Patric Raemy* and Manuel Puppis

The impacts of deepfakes on journalism: Risks, opportunities, and changes in news organizations

https://doi.org/10.1515/commun-2025-0211

Abstract: Deep learning technology allows for the creation of synthetic audio and video content that appears authentic, offering both risks and opportunities for journalism. Such synthetic content may enhance personalization, visualization, and immersion, but challenge the detection of disinformation. To date, there is limited empirical knowledge about the impacts of deepfakes on journalism. Thus, this study's goal is to explore how journalists assess potential risks and opportunities and how news organizations respond to deepfakes or synthetic content. Interviews with representatives of major Swiss news organizations show that journalists predominantly consider deepfakes as a form of disinformation, and that they have trouble imagining the technology's potential opportunities for journalism. Moreover, although specialized technological knowledge in news organizations is crucial, mitigating the risks posed by deepfakes and taking advantage of the opportunities synthetic content also require a reorientation toward professional journalistic norms and raising awareness among journalists and audiences alike. These results highlight that deepfakes make it necessary for journalists to renegotiate the role of journalism, professional norms, and the boundaries of AI manipulation.

Keywords: deepfakes, journalism, disinformation, news organizations, artificial intelligence, synthetic content

1 Introduction

Manipulations and deceptions have existed long before it became possible to create synthetic content using artificial intelligence (AI). Yet today's AI-driven, complex manipulations of images, words, and voices stand out for how easily synthetic

***Corresponding author: Patric Raemy**, Department of Communication and Media Research, University of Fribourg, Switzerland, E-mail: patric.raemy@unifr.ch. https://orcid.org/0000-0002-0771-4896
Manuel Puppis, Department of Communication and Media Research, University of Fribourg, Switzerland, E-mail: manuel.puppis@unifr.ch. https://orcid.org/0000-0002-4139-251X

video and audio content can be produced and for how authentic it can seem. These AI manipulations, known as deepfakes, refer to videos that use deep learning to hybridize or generate human bodies and faces (Paris and Donovan, 2019). The rise of social media has led to an increase in audio and visual data shared by users, providing AI with vast raw materials to recognize, classify, and recreate images–and to create deepfakes (Meikle, 2023).

As deepfakes push disinformation from mostly written text and static images to synthetic video and/or audio content, they threaten to exacerbate existing difficulties to detect manipulated information (Hameleers, 2023; Schapals and Bruns, 2022). Deepfakes are therefore posing a challenge to public trust, which is the foundation of communication and a central pillar of journalism. To explore the implications of deepfakes for journalism, this study examines journalists' assessment of risks and opportunities, as well as measures taken by news organizations.

Deepfakes have several implications for journalism. Deepfakes depict things that appear real and authentic but in fact never happened in reality. Today, based on short prompts, generative AI tools are able to produce texts, images, audios and videos which could challenge the accuracy, authenticity, and credibility of journalism (Cools and Diakopoulos, 2024). For news organizations committed to objective reporting, it is therefore imperative to be able to detect deepfakes and to not contribute to the spread of false or manipulated information (Godulla et al., 2021). In addition to the challenge of verifying information and detecting deepfakes, news organizations and journalists must also be prepared to defend themselves against attacks and discrediting caused by deepfakes (e.g., Compton, 2021; Godulla et al., 2021; Posetti, 2018a). While the term "deepfake" alludes to undue manipulation, the technology behind it may also offer opportunities for journalism. AI-generated synthetic content offers potential to increase the personalization, visualization, and immersion of news (Wahl-Jorgensen and Carlson, 2021). To better understand the implications of deepfakes on journalism, it is thus necessary, as with previous technological innovations, to investigate the negotiations within news organizations between transformative and stable aspects of journalism (Zelizer, 2019). Based on interviews with representatives of major Swiss news organizations, this study investigates how journalists assess the risks and opportunities of deepfakes or synthetic content, and whether and how news organizations take measures to detect deepfakes or mitigate associated risks.

The paper begins by discussing journalism's role in society in the context of disinformation and deepfakes before presenting the current state of research on deepfakes and synthetic content in journalism. Following a discussion of research questions and methods, the findings will be presented. The study shows that journalists predominantly consider deepfakes as a form of disinformation and that they have trouble imagining the technology's potential opportunities for journalism. Moreo-

ver, although specialized technological knowledge in news organizations is crucial, mitigating the risks posed by deepfakes and taking advantage of the opportunities also require a reorientation toward professional journalistic norms and raising awareness among journalists and audiences alike. The final section then discusses these findings, emphasizing the need to renegotiate the role of journalism in the context of synthetic content.

2 Deepfakes and journalism

Journalism plays a crucial role in democratic systems, as it is expected to contribute to citizens' opinion formation by informing them about relevant issues and positions, providing a forum for political debate, and performing a watchdog role over powerful individuals and institutions (Christians et al., 2009). Its foundational principles–public service, objectivity, independence, immediacy, and ethics–are vital to democracy (Deuze, 2005). They are also vital to journalism itself. Sjøvaag (2010) conceptualizes a "social contract" between journalism and the public. It is an implicit pact that grants journalism autonomy and authority–such as freedom of expression and agenda-setting power–in return for a commitment to uphold ethical standards. Society consents to the social contract so long as journalists honour their obligations to inform citizens truthfully and responsibly, thereby maintaining the public's trust and confidence in the media.

However, trusting news media means that the audience relies on the information they provide, even though the audience cannot oversee every editorial decision. This trust may suffer if the media fall short of expectations, providing the audience with incorrect information and false beliefs (van Dalen, 2020). Yet critically examining and trying to verify information based on professional norms has always been difficult for news media, and it has become even more challenging in the context of disinformation campaigns and information manipulation in the digitized public sphere. Disinformation refers to the intentional creation and dissemination of factually false or misleading information. In contrast, misinformation refers to information that is factually incorrect but is believed to be true by the sender and thus does not have a deceptive intent (Hameleers, 2023; High-Level Expert Group on Fake News and Disinformation, 2018; Kalsnes et al., 2021). With politically motivated disinformation, malevolent actors may aim at delegitimizing democracy by, for instance, sowing discord, amplifying polarized divides, and fueling cynicism (Hameleers, 2023).

Given the rise of generative AI, disinformation has evolved from mostly written articles and static images to synthetic video and/or audio content that is hard to

distinguish from unaltered media and from mis- and disinformation (Schapals and Bruns, 2022). Deepfakes allow for the creation of deceptive synthetic video and audio content using AI to appear authentic, often depicting real persons, objects, or events in artificial scenarios. As a result, deepfakes can be used to make authentic actors do and express things they never did in real life (Bateman, 2020; Hameleers, 2023; Vaccari and Chadwick, 2020; Weikmann and Lecheler, 2023). Prominent examples in Switzerland include fraudulent advertisements that use Swiss news anchors to lure people (Flach, 2024) as well as a political advertisement featuring a politician from an opposing party to promote statements that contradict those individual's actual political stance (Schweizer Radio und Fernsehen, 2024). In contrast to so-called cheapfakes or shallowfakes that can be created with fewer resources, well-made deepfakes are much harder to detect (Brady and Meyer-Resende, 2020; Paris and Donovan, 2019).

Hence, in contrast to other forms of communication, journalism is strongly tied to a broader democratic purpose of verification, validating facts, and truth-telling (Porlezza, 2019). Yet, by producing or inadvertently disseminating deepfakes, journalism risks challenging its foundational norms, threatening objectivity and accuracy, potentially eroding public trust, and blurring the line between fact and fiction (Citron and Chesney, 2018; Vaccari and Chadwick, 2020), contesting journalism's role of truthfulness and trustfulness (Canella, 2023; van Dalen, 2020). To counter these risks, journalism needs to face AI-driven changes by making innovation vital for its survival (Meier et al., 2024) and simultaneously upholding its foundational norms of accuracy and truth (Porlezza, 2019).

3 Literature review: Opportunities and risks offered by deepfakes and synthetic content for journalism opportunities

While studies mostly discuss deepfakes as a form of disinformation and the problems this may create, it seems important for news organizations to also examine potential opportunities for journalistic work. Duffy and Ang (2019) argued that the informational uncertainty created by unchecked sources on the internet could result in audiences relying more on professional media outlets that they deem more credible. Similarly, Wahl-Jorgensen and Carlson (2021) argued that the successful handling of deepfakes allows journalism to differentiate itself from unfiltered communication on social networks. In this vein, media companies and professionals can try to promote their own legitimacy as trustful sources by raising

public awareness about deepfakes (Diakopoulos and Johnson, 2021; Godulla et al., 2021; Wahl-Jorgensen and Carlson, 2021), by publishing and discussing examples of deepfake cases with their users (Vizoso et al., 2021), and by publishing manipulated alongside authentic videos for comparative purposes (Westerlund, 2019).

Beyond these possibilities to position journalism as a trustworthy source, AI as the main technology behind deepfakes also allows for useful applications in journalism. Recent studies indicate that news organizations may use various types of AI such as machine learning, computer vision, and planning, scheduling, and optimization (de-Lima-Santos and Ceron, 2022) and that generative AI has applications in news gathering, production, verification, and distribution (Cools and Diakopoulos, 2024). Focusing on deepfakes specifically, AI such as computer vision can–on the one hand–be used for detecting manipulated content (e.g., Cools and Diakopoulos, 2024; Jarrahi and Safari, 2022; de-Lima-Santos and Ceron 2022; Peng and Xintong, 2022; Rohera et al., 2022), even though detection might always be a step behind creation. On the other hand, news organizations might make use of generative AI to create synthetic video or audio content themselves. The technology offers potential for entertainment and creative purposes as well as for directly addressing users, thereby improving the personalization, visualization, and immersion of news (Collins and Ebrahimi, 2021; Cools and Diakopoulos, 2024; Gamage et al., 2022; Habgood-Coote, 2023; Hellyer, 2022; Wahl-Jorgensen and Carlson, 2021). AI also simplifies the production of content in multiple languages (Cools and Diakopoulos, 2024; Diakopoulos and Johnson, 2021) or the translation of content into sign language with avatars (SwissTXT, n.d.). However, despite its potential, the automation and alteration of journalistic content raises important ethical and copyright issues (Montal and Reich, 2017). In sum, while the literature on the potential benefits of deepfakes for journalism remains scarce and often focuses on AI rather than deepfakes specifically, this paper focuses on a specific AI application–the creation of synthetic video and audio content–that presents opportunities but also entails considerable risks and ethical concerns.

Risks to the role of journalism in a democracy

Unsurprisingly, most studies in journalism research focus on risks posed by deepfakes, including the difficulties in detecting them and the implications for trust in journalism. The proliferation of deepfakes, in conjunction with the platformization of the public sphere, complicates information verification, leading some scholars to call for a reassessment of journalism's role in democracy (Waisbord, 2018).

Deepfakes can be spread quickly on social networks and video-sharing platforms, making it crucial for news organizations to verify videos before the public

perceives them as real (Sohrawardi et al., 2020). Verification is also important to prevent the media from contributing to the spread of deepfakes themselves. Once deepfakes are published, it becomes difficult for news organizations to correct the erroneous reporting. An especially dangerous scenario is if a deepfake disseminated by a news outlet is further spread through other channels and is cited as a credible source, because deepfakes can lead to a loss of control by news organizations (Godulla et al., 2021). Moreover, spreading deepfakes can lead to reputation damage for news organizations themselves (Cool and Diakopoulos, 2024).

Because of their growing sophistication, deepfakes complicate the differentiation between true and false information and weaken the sense of social reality (Yadlin-Segal and Oppenheim, 2021), create uncertainty among media users about what information can be trusted (Vizoso et al., 2021), and hinder the assessment of news quality (Godulla et al., 2021; Vizoso et al., 2021). In addition, deepfakes may influence public opinion (Godulla et al., 2021) and may be used for propaganda (Wahl-Jorgensen and Carlson, 2021).

There is a risk that (para-)journalistic offerings, characterized by one-sided or misleading reporting, might use deepfakes to undermine public trust. For instance, Humprecht et al. (2023) showed that populist actors and their (journalistic) communication channels often spread disinformation to discredit their opponents. These developments pose risks to the role of journalism in a democracy.

Adapting editorial workflows to detect deepfakes

To detect deepfakes and avoid their unintentional dissemination, news organizations may need to adapt their editorial processes and routines. Scholars suggest recognizing deepfakes more effectively requires cooperation with research institutions, platform operators, and other media organizations (Vizoso et al., 2021). Shared databases (Sylvester, 2021) and blogs (Vizoso et al., 2021) can be helpful as well. Moreover, scholars recommend the use of new technology to detect deepfakes, such as machine learning (Gregory, 2019; Leibowicz et al., 2020; Wahl-Jorgensen and Carlson, 2021), algorithms (Collins and Ebrahimi, 2021), computational methods (Vizoso et al., 2021), forensic techniques, digital watermarks (Westerlund, 2019), as well as blockchain-based deepfake scanners (Sylvester, 2021). However, there is a risk that tools for detecting deepfakes may not function accurately and fail to detect all forgeries (Sohrawardi et al., 2020). In addition, such tools involve the risk of false positives (i. e., falsely identifying non-manipulated content as a deepfake; Lewis et al., 2023), which has implications for freedom of expression (Diakopoulos and Johnson, 2021; Sohrawardi et al., 2020). Scholars have also discussed the importance of raising awareness among media professionals and promoting training in news

organizations for effectively dealing with deepfakes (Marconi and Daldrup, 2018). This can be done by working on training data sets to experimentally produce deepfakes and understand their functioning (Gregory, 2019; Lyu, 2020; Malik et al., 2022). Godulla and colleagues (2021) compiled various detailed steps for detecting deepfakes, such as focusing on key points in the video or analyzing facial expressions and gestures.

The use of deepfakes for slandering journalists

Deepfakes not only bring risks to public opinion formation and media trust, but they can also directly affect individual media professionals. In a global context, journalists are often under attack for doing their job–scrutinizing those in power and uncovering injustice (Sammut et al., 2023). Although the situation in Switzerland is comparatively safe for journalists, federal measures have been taken to increase the protection of journalists from threats and violence (Federal Office of Communications, 2024). With deepfakes, another tool is available for such attacks: Journalists risk being discredited through deepfakes to prevent critical reporting (Posetti, 2018a, 2018b). For example, it is possible to integrate images of journalists into pornographic deepfakes (Ayyub, 2018; Razek, 2018), pair them with false interview partners, use them as false interviewers, or manipulate their actions in interviews (Compton, 2021; Marconi and Daldrup, 2018). Finally, deepfakes can also misuse trustworthy personalities like journalists working for reliable news outlets for advertising, spam, and fraudulent purposes (Flach, 2024).

4 Exploring the impacts of deepfakes in Swiss news organizations

Despite this body of literature on risks and opportunities of deepfakes or synthetic content and measures taken to detect them, little is known about the impacts of deepfakes on journalistic work. So far, only few studies offer empirical insights into how news organizations assess and deal with deepfakes. Likewise, there are no studies on how news organizations deal with situations in which journalists themselves are affected by deepfakes. Hence, focusing on the case of Switzerland, this exploratory study looks into how representatives of Swiss news organizations evaluate the impacts of deepfakes on journalism. Specifically, two research questions were investigated:

RQ1: How do representatives of Swiss news organizations assess the risks and opportunities of deepfakes for journalism?
RQ2: What measures have Swiss news organizations taken to detect deepfakes and to mitigate their risks?

To understand how news organizations handle deepfakes, expert interviews–a special form of semi-standardized interviews–were conducted. Experts are defined as individuals who belong to a functional elite within an organizational context, meaning they have influence due to their knowledge and position (Bogner and Menz, 2009). In September 2022, the ten largest news outlets owned by the six largest Swiss media organizations (CH Media, Fondation Aventinus, NZZ Mediengruppe, Ringier, the public service media organization SRG SSR, and TX Group) were contacted directly and asked to provide representatives who would be able to discuss how their editorial teams evaluate the opportunities and risks posed by deepfakes. While one news outlet decided not to participate, ten individuals from the other nine news organizations (*Blick*, Heidi.news, *20 Minuten*, *20 minutes*, *Neue Zürcher Zeitung*, RTS, SRF, Tamedia Zentralredaktion Deutschschweiz, *Watson*) were interviewed (see Appendix). This sample covers the most important news organizations reaching a majority of the population. Given the small country and media market size of Switzerland, the sample of 10 interviews is small but aligns with the exploratory nature of this study, which focuses on a specific group of experts who represent their news organizations and provide in-depth insights into how deepfakes are handled within their organizational contexts.

The interviews focused on how the experts perceive the opportunities and risks of deepfakes in journalism and whether and how measures to deal with deepfakes were implemented in their news organizations. An interview guide was developed based on the current state of research. To address the research questions, the interview guide consisted of thematic areas and specific questions to evaluate the risks, opportunities, and changes in news organizations in the context of deepfakes. The first theme concerned opportunities and risks of deepfakes. Participants were first openly asked to assess the opportunities and risks associated with deepfakes. Since the term deepfake often carries a negative connotation, potential opportunities and risks identified during literature review were summarized in a list. The journalists were then asked to discuss this list, ensuring that not only risks but also opportunities would be explicitly discussed even if not initially mentioned by the interviewees themselves. The second theme concerned procedures to detect deepfakes implemented by news organizations. Questions focused on specific measures and training among journalists in different stages of the media production process and the prevention of the spread of deepfakes. Here, the focus was on procedures, routines, and measures to select and verify information. Questions also addressed

possible options for intervention and redress in cases where deepfakes have been inadvertently disseminated. The third theme concerned measures taken when journalists themselves were targeted by deepfakes. The fourth and final theme addressed how to raise awareness of deepfakes among media professionals and their audiences.

The interviews were conducted face-to-face (except for two online interviews) between October 2022 and January 2023 and lasted about 60 minutes each. All interviews were recorded with the interviewees' permission and subsequently transcribed by student assistants. An inductive qualitative content analysis was used to evaluate the interviews, developing categories through a combination of inductive and deductive coding (Puppis, 2019; Mayring, 2022). Using the software MAXQDA, all transcripts were coded. Overarching themes drawn from literature formed the basis for an initial broad-level coding, which was then refined and expanded with inductively derived themes. Although one researcher performed the coding, the codes, interpretations, and results were discussed collaboratively by the entire team. All citations from the interviews were presented to the interviewees for verification and approval. Although the interview data were anonymized, it might still be possible to identify the participants due to the smallness of the Swiss media system and the limited number of experts in the field. The interviewees were briefed accordingly and gave their informed consent to participate in the study.

5 Results

During the interviews, it became apparent that journalists' views on deepfakes were shaped by two circumstances: the rarity of the phenomenon, and their sensitivity to disinformation. First, there have been only few domestic deepfake cases, as mentioned by one interviewee: "I am not aware of deepfakes that affect our work directly–for example, manipulated images of a Swiss politician" (I-3). Several interviewees attributed this to the political culture in Switzerland and to the rather low reach and popularity of Swiss media and their journalists on a global scale. Given the rare cases of deepfakes, interviewees were more worried about attempts to cast journalism as untrustworthy. For them, "the constant repetition that the media are fake" (I-2) is more worrisome than deepfakes. Many in the news audience do not trust the media regardless of deepfakes. As one journalist said: "They do not believe our articles even if they are truthful and based on intensive investigation. They think that journalists are lying" (I-1). However, some interviewees would "not completely reject the possibility of an attack on people or journalists who publicly express controversial opinions" (I-7). The only deepfakes that interviewees were

aware of at the time concerned foreign events, especially war propaganda (I-7). One journalist said: "We have had some deepfakes, but we identified them or treated the images as potential deepfakes right from the beginning" (I-6).

Second, the interviewees mostly considered deepfakes as a special form of disinformation. Since the topic of disinformation has been prominently discussed in journalism for several years, the respondents seemed to be especially sensitive to this issue. Consequently, the interviewees discussed deepfakes mostly as a threat for society and journalism. They only cautiously discussed opportunities of the technology, emphasizing the need to comply with journalistic norms and standards.

Against this backdrop, four themes emerged from the interviews: (1) how news organizations are dealing with deepfakes, (2) how media brands and journalists may be misused, (3) how awareness can be promoted among audiences, and (4) opportunities offered by deepfakes or synthetic content.

Dealing with deepfakes in news organizations

Most interviewees considered deepfakes as a form of disinformation and thus as a challenge for journalism. Most of them had encountered only a few deepfakes and advised against overestimating the risks posed by them (I-9). Several interviewees noted that creating convincing deepfakes requires significant effort and that it is difficult to imagine laymen producing perfect deepfakes that journalists are not able to detect (I-1). In the view of one respondent, the risk in Switzerland is lower because deepfakes are less common, journalists are well trained and equipped, and the level of education [of the Swiss population] is comparatively high (I-4). However, many interviewees also expressed concerns. They worried that the existing problem of disinformation could become even worse with deepfakes and warned against underestimating technological developments (I-10) and the danger of deepfakes for journalism (I-5).

To successfully prevent the unintentional spread of deepfakes by news organizations, interviewees emphasized that it is essential to first realize that this kind of manipulation is possible. One journalist explained: "The key point is: Are you aware that this exists? If not, you are going to be fooled" (I-2). Interviewees often mentioned the need to raise awareness in news organizations (I-9). As one of the experts recounted, this can be done, for example, in the form of a news story that discusses how deepfakes are produced, "to understand the technology from A to Z–how much it costs, how it is done, what the technology is able to do, and for what purpose" (I-3).

Since the verification of information is a central element of journalistic work, training and/or further education with respect to fact-checking existed in all the

news organizations included (I-1). Depending on an organization's size and financial resources, such training addressed individuals, teams, or divisions. These efforts touched on the spread of disinformation in general but not specifically on deepfakes. Special training for detecting deepfakes is not yet widely used, as there is a feeling that most deepfakes are still easy to detect (I-4). Interviewees often mentioned that traditional journalistic standards and verification strategies provide "important basics for detecting deepfakes" (I-2) such as "the two-source principle or the careful verification of user-generated content" (I-8). Unknown foreign online sources were perceived to pose the highest risk of deepfakes, so the interviewees emphasized the importance of being able to rely on a network of trusted foreign correspondents and partner news organizations (I-7).

Yet most interviewees were aware that deepfakes add "complexity to the daily work of checking and verifying information" (I-3). Several of them stated that fact-checking for deepfakes goes "one step further because of their complexity" (I-6). One interviewee explained their news organizations' stance toward videos as follows: "I call our people 'bullshit detectors.' First thing we want them to do for every video is to assume that it is fake" (I-7). The interviews revealed that open-source intelligence is seen as an important instrument to detect deepfakes. One big news organization thus invested in relevant training: "All of our editors are trained for three days in the basic principles of fact-checking with open-source intelligence. This includes how to do research and how to spot deepfakes" (I-4).

According to interviewees, many Swiss news organizations have dedicated persons, teams, or divisions dealing with the verification of online information: "They know how to produce a deepfake, and they can detect deepfakes" (I-2). However, the need for expertise seems to be increasing, which is affecting the workload of those specialists who are helping their colleagues "on top of our daily work" (I-8) and who are "constantly overwhelmed with requests from colleagues" (I-4). Thus, larger news organizations have started to expand their specialized units (I-1). Not all the analyzed news organizations are able to do so though, as they are operating under tight economic conditions (I-9). As a result, smaller local media rely more on trustful information from partner organizations and news agencies such as the Swiss news agency Keystone-SDA (I-7).

The need for (even more) thorough journalistic work was often echoed by the interviewees: "In the context of new technologies, it is important to remember that we are not immune to being fooled even if we are doing our job well" (I-5). Verification and thorough journalistic work, however, require time, which is scarce in journalism. In this context, some interviewees also observed a welcome change to journalistic working conditions:

> Journalism used to be about speed, but now I see an opposite trend. A central pillar in the fight against disinformation and fakes is a certain deceleration. It is better to think twice before publishing a piece. We need to slow down a bit. (I-1)

Misuse of media brands and journalists

The interviewees were not aware of any case in Switzerland in which deepfakes were used to discredit journalists. One interviewee is "not afraid that this could happen" (I-2) in Switzerland, not only because Swiss journalists are believed to be less influential and famous than their colleagues in larger countries working for big media corporations (and thus a less interesting target) but also because they perceive Switzerland to be less polarized than other countries (I-5). Although journalists in Switzerland experience criticism as well, some interviewees believed that the situation "is still very civilized" (I-2). However, others mentioned that hate on the internet is an issue (I-8) and considered attacks using deepfakes to be a scary prospect: "If you suddenly show up in a deepfake, not everybody will believe you when you say: 'That's not true!'" (I-1). Some respondents expressed concern that threats have increased in recent years: "This is something we discuss in our team. Yes, we take it seriously. I think in the past, it was often not taken seriously enough" (I-1). To date, there are no specific mitigation measures in place for when a journalist would be personally discredited by a deepfake. Such an attack would be treated as part of already existing strategies to defend a news organization's employees: "Quickly communicating our point of view, positioning ourselves and, if needed, filing a lawsuit" (I-3). Another risk identified by the interviewees was that their own media outlet's content and news anchors could be altered and misused for a deepfake (I-6).

However, several interviewees questioned the risks posed by deepfakes: "This is still a technology that is not completely ready yet and not available to everyone" (I-9). Therefore, journalists contrasted deepfakes with simpler and older forms of information manipulation such as cheapfakes and shallowfakes that distort and negate facts with simple manipulations such as adjustment of the video speed, cutting of footage, restaging of the scene, or re-contextualized content with little resources (Brady and Meyer-Resende, 2020; Paris and Donovan, 2019). One journalist even mentioned that such cheapfakes and shallowfakes might pose a larger problem than deepfakes: "Deepfakes require significant effort, time, engineering, and technological know-how. However, the same goal can be achieved with shallowfakes, which are traditional disinformation campaigns that are very basic yet just as effective in fostering distrust in media" (I-4).

Awareness among audiences

Given the potential dangers posed by disinformation and deepfakes, interviewees identified two important aspects in raising the awareness of their audiences. The first one is to provide transparency and to better explain how journalism works (I-3): "We are trying to build a community around our news organization and to be open and transparent. Readers can take part in editorial meetings and ask questions either online or in person" (I-9). Many interviewees believed that highlighting the added value of journalistic work compared to other sources of information is a sound response to disinformation: "Journalism costs time and money. It is important that the audience also realizes that there is a great deal of work behind quality reporting that has to be paid for" (I-1). The second aspect involves specific strategies to raise awareness among audiences. Reporting on deepfakes was often echoed as a promising strategy: "In some stories, we say: 'This is a deepfake.' Readers learn how to decide which content is potentially a deepfake and they get aware of potential risks" (I-7). Another interviewee explained that in their reporting, they "often discuss general fact-checking measures" to show how to debunk disinformation (I-8). This is seen as an opportunity for journalists to share their expertise with their audience (I-6).

Nevertheless, there might be limitations in promoting awareness among audiences. Media-skeptical individuals might rather trust unverified sources on social networks than news organizations. One interviewee stated: "Once people are down that rabbit hole, it is very difficult for them to get out. People stick to their opinion" (I-4). Some interviewees were even skeptical about whether raising awareness would run the risk of increasing distrust in the media: "When news organizations adopt terms such as 'fake news' and 'alternative facts,' they sort of support the belief that the media are not trustworthy" (I-2).

Opportunities offered by deepfakes or synthetic content

When specifically asked about potential opportunities of deepfakes or synthetic content for journalism, interviewees overwhelmingly expressed critical views. Some interviewees considered the production of non-authentic content by media themselves to be in conflict with the journalistic profession and its norms: "The goal of journalism is to find truth, but this technology makes it possible to modify reality and create an alternative truth" (I-3). They said that rather than making use of deepfakes, the media should "stand up as defenders of the truth, of truth seeking, of the accuracy of facts" (I-3). Others pointed to the crucial importance of journalism for society and emphasized that a critical attitude toward technology

and synthetic content is needed (I-4). One interviewee stated: "I am a fan of working with technology, such as artificial intelligence and algorithms. But deepfakes are something quite different" (I-3).

Nevertheless, interviewees offered some ideas on how to use synthetic content in journalism. For example, deepfakes might be used for short explanatory videos on news sites "but such use needs to be really obvious and readers should not be misled" (I-1). One person mentioned the advantages of deepfakes for undercover reporting, such as for creating fake profiles (I-1). Despite their critical stance toward using deepfake technologies in journalistic work, many interviewees still did not want to miss out on new trends and opportunities: "We need to keep up with this technological progress. Instead of completely shutting it out, we should also recognize and consider the opportunities it presents" (I-1). One solution offered was to make news more attractive:

> It is sometimes challenging to make dull topics, like politics, more interesting. Here, we could employ alternative storytelling techniques to engage people with the content and to present it in an intriguing way to those who might not otherwise consume news or factually relevant information. (I-2)

Apart from using the technology for journalistic work itself, many interviewees saw a major benefit for media in signaling their successful handling of deepfakes. According to the interviewees, the role of the media in society can be strengthened by demonstrating their role as trustful gatekeepers and by explaining what deepfakes are and how to detect them (I-1). One interviewee offered up an example: "There is a famous video of a girl in Kiev showing a missile hitting behind her. We were able to verify the content with 99.9 % probability and explain that to our audience" (I-7).

6 Discussion and conclusion

This study's goal was to offer insights into how representatives of Swiss news organizations assess the risks and opportunities of deepfakes for journalism and how news organizations adapt to these advances in image processing to mitigate the risks associated with deepfakes.

RQ1 addressed how the representatives assess the risks and opportunities of deepfakes for journalism. Most interviewees predominantly considered deepfakes to pose risks. While convinced that Swiss news organizations are well prepared to detect deepfakes, interviewees also feared that deepfakes will further exacerbate the loss of trust in media and the unfounded questioning of established facts by

some users. However, in contrast to previous studies (Compton, 2021; Marconi and Daldrup, 2018; Posetti, 2018a, 2018b; Razek, 2018), Swiss journalists were not particularly concerned about the potential use of deepfakes to attack journalists or the alteration of their media outlet's news programs and anchors, arguing that there are much simpler ways to discredit news professionals or media brands. Indeed, in Switzerland, no attacks on journalists involving deepfakes have been reported to date. However, videos of a public service media news anchor have been used in attempts to commit fraud (Flach, 2024).

Because the term ends with "fake" and the connection they established to disinformation, it was difficult for the interviewees to imagine potential opportunities of synthetic content. Nevertheless, opportunities were explicitly discussed in all the interviews and the participants perceived various possibilities such as making news more attractive, improving visualization, and better explaining facts. However, the interviewees believed that any use of such technology in journalism needs to be guided by journalistic ethics. This finding aligns with Habgood-Coote's (2023) conclusion that deepfakes are not inherently bad but rather present a social challenge concerning how we manage our practices of producing and receiving recordings. Potentially using a more neutral term like synthetic content might allow for less negative discussions about the potential opportunities offered by AI-driven image processing and media production. Still, whether called deepfakes or not, manipulating audio and video can conflict with journalistic norms, thus eliciting negative perceptions of journalists–a key finding of this study. Hence, despite its potential benefits for journalistic practice, generative AI such as deepfakes threatens the most valuable journalistic capital, namely trust (Fink, 2019).

Regarding RQ2, which addressed the measures taken by Swiss news organizations to detect deepfakes, the present research findings partly put previous studies' concerns about increased difficulties to identify information into perspective (Godulla et al., 2021; Vizoso et al., 2021; Yadlin-Segal and Oppenheim, 2021). The interviews revealed journalists' firm belief that a strong orientation toward journalistic norms helps to handle information carefully and to deal with deepfakes. Most of them considered that existing measures against disinformation implemented by news organizations are sufficient to deal with deepfakes as well. The findings reveal a need for specialized individuals or divisions responsible for complex video verification. Given the scarce financial resources in journalism, creating or expanding internal verification teams may not be realistic for all media organizations. There is also uncertainty about how reliable technical tools for detecting deepfakes will be in the future. In addition to having internal procedures for verifying information, interviewees also believed that raising user awareness is crucial. Previous studies discussed the importance of raising awareness about deepfakes and of promoting media trust (Collins and Ebrahimi, 2021; Godulla et al.,

2021; Wahl-Jorgensen and Carlson, 2021; Yadlin-Segal and Oppenheim, 2021), calling for journalism to engage more in media literacy promotion (Kalsnes et al., 2021; Schapals and Bruns, 2022), increase transparency in news production, and enhance the visibility of professional norms and practices (Jahng et al., 2023; Koliska and Assmann, 2021; Schapals and Bruns, 2022). This is in line with the findings of this research, which show that promising ways to promote media trust include providing explanations about deepfakes as well as informing audiences about media verification practices. The interviewees also highlighted the need for promoting media literacy to enable critical source evaluation and increase society-wide awareness of manipulated information. Hence, the interviewees argued that journalistic media could play a crucial role in sensitizing audiences to deepfakes. However, they also pointed to certain limitations, especially that raising awareness needs to be negotiated with other priorities in the news organizations and the role of journalism. Furthermore, some interviewees pointed out that by reporting on deepfake cases, journalists run the risk of bringing too much attention to harmful information and unintentionally supporting propaganda.

This study explores how Swiss news organizations deal with deepfakes. During data collection, most journalists were not yet aware of generative AI. ChatGPT was released at the end of 2022, and only a few journalists in our sample had used it before participating in this study. While more experience with generative AI might have changed the views on pervasiveness of the problem and media's ability to detect deepfakes, the main insights provided by this study regarding the opportunities and risks of deepfakes–as the AI-driven synthetic creation of audio and visual content for journalism–still stand. These core findings remain unaffected, as the study's focus was on how deepfakes are evaluated through the lens of journalistic standards and norms. As the technology becomes more widespread and accessible, the need for specific training, promotion of algorithmic literacy, and ongoing monitoring to ensure responsible AI usage in journalism will increase even more (Cools and Diakopoulos, 2024).

Previous studies show that users, experts, and journalists struggle to recognize manipulations and computer-generated content (Vogler et al., 2024; Lehmuskallio et al., 2019). As technology advances, detecting deepfakes will likely become harder. While some suggest using deep learning techniques (Rana et al., 2022), it is uncertain if these tools will remain effective against sophisticated deepfakes (Ebbers et al., 2024). Given the limitations of technological solutions, the importance of following and renegotiating professional norms becomes evident. This study provides initial insights into how discourses about deepfakes might lead journalists to rethink the essence of journalism and redefine their role in the context of deepfakes. Long-held assumptions, norms, and standards might be subject to discussion, paving the way for institutionalizing new ideas and roles in journalism (Hanitzsch and Vos,

2017). The present research findings revealed that more comprehensive, elaborate context checks are needed to deal with deepfakes. Slowing down journalism, a need long advocated by scholars (e.g., Le Masurier, 2019) and echoed in this study's interviews, might be an example of how a new situation can trigger an idea that might eventually become a new norm.

However, in news organizations owned by commercial media companies, time is money. With many media companies struggling to survive, cost pressures are increasing (Curran, 2010; McChesney and Nichols, 2010; Meier, 2017), which also puts journalists under time pressure and affects their capabilities to thoroughly verify information (Puppis et al., 2017). For commercial media to prioritize professional norms, a major rethink of economic viability is needed. As platforms and messaging services facilitate the spread of disinformation through deepfakes, journalism has the chance to stand out as a trusted source. Emphasizing professional norms and transparency could influence both audiences' willingness to pay for news and news organizations' investment in journalism. Future studies should explore whether and how these changes are implemented in an AI-driven future.

Although Swiss political culture has unique features that, according to interviewees, may reduce the threat of deepfakes, many of our findings have broader relevance for other media systems. Similarly to Switzerland's journalistic culture, other journalistic cultures also uphold core journalistic norms such as objectivity, accuracy, and truth-telling (Hanitzsch et al., 2019), which likely influence how they evaluate the opportunities and risks posed by deepfakes. Journalism worldwide is facing the same challenges, although responses of news organizations may vary according to media system characteristics. Nevertheless, we anticipate many common responses–such as thorough fact-checking–across media systems. Moreover, we observed no significant differences among news organizations within Switzerland or across the two observed language regions, suggesting that our findings could apply in other contexts as well.

Despite its risks, deepfake technology offers intriguing possibilities for journalism. However, interviewees were generally skeptical, likely due to the negative connotations of the term "deepfake." Future studies should aim to better categorize deepfakes and AI-manipulated media. Importantly, not all deepfakes are created with deceptive intent; they can also be used for entertainment, satire, or have artistic merit. Additionally, deepfakes could help personalize and translate news content. It was not surprising that interviewees often emphasized the tension between editorial teams creating manipulated content and upholding journalistic norms. Journalistic norms are being challenged by the rise of deepfakes and need to be renegotiated–a process in which journalists balance, rethink, and revalue various aspects of the oughts and needs of journalistic work (Raemy and Vos, 2021). Discussing potential opportunities of deepfakes led the interviewees to revisit the

question of what journalism should be and what it should do. Furthermore, the existence of deepfakes fuels old ontological debates in journalism about how media can depict reality. This study provided initial insights into when the interviewees considered the line being crossed when it comes to image processing and synthetic content creation. Hence, the negotiation of economic goals, boundaries of manipulation, and journalistic norms become increasingly important when considering possibilities to make the news more attractive. Similarly, future research might investigate whether skepticism toward deepfakes is justified from a normative perspective. This study already shows that adhering to professional journalistic norms is crucial for handling deepfakes responsibly, especially considering media's influence on public opinion formation.

Funder Name: Swiss Foundation for Technology Assessment (TA-Swiss), Funder Id:, Grant Number: https://www.ta-swiss.ch/en/deepfakes
Please use Ringold ID (when given): 27211

References

Ayyub, R. (2018, November 21). I was the victim of a deepfake porn plot intended to silence me. *Huffington Post*. https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316

Bateman, J. (2020, July). *Deepfakes and synthetic media in the financial system: Assessing threat scenarios* [Cyber Policy Initiative Working Paper Series No. 7]. https://carnegieendowment.org/files/Bateman_FinCyber_Deepfakes_final.pdf

Bogner, A., & Menz, W. (2009). The theory-generating expert interview: Epistemological interest, forms of knowledge, interaction. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing experts: Research methods series* (pp. 43–80). Palgrave Macmillan.

Brady, M., & Meyer-Resende, M. (2020). *Deepfakes: A new disinformation threat*. Democracy Reporting International.

Canella, G. (2023). Journalistic power: Constructing the "truth" and the economics of objectivity. *Journalism Practice*, *17*(2), 209–225. https://doi.org/10.1080/17512786.2021

Christians, C. G., Glasser, T. L., McQuail, D., Nordenstreng, K., & White, R. A. (Eds.). (2009). *Normative theories of the media: Journalism in democratic societies.* University of Illinois Press.

Citron, D. K., & Chesney, R. (2018). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*, 1753–1820. https://scholarship.law.bu.edu/faculty_scholarship/640

Collins, A., & Ebrahimi, T. (2021). *Risk governance and the rise of deepfakes.* International Risk Governance Center. EPFL. https://doi.org/10.5075/epfl-irgc-285637

Compton, S. (2021, March 16). More and more women are facing the scary reality of deepfakes. *Vogue.* https://www.vogue.com/article/scary-reality-of-deepfakes-online-abuse#:~:text=Of%20course%2C%20it's%20not%20really,most%20countries%2C%20it's%20entirely%20legal

Cools, H., & Diakopoulos, N. (2024). Uses of generative AI in the newsroom: Mapping journalists' perceptions of perils and possibilities. *Journalism Practice*, 1–19. https://doi.org/10.1080/17512786.2024.2394558

Curran, J. (2010). The future of journalism. *Journalism Studies, 11*(4), 464–476. https://doi.org/10.1080/14616701003722444

de-Lima-Santos, M.-F., & Ceron, W. (2022). Artificial Intelligence in news media: Current perceptions and future outlook. *Journalism and Media, 3*(1), 13–26. https://doi.org/10.3390/journalmedia3010002

Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism: Theory, Practice & Criticism, 6*(4), 442–464. https://doi.org/10.1177/1464884905056815

Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society, 23*(7), 2072–2098. https://doi.org/10.1177/1461444820925811

Duffy, A., & Ang, P. H. (2019). Digital journalism: Defined, refined, or re-defined. *Digital Journalism*, *7*(3), 378–385. https://doi.org/10.1080/21670811.2019.1568899

Ebbers, F., Karaboga, M., Runge, G., & Friedewald, M. (2024). Ist- und Trendanalyse [Status and Trend Analysis]. In M. Karaboga, N. Frei, M. Puppis, D. Vogler, P. Raemy, F. Ebbers, G. Runge, A. Rauchfleisch, G. de Seta, G. Gurr, M. Friedewald & S. Rovelli (Eds.), *Deepfakes und manipulierte Realitäten. Technologiefolgenabschätzung und Handlungsempfehlungen für die Schweiz* (pp. 83–123). vdf Hochschulverlag AG. https://doi.org/10.3218/4186-6

Federal Office of Communications. (2024). *National Action Plan for the safety of media professionals in Switzerland*. OFCOM. https://www.bakom.admin.ch/bakom/en/homepage/electronic-media/media-policy/national-action-plan.html#-1867165064

Fink, K. (2019). The biggest challenge facing journalism: A lack of trust. *Journalism, 20*(1), 40–43. https://doi.org/10.1177/1464884918807069

Flach, L. (2024, June 1). *Fakes mit SRF-Moderatoren: Das steckt hinter der Betrugsmasche*. [Fakes with SRF anchormen: Uncovering the fraud]. SRF Schweizer Radio und Fernsehen. https://www.srf.ch/news/gesellschaft/werbung-auf-sozialen-medien-fakes-mit-srf-moderatoren-das-steckt-hinter-der-betrugsmasche

Gamage, D., Chen, J., Ghasiya, P., & Sasahara, K. (2022). Deepfakes and society: What lies ahead? In M. Khosravy, I. Echizen, & N. Babaguchi (Eds.), *Frontiers in fake media generation and detection* (pp. 3–43). Springer Nature.

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

Gregory, S. (2019). *Preparing for deepfakes against journalism*. WITNESS Media Lab. https://lab.witness.org/wp-content/uploads/sites/29/2019/12/case_23_preparing_for_deepfakes_against_journalism.pdf.

Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese, 201*(103). https://doi.org/10.1007/s11229-023-04097-3

Hameleers, M. (2023). Disinformation as a context-bound phenomenon: Toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory, 33*(1), 1–10. https://doi.org/10.1093/ct/qtac021

Hanitzsch, T., Hanusch, F. Ramaprasad, J., & de Beer, A. S. (2019). *Worlds of journalism: Journalistic cultures around the globe*. Columbia University Press.

Hanitzsch, T., & Vos, T. P. (2017). Journalistic roles and the struggle over institutional identity: The discursive constitution of journalism. *Communication Theory, 27*(2), 115–135. https://doi.org/10.1111/comt.12112

Hellyer, F. (2022, March 22). Deepfakes: The new ticket to immortality? *Rolling Stone*. https://www.rollingstone.com/culture-council/articles/the-new-ticket-to-immortality-1324513/

High-Level Expert Group on Fake News and Disinformation. (2018). *A multi-dimensional approach to disinformation: Report of the independent high-level group on fake news and online disinformation.* https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1

Humprecht, E. (2023). The role of trust and attitudes toward democracy in the dissemination of disinformation–a comparative analysis of six democracies. *Digital Journalism, 13*(5), 931–948. https://doi.org/10.1080/21670811.2023.2200196

Jahng, M. R., Eckert, S., & Metzger-Riftkin, J. (2023). Defending the profession: U.S. journalists' role understanding in the era of fake news. *Journalism Practice, 17*(2), 226–244. https://doi.org/10.1080/17512786.2021.1919177

Jarrahi, A., & Safari, L. (2022). Evaluating the effectiveness of publishers' features in fake news detection on social media. Multimedia Tools and Applications, 82, 2913–2939. https://doi.org/10.1007/s11042-022-12668-8

Kalsnes, B., Falasca, K., & Kammer, A. (2021). Scandinavian political journalism in a time of fake news and disinformation. In E. Skogerbø, Ø. Ihlen, N. N. Kristensen, & L. Nord (Eds.), *Power, communication, and politics in the Nordic countries* (pp. 283–304). Nordicom.

Koliska, M., & Assmann, K. (2021). Lügenpresse: The lying press and German journalists' responses to a stigma. *Journalism, 22*(11), 2729–2746. https://doi.org/10.1177/1464884919894088

Le Masurier, M. (Ed.). (2019). *Slow journalism*. Routledge.

Lehmuskallio, A., Häkkinen, J., & Seppänen, J. (2019). Photorealistic computer-generated images are difficult to distinguish from digital photographs: A case study with professional photographers and photo-editors. *Visual Communication, 18*(4), 427–451. https://doi.org/10.1177/1470357218759809

Leibowicz, C., Stray, J., & Saltz, E. (2020, July 13). *Manipulated media detection requires more than tools: Community insights on what's needed*. Partnership on AI. https://www.partnershiponai.org/manipulated-media-detection-requires-more-than-tools-community-insights-on-whats-needed/

Lewis, A., Vu, P., Duch, R. M., & Chowdhury, A. (2023). Deepfake detection with and without content warnings. *Royal Society Open Science, 10*(11), 1–13. https://doi.org/10.1098/rsos.231214

Lyu, S. (2020). Deepfake detection: Current challenges and next steps. *Proceedings of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW), UK*, 1–6. https://doi.org/10.48550/arXiv.2003.09234

Malik, A., Kuribayashi, M., Abdullahi, S. M., & Khan, A. N. (2022). DeepFake detection for human face images and videos: A survey. *IEEE Access, 10*, 18757–18775. https://doi.org/10.1109/ACCESS.2022.3151186

Marconi, F., & Daldrup, T. (2018, November 15). *How The Wall Street Journal is preparing its journalists to detect deepfakes*. Nieman Lab. https://www.niemanlab.org/2018/11/how-the-wall-street-journal-is-preparing-its-journalists-to-detect-deepfakes

Mayring, P. (2022). *Qualitative content analysis: A step-by-step guide*. Sage Publications.

McChesney, R. W., & Nichols, J. (2010). *The death and life of American journalism: The media revolution that will begin the world again*. Nation Books.

Meier, W. A. (Ed.). (2017). *Abbruch–Umbruch–Aufbruch. Globaler Medienwandel und lokale Medienkrisen* [Breakdown–transformation–new beginning: Global media change and local media crises]. Nomos.

Meier, K., García Avilés, J. A., Kaltenbrunner, A., Porlezza, C., Wyss, V., Lugschitz, R., & Klinghardt, K. (2024). *Innovations in journalism: Comparative research in five European countries.* Routledge.

Meikle, G. (2023). *Deepfakes.* Polity Press.

Montal, T., & Reich, Z. (2017). I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism, 5*(7), 829–849. https://doi.org/10.1080/21670811.2016.1209083

Paris, B., & Donovan, J. (2019). *Deepfakes and cheap fakes. The manipulation of audio and visual evidence.* Data & Society. https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf

Peng, X., & Xintong, B. (2022). An effective strategy for multi-modal fake news detection. *Multimedia Tools and Applications, 81*(10), 13799–13822. https://doi.org/10.1007/s11042-022-12290-8

Porlezza, C. (2019). Accuracy in journalism. In *Oxford Research Encyclopedia of Communication*. https://doi.org/10.1093/acrefore/9780190228613.013.773

Posetti, J. (2018a). Combatting online abuse: When journalists and their sources are targeted. In C. Ireton & J. Posetti (Eds.), *Journalism, "fake news" & disinformation: Handbook for journalism education and training* (pp. 115–127). UNESCO.

Posetti, J. (2018b). News industry transformation: Digital technology, social platforms and the spread of misinformation and disinformation. In C. Ireton & J. Posetti (Eds.), *Journalism, "fake news" & disinformation: Handbook for journalism education and training* (pp. 57–72). UNESCO.

Puppis, M. (2019). Analyzing talk and text I: Qualitative content analysis. In H. van den Bulck, M. Puppis, K. Donders, & L. van Audenhove (Eds.), *The Palgrave handbook of methods for media policy research* (pp. 367–384). Palgrave Macmillan.

Puppis, M., Schenk, M., & Hofstetter, B. (Eds.). (2017). *Medien und Meinungsmacht* [Media and opinion authority]. Vdf Hochschulverlag. https://vdf.ch/medien-und-meinungsmacht-e-book.html

Raemy, P., & Vos, T. P. (2021). A negotiative theory of journalistic roles. *Communication Theory, 31*(1), 107–126. https://doi.org/10.1093/ct/qtaa030

Rana, M. S., Nobi, M. N., Murali, B., & Sung, A. H. (2022). Deepfake detection: A systematic literature review. *IEEE Access, 10*, 25494–25513. https://doi.org/10.1109/ACCESS.2022.3154404

Razek, A. (2018, November 15). *Seeing isn't always believing.* BBC. https://www.bbc.co.uk/blogs/aboutthebbc/entries/bee04c43-d896-4e36-8f02-244cb0db1c08

Rohera, D., Shethna, H., Patel, K., Thakker, U., Tanwar, S., Gupta, R., Hong, W.-C., & Sharma, R. (2022). A taxonomy of fake news classification techniques: Survey and implementation aspects. *IEEE Access, 10*, 30367–30394. https://doi.org/10.1109/ACCESS.2022.3159651

Sammut, F., Bezzina, M., & Scerri, J. (2023). Under attack in the cyber battlefield: A scoping review of journalists' experiences of cyberharassment. *Journalism Practice, 19*(8), 1839–1867. https://doi.org/10.1080/17512786.2023.2294290

Schapals, A. K., & Bruns, A. (2022). Responding to "fake news": Journalistic perceptions of and reactions to a delegitimising force. *Media and Communication, 10*(3). https://doi.org/10.17645/mac.v10i3.5401

Schweizer Radio und Fernsehen. (2024, January 5). *Nationalrat Andreas Glarner muss wegen Fake-Video zahlen* [National Councilor Andreas Glarner must pay a penalty because of a fake video]. SRF. https://www.srf.ch/news/schweiz/klage-von-sibel-arslan-nationalrat-andreas-glarner-muss-wegen-fake-video-zahlen

Sjøvaag, H. (2010). The reciprocity of journalism's social contract: The political-philosophical foundations of journalistic ideology. *Journalism Studies, 11*(6), 874–888. https://doi.org/10.1080/14616701003644044

Sohrawardi, S. J., Sovantharith, S., Chintha, A., Thai, B., Ptutcha, R., & Wright, M. (2020, August 9-11). *DeFaking deepfakes: Understanding journalists' needs for deepfake detection* [Conference presentation]. USENIX Symposium on Usable Privacy and Security (SOUPS) 2020, Boston MA, United States. https://www.usenix.org/system/files/soups2020_poster_sohrawardi.pdf

Swiss TXT. (n.d.). *Less constraints, cheaper and more flexible: The sign language of tomorrow*. Retrieved August 15, 2023, from https://www.swisstxt.ch/en/services/accessibility/sign-language/the-sign-language-of-tomorrow/

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, 73*(3), 369–392. http://www.fclj.org/wp-content/uploads/2021/05/73.3.1-Dont-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.pdf

Udris, L., Rivière, M., Vogler, D. & Eisenegger, M. (2022). *Reuters Institute Digital News Report 2022. Länderbericht Schweiz* [Reuters Institute Digital News Report 2022. Country report Switzerland]. fög – Forschungszentrum Öffentlichkeit und Gesellschaft. https://www.foeg.uzh.ch/dam/jcr:04d547fc-0fc3-42fc-bbf8-cee04a06ebcf/DNR_22_Schweiz.pdf

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1). https://doi.org/10.1177/2056305120903408

van Dalen, A. (2020). Journalism, trust, and credibility. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (2nd ed., pp. 356–371). Routledge. https://doi.org/10.4324/9781315167497-23

Vizoso, Á., Vaz-Álvarez, M., & López-García, X. (2021). Fighting deepfakes: Media and internet giants' converging and diverging strategies against hi-tech misinformation. *Media and Communication, 9*(1), 291–300. https://doi.org/10.17645/mac.v9i1.3494

Vogler, D., Rauchfleich, A., & de Seta, G. (2024). Wahrnehmung von Deepfakes in der Schweizer Bevölkerung [Swiss citizen' perceptions of deepfakes]. In M. Karaboga, N. Frei, M. Puppis, D. Vogler, P. Raemy, F. Ebbers, G. Runge, A. Rauchfleisch, G. de Seta, G. Gurr, M. Friedewald, & S. Rovelli (Eds.), *Deepfakes und manipulierte Realitäten. Technologiefolgenabschätzung und Handlungsempfehlungen für die Schweiz* (pp. 125–151). vdf Hochschulverlag. https://doi.org/10.3218/4186-6

Vu, H. T., & Saldaña, M. (2021). Chillin' effects of fake news: Changes in practices related to accountability and transparency in American newsrooms under the influence of misinformation and accusations against the news media. *Journalism & Mass Communication Quarterly, 98*(3), 769–789. https://doi.org/10.1177/1077699020984781

Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing fearful futures: Journalistic discourses on deepfakes. *Journalism Practice, 15*(6), 803–820. https://doi.org/10.1080/17512786.2021.1908838

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies, 19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Weikmann, T., & Lecheler, S. (2023). Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society, 25*(12), 3696–3713. https://doi.org/10.1177/14614448221141648

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 40–53. https://doi.org/10.22215/timreview/1282

Yadlin-Segal, A., & Oppenheim, Y. (2021). Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence: The International Journal of Research into New Media Technologies*, *27*(1), 36–51. https://doi.org/10.1177/1354856520923963

Zelizer, B. (2019). Why journalism is about more than digital technology. *Digital Journalism*, *7*(3), 343–350. https://doi.org/10.1080/21670811.2019.1571932

Appendix

Overview of the interviewees: Their position and expertise in a Swiss news organization

Interviewee	Position and expertise
I-1	Head of Investigation
I-2	Head of Editorial Department
I-3	Digital News Coordination Manager
I-4	Open Source und Visual Investigations
I-5	Head of Editorial Department
I-6	Knowledge Editorial Department
I-7	Head of the Video Editorial Department
I-8	Head of Fact-Checking
I-9	Editorial Department in Science and Technology
I-10	Product Manager of Artificial Intelligence