well as the difficulties in regulatory responses, policing, and mitigation
9	Flynn <i>et al.</i> (2021)	Examined the details surrounding victimization and perpetration of deepfakes and digitally altered explicit imagery	Semi-structured interviews with 75 image-based sexual abuse victim-survivors and 43 stakeholders	UK, Scotland, New Zealand, Australia	

(continued)

Table A2.

Sr. No	Study	Description of the study	Details on method(s) and sample(s) utilized for the study	Country of the research project	Key finding(s)
10	<a href="#">Ghazi-Tehrani and Pontell (2021)</a>	Explored the evolving nature of phishing as a cybercrime and the role of deepfakes within this context	Semi-structured interviews with 62 professionals from the information technology security domain, hackers, and academic scholars	USA	Considerable agreement was found on the contribution of technology developments to the rise of phishing attacks and their links to financial fraud and ransomware, while the same advancements also aid in their detection. Highlighted the importance of deepfakes in aggravating the phishing problem and emphasized the necessity for public education to alleviate the crisis
11	<a href="#">Gregory (2022)</a>	Highlights key activities of WITNESS, a human rights and civic journalism network in the wake of tackling new media challenges such as deepfakes	Case study drawing upon author's own experience and those of colleagues at WITNESS	USA	Explicated on authenticity infrastructure as a mechanism to track media provenance, integrity, and novel methods such as verification incentives that would allow distant witnesses to correctly interpret eyewitness material. Emphasizes the hazards to information posed by deepfakes and the necessity to tackle this problem
12	<a href="#">Eelmaa (2022)</a>	Examined the concerns related to deepfake enabled explicit content of children on online forums	Thematic analysis of 13,293 user comments posted to a thread on Reddit related to a sitewide rule change	Estonia	A gendered discourse, considerable ambiguity about computer-generated explicit content depicting children, and discrepancies in permissibility regarding animated and deepfaked videos were highlighted through the analysis

*(continued)*



Sr. No	Study	Description of the study	Details on method(s) and sample(s) utilized for the study	Country of the research project	Key finding(s)
13	<a href="#">Gamage et al. (2022)</a>	Explored how online communities perceive deepfakes and linkages to societal concerns	Thematic analysis of 6,638 posts and 86,425 comments related to conversations on deepfakes on Reddit from 2018 to 2021	Japan	Discussions revealed the pro-deepfake nature of Reddit conversations and the support for the creation and transmission of deepfake artifacts, combined with mounting societal concerns around the phenomenon Revealed that regardless of the platform a journalist specializes in, deepfakes were found to be among the most difficult to discern
14	<a href="#">Himma-Kadakas and Ojaments (2022)</a>	Analyzed fact-checking skills of journalists and their abilities to debunk false information including deepfakes	Semi-structured interviews with 20 journalists from online, radio, and television newsrooms	Estonia	Revealed general consensus around the loss of public trust due to deepfakes, a variety of fears surround the phenomenon while some felt reassured with the emergence of deepfake detection apps. Also discussed procedures employed by individuals to verify videos for authenticity
15	<a href="#">Sabrina (2022)</a>	Examined fake news and deepfakes as phenomena, their societal implications including democratic harms and responses to mitigate such negative ramifications	Semi-structured interviews with 10 university students (5 South Korean and 5 expatriates)	South Korea	Found that notions of unrestricted code usage, beliefs surrounding neutrality and inevitability of technology drove how members framed responsibilities within the community, including use and misuse of deepfakes and motivations for ethical actions
16	<a href="#">Widder et al. (2022)</a>	Examined the practices around transparency and accountability in open-source AI communities	Semi-structured interviews with 11 professionals	USA, Europe	