Deepfakes in documentary film production: images of deception in the representation of the real

Article

Published Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Open Access

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1080/17503280.2023.2284680

Publisher: Intellect

www.reading.ac.uk/centaur

CentAUR

Deepfakes in documentary film production: images of deception in the representation of the real

Dominic Lees

Published online: 28 Nov 2023.

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Deepfakes in documentary film production: images of deception in the representation of the real

Dominic Lees 🆔

Department of Film, Theatre and Television, University of Reading, Reading, UK

ABSTRACT

Deepfakes are a technological innovation that might be understood to violate the documentary film's relationship with the real. Yet documentary makers have been among the first screen producers to adopt this technology, using it to swap the faces or voices of people they present to their audiences, further disrupting the already complex relationship between the filmmaker and their subject. This paper investigates the growing practice of documentary deepfakes, using two case studies of documentary films from the period 2019–2022 that have used varied forms of deepfakery, examining the intentions of the filmmakers, the technical processes and the ways in which creative choice is both expanded and limited by the technology. An interview-based research methodology provides original insights into filmmakers' rationales when creating deepfakes. The paper reveals the contradictions inherent in deepfake practice, as described by one documentary filmmaker: 'I'm looking at deepfake as a way of telling the truth' (Benjamin Field, producer). A central theme is the ethics of documentary practice using manipulative AI. This paper contributes new insights to theoretical discourse around the digital manipulation of the moving image, discussing the growing disruption caused by deepfakes to documentary film culture and audiences.

Introduction

> I'm looking at deepfake as a way of telling the truth […] or at least making the truth accessible. (Benjamin Field, documentary director/producer)

In autumn 2022 in the UK, BBC3 broadcast a documentary, Jess Davies's (2022) *Deepfake Porn: you could be next*, raising current and urgent concerns about the abusive use of the deepfakes AI system for replacing faces in digital video. The issue of deepfakes had already reached a high profile in UK television through the BBC drama, *The Capture* (2021–2022), a hugely popular thriller that engages speculatively in the near-future possibilities of digital face replacement technology and its potential disruption of law

enforcement and state security. Yet the drama's producers used no deepfakes in the creation of their series; indeed, TV fiction producers have been hesitant to use the technology, with a rare exception being Disney's *The Book of Boba Fett* (2022). In commercial filmmaking, the first use of deepfakes was in 2021, when Bruce Willis gave permission to the firm Deepcake to insert his deepfake likeness into an advert for the Russian mobile phone network, Megafon. In Hollywood movies, despite several examples of face replacement using VFX processes, the first significant use of deepfakes is Robert Zemeckis' film, *Here* (2024), for which the company Metaphysic.ai has used the technology to age and de-age the film's star Tom Hanks in real time (in production at time of writing). Contrast this with the field of documentary, in which filmmakers have pioneered the use of deepfakes, acting as a technological vanguard and creating broadcasts for mainstream terrestrial and SVOD channels. This presents us with a conundrum, since the nature of deepfakes apparently puts the technology at odds with important principles of documentary media. As Craig Hight notes, 'Documentary makers are assumed to be important stake-holders in truth and trust in nonfictional forms of representation and engagement' (2022, 398). So major questions arise: how is it that this ethical problem has not deterred documentary filmmakers? Why is it that documentarists, whose work is concerned with forms of representing the real, have been more ready to adopt 'deepfakery' than their counterparts who create screen fictions for streaming, broadcast and theatrical platforms? The phenomenon of documentary practitioners using deepfakes warrants a concerted inquiry from academics in this field. While the number of documentaries to have used deepfakes is still small, these films provoke important questions in the evolving landscape of synthetic media. This article analyses recent examples of deepfakes in documentaries, using production case study and interview methodologies to understand the creative intentions and technological processes undertaken by key documentary filmmakers working in this field. The choice of methodologies is designed to develop an ethnographic approach to this research, foregrounding the voices of filmmakers within the discourses around deepfakes, and enabling them to express the cultural and ethical concerns that arise from their practice. This focus on the filmmakers themselves follows the methodological example of the MIT Open Documentary Lab, which through its series of open talks (MIT 2020) gave voice to pioneering artist filmmakers (Bill Posters, Francesca Panetta, and others) who have used deepfakes to provoke debate (see also Peele 2018). In its final section, this article develops further the theoretical debates initiated by authors examining deepfakes and documentary (Hight 2022), before broadening the discussion by relating it to debates on the role of AI in documentary filmmaking expounded by Kapur and Ansari (2022).

Defining deepfakes: a technology of deception

Deepfakes are automated manipulations of human images and voices in digital video and are part of the growing field of 'synthetic media'. The dominant understanding of deepfakes is as a technique to swap the original face recorded in a video with that of another person using machine learning, but popular use of the term also links it to the manipulation of lip movements to match altered speech, and to audio deepfakes (or voice cloning) (Ajder and Glick 2021, 9). The origin of deepfakes was in 2017 when machine learning processes were used to swap Hollywood actresses' faces into adult

movies, a phenomenon revealed in the groundbreaking report by Samantha Cole (2017) that exposed the scale of non-consensual abuse of women's images in 'deep porn'. Deepfakes are achieved through the application of artificial neural networks and there are multiple approaches proposed by computer scientists (Tolosana et al. 2020) with recent outcomes demonstrating impressively convincing examples at high-definition levels of image quality. A second, related category of audiovisual manipulation is known as 'cheapfakes', described by Aneja et al. as 'a general term that encompasses many non-AI ("cheap") manipulations of multimedia content, created without using deep learning methods' (2021, 2). Paris and Donovan have charted a spectrum of AV manipulation from the sophisticated, AI-intensive deepfakes to the simple cheapfakes 'that use conventional techniques like speeding, slowing, cutting, re-staging, or re-contextualizing footage' (2019, 6). In this article I will exclude 'cheapfakes' or 'shallowfakes' from the scope of my inquiry, focusing instead on mainstream documentary cultures and broadcast media creating deepfake synthetic media at high resolution.

In all cases, 'deepfake' is a process of altering the recorded subject and representing to the audience a synthesised image (and sometimes voice) of a person that is a composite of the profilmic and the computer-generated. This description could equally apply to the creation of VFX sequences in films. However, deepfakes involve automated image manipulation, in contrast to the creative engagement of the VFX artist in building computer generated images: the machine learning technologist sets up a Generative Adversarial Network (GAN) with key images of subject and target, and then leaves the neural network to generate new images independently. Within this process, deception is embedded in deepfakes at the most fundamental level. GANs are a form of machine learning in which a *generator* neural network is trained to develop fake images to the point at which its rival *discriminator* neural network can no longer distinguish between the fake image presented to it and a real one. This intra-computing self-deception, built into the process of generating deepfakes, mirrors audience responses to deepfakes in which viewers delight in the borderline between the visibly and imperceptibly fake.[1]

Mainstream deepfakes in documentary and factual production have shown an acute awareness of their responsibilities in relation to deepfake deception, very clearly signalling their use of image manipulation. Indeed, the first broadcast deepfake on terrestrial television in the UK, *The Alternative Christmas Message* (Channel 4 2020)*,* was explicitly a warning about the deceptive power of deepfakes. Opening as a broadcast to the nation by the Queen, midway through the piece the nonagenarian sovereign leaps onto her desk and dances, and the programme ends with a reveal demonstrating the techniques of the deception (Figure 1). Nevertheless, despite the honest full disclosure, deception is intrinsic to this and all deepfakes. The pleasures for the audience include the enjoyment of feeling deceived, and appreciation of the technological skill involved in achieving this deception. There is a parallel with our appreciation of the skills of a secular magician conducting card or conjuring tricks, despite our knowledge that what we are witnessing is trickery. Directed by William Bartlett, *Christmas Message* falls into the category of mockumentary and is a scripted production, so sits outside this article's consideration of deepfakes in documentary film, however it highlights issues of deception and the ethical responsibility felt by filmmakers towards their audiences. Levels of honesty, the issue of how a film should signal its deepfake to the spectator, are vital ongoing

Figure 1. Revealing the layers of deception: The Alternative Christmas Message (Channel 4 2020).

debates for producers and broadcasters. A discussion of ethical questions encountered by documentary filmmakers deploying deepfakes will be developed in a later section of this paper.

Scholarship since the inception of deepfakes in 2017 has been dominated by writing around the negative social, gender and political implications (Vaccari and Chadwick 2020). Within computer science, which by volume dominates publications on deepfakes, the emphasis of research is on the detection and prevention of deepfakes (Lyu 2022; Rana et al. 2022; Yu et al. 2021). Danry et al. (2022) and Gaur (2022) represent writings on the science of ethical applications of deepfakes, and in philosophy, De Ruiter (2021) and Rini (2020) have developed evaluations of the moral status of the creation and dissemination of deepfakes. Law scholars, including Pavis (2021), have published on the reforms to legal frameworks that will be required to regulate deepfakes. Research in cultural studies has begun to focus on creative applications of the technology (Lees, Bashford-Rogers, and Keppel-Palmer 2021; Mihailova 2021), and the MIT Media Lab hosted a conference event on 'Putting Deepfakes to Good Use' in summer 2022; in the UK the Synthetic Media Research Network developed these themes with its symposium, 'Synthetic Media and the Screen Industries' in July 2023. A further distinct approach to deepfakes situates them within communication studies: Graham Meikle firmly roots them in this discipline with his declaration that 'Deepfakes are first of all a communication phenomenon: they are about new ways of making meanings and they are also about challenges to settled understandings of how meaning gets made' (2022, 23–24). Craig Hight (2022) brings the discourse to bear on documentary film, relating deepfakes to broad themes of contemporary misinformation culture. He notes that deepfakes sit within a continuity of earlier practices of documentary filmmakers: 'there has always been a tension within documentary practice with an inherent need to manipulate evidence in the pursuit of accessible and coherent storytelling' (2022, 398). The use of

deepfakes thus links to an ongoing discourse in documentary practice, in which the deception involved in image manipulation has been justified by filmmakers' understanding that such techniques serve a greater cause of delivering strong documentary narratives to the audience.

The use of AI processes such as deepfakes within documentary production increases the extent to which the filmmaker is embedded into the online world. The concept of this integration of documentary and its digital context is well-established: Vinicius Navarro has described how the focus of the study of documentary film is changing: 'the emphasis has shifted from individual projects to the environments in which the documentary materials circulate' (2020, 92). But the issue is no longer limited to the circulation of documentary online, a new fluidity of the distribution of content. In this article, we will see how the use of deepfake AI extends Navarro's concept: the online world becomes part of the means of production for the documentarist, with practitioners pulling computer code from open source repositories such as DeepFaceLab and using the internet to find training data that will help generate their deepfake images.

Using deepfakes in documentary practice

In this section, I will examine two case studies of documentary films that have adopted deepfake technology in their production. Although made just three years apart, the films represent one of the earliest and one of the more recent instances of deepfakes in the documentary form. *In Event of Moon Disaster* was directed by Francesca Panetta and Halsey Burgund and released at the International Documentary FilmFestival Amsterdam (IDFA) in 2019; *Gerry Anderson: A Life Uncharted* was directed by Benjamin Field and has streamed on Britbox since 2022. I will use the first film to illuminate the technical processes of deepfake audio and how these are integrated into the production. My interview with Oleksandr Serdiuk, CEO of the Ukrainian company Respeecher that created the deepfake voice of Richard Nixon for the film, enables an understanding of the technological progression of deepfake audio since 2019 and its continuing limitations. The case of *Gerry Anderson: A Life Uncharted* provides an illuminating study of how deepfakes moved rapidly from a status as a pariah technology to being commercially desirable. My interviews with both the director and the film's deepfake/VFX technologist, Christian Darkin, illustrate the accessibility of deepfakes to a small-scale independent production company. The analysis of both case studies develops an understanding of the creative opportunities for documentary film afforded by this technology, as well as the shifts in documentary form that are emerging.

In event of moon disaster

The Apollo 11 mission, the first manned flight to the moon, was fraught with dangers, with the serious prospect that technical error or accident would prevent the astronauts from leaving the moon. As a contingency against the worst outcome, Richard Nixon's speechwriter, Bill Safire, composed an address to the nation that could be delivered by the President 'in event of moon disaster' (Safire 1969). In 2019, Massachusetts Institute for Technology (MIT) commissioned a collaborative project between the XR Creative

Director at its Centre for Advanced Virtuality, Francesca Panetta, and Halsey Burgund, Fellow of MIT's Open Documentary Lab, to use deepfake technology to animate the speech that was never delivered. *In Event of Moon Disaster* (Panetta and Burgund 2019) is a multimedia project that won the News and Documentary Emmy Award for Outstanding Interactive Media: Documentary in 2021. It comprises a website, a seven-minute film, and an installation. For the purposes of this article, my analysis will concentrate on the film element of the project and the creative decisions in its use of deepfakes.

Two elements of audiovisual manipulation were required to achieve the project. These were separate AI processes: the first was voice cloning to create the voice of Richard Nixon delivering a speech that he never made; the second was deepfakes to create a video of the speech from an alternative piece of contemporary news footage. The filmmakers worked with two tech companies, Respeecher to build the audio and CannyAI to create the video. In 2019, the machine learning for deepfakes was in its infancy, and it is instructive to observe how the state of the technology imposed severe limitations on the creative choice of the filmmakers. Halsey Burgund has described how CannyAI outlined the restrictions within which deepfake image-making could operate:

> they gave us explicit instructions as to what they needed from us and told us that the target video (the video of Nixon that we wanted to manipulate) had to have certain characteristics. It basically needed to be a still shot of Nixon talking. No close ups, and motion. (qtd in Pietrobon 2020)

Panetta and Burgund had already chosen archive footage of Nixon that included a slow zoom, but this had to be abandoned. Choice of the target video was foundational to this project; indeed, in creating a deepfake, the target video is of major significance in determining the ultimate qualities of the synthetic media. While several authors have written about *In Event of Moon Disaster* as an important moment in the development of deepfakes (Ajder and Glick 2021; Mustak et al. 2023), insufficient critical commentary has focused on the archive that the filmmakers ultimately chose as the target video. Panetta and Burgund chose the resignation speech made by Richard Nixon following the Watergate scandal, delivered on August 8, 1974. The choice was an important one, because this speech embodies different but parallel emotions to those inherent in the Safire speech. In his resignation broadcast, Nixon is dispirited, defeated, but maintains a strong element of defiance. These very personal expressions have become elements of the *Moon Disaster* film artefact that we view, the politician's emotions now translated into a moving screen performance within the mise-en-scene of the documentary footage. In this use of performance, *In Event of Moon Disaster* is particularly relevant to developing theories of documentary performance (Lyons 2020; Marquis 2013). A major creative achievement of Burgund and Panetta's work is identifying this speech as the target video, enabling the emotions of Nixon in 1974 to translate into the tragedy of the alternative history that they are telling in their film. *In Event of Moon Disaster* is an intermedial adaptation between political archive and deepfake synthetic media. In the next part of this section, the creative and technological practice involved in making this adaptation are examined, with particular focus on the audio process.

Respeecher – voice cloning the president

The opportunity that new machine learning technologies provided to Panetta and Burgund was to have Richard Nixon's own voice speak words that he never uttered. This ability to create speech patterns, phonemes and mannerisms that are indistinguishable from the real voice is at the heart of the 'fake'. In the hand of bad actors, the technology is an irritating disruption of political discourse (see the frequency of fake speeches by Joe Biden in 2023).[2] However, when creatively deployed by an experimental artist and curator such as Francesca Panetta, an expert in immersive storytelling, the technology becomes capable of more nuance and greater meaning, as well as allowing her to move across genres and forms. In her interview with this author, Panetta describes deepfakes in this documentary:

> For me, deepfakes offer the opportunity to imagine both speculative histories and speculative futures. They can be documentary-like, but the creative use of synthetic media can also help us enter the grounds of magical realism. Technology has the power to blur the boundaries of truth and fiction, create ambiguity between reality and non-reality.

Respeecher is a company based in Kjiv, Ukraine, which has established itself as a world leader in voice cloning, marketing its expertise to film, video and online content creators. The CEO of Respeecher, Alex Serdiuk, and the company's lead for Ethics and Partnerships, Anna Bulakh, contributed interviews for this article. The role of Respeecher in creating *In Event of Moon Disaster* was early in the workflow. Their work began with audio 'raw material' – Richard Nixon's speeches – provided by the team at MIT. These voice files became a dataset that could be used to train the machine learning algorithm to learn the characteristics of Nixon's voice. Serdiuk describes the next stage of the process: 'Then we needed another human who would be a performer for Nixon, who would be doing voice over for new lines'. Panetta and Burgund hired the actor Lewis D. Wheeler to voice the Bill Safire speech. Serdiuk describes him as,

> a great performer who was able to reproduce the speaking style, the accent the way how Nixon spoke back then. And we asked that performer to record the same data set we had for Richard Nixon, so he had to go through sentence by sentence, piece by piece, recording the same data set after Nixon. Keeping the emotions the same.

This comment contains two important points about the creation of synthetic media characters. Unlike normal acting roles that Lewis D. Wheeler plays, in which his performance is a personal interpretation of the character in which he is cast, for *In Event of Moon Disaster* he was asked to deliver the Bill Safire speech using a careful impersonation of Richard Nixon. The performance required Wheeler to imagine how Nixon would have delivered the *Moon Disaster* text, and to record this faithfully, recreating the President as a man, not as a fictionalised character that he would play in a biopic. This principle of impersonation narrowed the gap between the target video and source, aiding the work of the voice conversion system that was then deployed to combine the actor's delivery and the voice of Richard Nixon.

The second insight that Serdiuk provides is that a key feature of synthetic media technology is its limitation in terms of replicating human emotions. Deepfakes are synthetic media, but for the audience they are screen performances and will be expected to exhibit

natural emotive features, without which the synthetic characters will be unconvincing or robotic. Serdiuk emphasises that,

> with emotions we rely on humans, so humans are best in terms of the exact way they have to perform and we basically don't change emotions when we do our conversion, we just apply a different vocal apparatus. So it's usually a question of casting a good actor who can perform in the exact way or can reproduce the particular emotions in the speech that are common for that target voice.

Respeecher signals its capacity to preserve the human emotions found in the source data through the quality of its proprietary 'speech to speech' technology. This is markedly different from 'text to speech' technology, which frequently sounds robotic – inevitable in a process that lacks the intervention of a human actor. One of the most impressive features of *In Event of Moon Disaster* is its emotional force, derived from the eloquent power of Bill Safire's script, the real human emotions in Richard Nixon's resignation, and the sense of tragedy that Lewis D. Wheeler brings to his vocal performance.

Discussing the project four years after its creation, Serdiuk outlined how much has changed with the technology. Speed and accuracy have improved, with processes that took weeks in 2019 now taking Respeecher just days: 'the system changed like 90%. So we are using the same approach but the way how we train our system, the requirements of data, the speed, quality of the output, robustness of the system – that all has been improved'. However, the central limitation of the technology, the production of emotion in cloned voices, remains. Serdiuk believes that this will be an enduring feature of voice cloning: 'I personally don't believe that performance, human performance, is something that could be reproduced to a fine grade within technology'.

President Nixon's lips

The role of the company CannyAI in the project was to manipulate the footage of Nixon's resignation speech so the President appears to be delivering the Bill Safire text. Deepfake is popularly understood to mean the swapping of faces in digital video, but such complete transference was not required for *In Event of Moon Disaster*. CannyAI has developed a neural rendering technology called Video Dialogue Replacement (VDR), a process which can replace a person's face with a new face of the same person, but speaking different words. CannyAI trained their system with the archive footage of President Nixon and the Respeecher audio; it then 'hallucinated' alternative images of Nixon's head in which the lip movements of Nixon match the speech patterns of the new voice. Finally, VDR returns the new head to the original video. To the viewer, the only visual difference between the original resignation speech and the Bill Safire speech is the lip movements of the president (Figure 2).

In the creation of a synthetic voice and deepfake video for *In Event of Moon Disaster*, machine learning technology was a unique enabler for the documentary filmmakers, allowing the deepfake President to deliver a speech that Nixon never spoke. However, we have also seen a key limitation of the technology: the deepfake process is dependent for its believability on the qualities of the actor who first delivers the script. The AI was able to deliver the emotions of Nixon's voice, but only by matching the performance of Lewis D. Wheeler. The important insight is that the affective quality of a synthetic character rests on the human skills of an actor.

Figure 2. Resignation speech of President Richard Nixon, August 8, 1974.

Resurrecting Gerry Anderson

The documentary film, *Gerry Anderson: A Life Uncharted* (Benjamin Field 2022), produced by The Format Factory and Anderson Entertainment, was commissioned by Britbox and streams on the platform. Its subject is the creative and family life of British children's TV programme maker Gerry Anderson, responsible for cult shows including *Thunderbirds* (ITC Entertainment 1965–6) and *Captain Scarlet and the Mysterons* (ITC Entertainment 1967–8). Although the documentary features a deepfake Gerry Anderson, the filmmakers did not set out to use the technology when they began development of the project. The producers were Benjamin Field and Jamie Anderson, the youngest son of the puppet master. They felt certain that the personal story of this very private man and hero of children's TV had not been revealed, and to support the documentary project the Anderson Estate could provide twenty-five hours of interview audio recordings that had never been heard by the public. The opportunity for a unique documentary insight into Anderson was clear. In an interview for this author, Field made clear that the commissioning editor of Britbox, Craig Morris, 'was pretty much sold on the idea of a documentary about Gerry way before we mentioned them [deepfakes]'. However, as director of the film Field grappled with the means to make the very old audio relevant to his film audience. He even considered using puppets to voice the twentieth century audio recordings of Gerry Anderson, but his co-producer assured him that the family would not approve the technique. These discussions were in 2021, shortly after Chris Ume released his groundbreakingly convincing deepfakes of Tom Cruise, which spread virally on TikTok (2021). Field comments on the impact:

I'd seen that just at the time that we were looking for ways to work with the audio archive that was supplied by the Anderson Estate when we were making *Gerry Anderson: A Life Uncharted* and the two just clicked. I thought: Right, actually this could be a way of bringing audio archive to life.

Field had identified a particularly useful application of deepfakes in documentary production, as a means of animating audio recordings. Sound archives are a vital form of audio testimony, relevant to multiple modes of documentary film and factual content, but extensive use of these recordings stretches the ability of filmmakers to create sufficient visual sequences as overlay for the voices. Using deepfakes to create an artificial to-camera interview with a documentary subject, speaking their words from the audio archive, can solve this problem.

Using deepfakes to put an image to an existing voice file was an early use of the technology, but almost always the audio was linked to an existing video clip. One of the playful uses of deepfakes is found on popular apps such as Reface and widely gamed on YouTube, in which people add their own face to an existing audiovisual text, most frequently lines from Hollywood movies. Benjamin Field was proposing a more complicated process, with his originating material being no more than old-fashioned audio cassettes. The deepfake project undertaken by The Format Factory contrasts with *In Event of Moon Disaster* in two significant ways. Firstly, it comes at a later stage of the technological development of deepfakes, allowing the filmmakers to use higher precision AI tools; second, the project was undertaken by a very small content production company, without the access to skills, technology and resources that were available to Panetta and Burgund through their work at MIT.

Following the commission by Craig Morris at Britbox, the producers sought further funding. A contract was signed with Abacus Media Rights which included a surprising stipulation: the completed film must include a minimum of ten minutes of high quality deepfake material. To Ben Field's surprise, the novelty of deepfake technology had become a marketable feature of this documentary. Field cast the actor Roly Hyde to be the body double for Gerry Anderson. It was important that the actor be as close to Gerry Anderson in terms of head size and shape, and this posed two considerable problems: Anderson was bald, unlike Hyde, and their head shapes were not a perfect match. The team decided to use traditional analogue makeup and hair techniques to narrow the gap between the two. In a two-day film shoot, a shot was composed of Hyde on a sofa, in an interior interview setting; the actor's role was to speak a total of 53 min of chosen material from the audio archive of Gerry Anderson, trying to lip sync as closely as possible to the original.

Responsible for creating the deepfakes for *Gerry Anderson: A Life Uncharted* was Christian Darkin. Darkin worked with an open source deepfake algorithm, DeepFaceLab, that had 'already been trained on hundreds of different other people's faces so [it knows] what a face looks like'. With this head start, Darkin could begin the specific process required by the *Gerry Anderson* film – 'you then have to retrain it on the person that you're trying to replace and the person you're trying to replace them with […] it's learning how to produce the combination of the two'.

However, with the restrictive resources available to The Format Factory, the 'training' process was particularly slow, and this would necessitate a change in approach, limiting the scale of digital transformation that Darkin was asking the AI to perform: 'I ran that
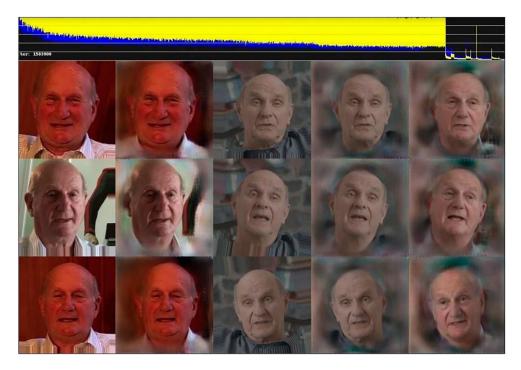
Figure 3. Actor Roly Hyde being 'deepfaked' into the late Gerry Anderson (Photo: Christian Darkin).

process for about three or four weeks on the whole head, and it still wasn't good. But then I tried it on just the face and after a couple of days it was looking very good' (Figure 3). Darkin emphasises that creating a deepfake screen character involves two separate technical processes: first, the AI neural rendering, and secondly with traditional Visual Effects (VFX), Darkin's other area of expertise. The latter is necessitated because the output generated by the deepfakes process leaves glitches in the image. In their overview of the full range of technical processes of deepfakes, Seow et al. note that 'Due to the instability of GAN training, most deepfake outputs consist of subtle traces or fingerprints, such as unusual texture artifacts or pixel inconsistency' (2022, 367). An example in this case study was the visual artifacts where the new face was attached to the target image. Darkin's deepfake process generated a new face of Gerry Anderson, then superimposed this onto the video frames of actor Roly Hyde in the staged interview, but this left clearly visible joins within the composite image:

> where you attach the two together, there's a line right between the two. You can blur that line so it's less noticeable. You can colour correct the skin of the original picture with the skin of the new stuff that you'll replace it with and make it match better.

Resolution of the image was a further problem for the filmmakers. Due to the computing power required to generate deepfakes, in the technology's early years they have been produced at very low resolution. This is an insignificant issue when disseminating deepfake videos on mobile devices, but very serious in the context of HD broadcast. While research is developing strategies for delivering high resolution deepfakes (Naruniec et al. 2020), Darkin found that using the tools available to him, the maximum face image resolution

Figure 4. Production still from the set of Gerry Anderson: A Life Uncharted: period television set used to present deepfake Gerry Anderson.

that he could generate gave a width of 320 pixels, compared with the full HD screen width of 1092 pixels. The implication was that the size of the deepfake Gerry Anderson's head in the frame would need to be very small. Benjamin Field devised a solution: in the film, the deepfake archive clips of Gerry Anderson would be seen on old televisions from the 1960s, within a wider set that would give viewers a sense of the period when the puppet animator was creatively active (Figure 4). Darkin was impressed by his director's decision:

> That was a clever, creative way of doing it, but basically what it enabled us to do was zoom out a little, push the deepfake a little bit back into the distance, reduce that number of pixels that it had to cover.

Furthermore, by introducing 1960s period elements into the documentary's mise-en-scene, the synthetic images could be given a similar contemporary look. Darkin would exploit this as a means to mask any irreparable visual artefacts created by the AI: 'if there were bits where it wasn't going to work, we could put a bit of static on it'.

In this case study, we have seen how deepfakes can be deployed by documentary filmmakers as a creative tool in animating the audio archive. By 2022, deepfakes had also become a marketable feature of a documentary, signalling that filmmakers had now found positive applications for the technology in spite of its dominant use online

as a means of victimising women in non-consensual porn. The technical restrictions faced by Christian Darkin in terms of image resolution have since been overcome, with high resolution broadcast deepfakes (*Deep Fake Neighbour Wars*, ITVX 2023), but his experience demonstrates the barriers that small production companies face when adopting new technology. The imperfections of deepfake image generation confirms a key feature of almost all deepfake practice: the need to improve the AI's output using traditional VFX.

Ethics in the use of deepfake technology

A discussion of the ethics of using deepfake technology within documentary film intersects with ongoing debates on the ethics of the digital (Floridi 2018) and specifically of Artificial Intelligence (Ashok et al. 2022), in particular the problems raised by machine learning (ML). Some authors pursue an ideal, regulated response to the technology, 'enabling the so-called dual advantage of 'ethical ML' – so that the opportunities are capitalised on, whilst the harms are foreseen and minimised or prevented' (Morley et al. 2020). The filmmaker-centred methodology of this article enables a more pragmatic approach, allowing us to listen to the personal responses of practitioners in grappling with the ethical dilemmas that are inherent in the use of synthetic media technologies.

The use of deepfakes in documentary filmmaking raises fundamental ethical questions, as Joshua Rothkopf (2020) states in the simplest terms: 'why should we trust a documentary that uses deepfakes?' The challenge is that a filmmaker deploying this technology undermines their film's claim to be a valid representation of its subject. Rothkopf was discussing the use of deepfakes in the documentary, *Welcome to Chechnya* (David France 2020), a film about the murderous pogrom against gays and lesbians in that country. Director David France used deepfakes to obscure the identities of the persecuted subjects of his documentary, allowing them to tell their stories without risk of identification and violent retribution. Like every mainstream documentary using deepfakes, a clear disclaimer is used at the opening to inform the viewer that they will see synthetic images in the film. France does not like the term deepfake: he has described it as a tool to change what people do, whereas his use of the technology liberates people to be themselves – his disclaimer says that they have been 'digitally disguised', not deepfaked. The sense of disguise was made clear by the film's VFX supervisor, Ryan Laney, who created a blurred effect for the disguised subjects, with a digital 'halo' around their faces, in order that the audience maintains an awareness of the technological disguise throughout the film. Such honesty towards the audience, a clear intention not to deceive, is one response of documentary filmmakers to the ethical issues surrounding the technology.

A primary issue for the producer is whether or not the documentary should use deepfakes at all. In the preproduction of *Gerry Anderson: A Life Uncharted*, the role of the subject's son, Jamie, as co-producer created a unique ethical environment for the film's director. In interview, Benjamin Field describes how this was a deciding factor in his adoption of AI.

> *Dominic Lees:* Do you think you would have used deepfakes if Jamie Anderson hadn't been involved? And if you didn't have that trusting relationship with him?

> *Benjamin Field*: No. I don't think we would have gone ahead with *Gerry*, because Gerry Anderson was incredibly private as an individual and his archives were locked away. He was very controlling over what was known about him publicly.

Throughout the making of the film and the creation of the deepfake Gerry Anderson, the role of Jamie Anderson as co-producer ensured in-built monitoring of his father's post-mortem rights and privacy. As co-producer/director, Field could explore an ambitious range of options for the visual realisation of the film, knowing that his more outlandish ideas (such as using a puppet to voice Gerry Anderson) would be ruled out by his close collaborator if they violated the values of the family and the man. When Field broached the concept of creating a deepfake Gerry, he was surprised at the enthusiasm of Jamie Anderson. They both felt that because the great children's TV pioneer was an enthusiast for new technology, the use of deepfakes in a biographical documentary about Gerry Anderson would be highly appropriate. Christian Darkin, in his role as technically responsible for the deepfakes, saw no ethical conflict in deepfaking Gerry Anderson: 'It was his son who was in charge of the production, we were using audio of interviews that were freely given, that you know clearly those were things that Gerry Anderson wanted to say and wanted to have heard'.

In discussing his work as a documentary filmmaker, Benjamin Field exhibits high levels of concern about the ethics of his practice. Framing a filmmaker's proposal to use deepfakes is the predominant use of this AI in non-consensual pornography: a 2019 research report found that 96% of deepfakes online were pornographic, and 100% of these were video images of women (Ajder et al. 2019). In interview, Field recalls the response of film commissioners when he first started to discuss using deep-fakes in 2021:

> You can see people visibly flinch in meetings [...] what Deepfake has struggled with for some time is being tarnished by negativity [...] it's made deepfake a dirty word or dirty term whereas actually deepfake can be a very useful tool in the armoury of a filmmaker and in an ethical way, it can be a great tool at our disposal.

To counter the negative perception of deepfakes, Field chooses to use the term 'ethical deepfake' in his work as a documentary filmmaker, distancing his own practice from the majority use of this technology.

A key ethical issue facing the producer is the rights of the actor who performs to camera. Following the filming of the actor's performance, their face is replaced by the deepfake technologist, raising major questions: what remains of the original perform-ance? What are the moral and legal rights of that performer? Pavis has discussed this problem in detail and describes how AI technology has outstripped the boundaries of UK law:

> Deepfake technology achieves something no recording technology has done before: they are able to produce high-quality, low-budget, realistic *imitations* of performances on scale. The imitation, or reproduction, of a performance is not protected by performers' rights, or any other intellectual property right strictly speaking. (2021, 849)

The reality is that filmmakers who decide to incorporate deepfakes into their work are operating in an environment of legal uncertainty, with a lack of clarity as to the responsibilities and obligations on those involved in creating synthetic media.

Legislation around deepfakes frequently relates to the control of abuses outside the context of mainstream media production (cyber crime, bullying, fraud, defamation); recent international agreements on AI, including The Bletchley Declaration (2023), have failed to address deepfakes (the word 'media' never enters the text of the Declaration). Richard Arnold (2022) has described the inadequacies of UK law in relation to deepfakes and performer rights, however such legal attention is on the nonconsensual use of images and voices. Most documentary filmmakers, including Benjamin Field, secure the agreement of their subjects (or in posthumous cases their estate) before filming, meaning the subjects surrender some of their customary personality and intellectual property rights by virtue of their consent. Naturally, consent is not a blank cheque for documentarists, but existing law gives insufficient guidance on filmmaker responsibilities towards subjects when deepfaking their image and/or voice. A complicating phenomenon in making deepfakes is that the character observed by the audience on screen is not a recording of a single performance. There are two individuals involved in a deepfake performance, the actor and the target individual, neither of whom is authentically represented in the resulting deepfake. In the process of deepfaking, the machine learning AI reads the target face and maps the movements of the originating performance onto it, creating what Pavis calls an 'imitation' of that target individual.

In the production of *Gerry Anderson: A Life Uncharted*, the documentary producers were faced with two layers of rights in the creation of their deepfake: that of the audio and that of the performer. Clearance to use the audio interview with Anderson could be obtained in a manner familiar to documentary filmmakers, however, rights issues meant the contracting of an actor to perform the interview was complex. Other documentary filmmakers have hired actors, for instance in order to film reconstruction sequences, but this deepfake project was a legally and definitionally difficult proposition. Benjamin Field discusses the difficulties he faced seeking to hire a performer to play Gerry Anderson in his film:

> The first three refused because nobody knew what deepfake was. We couldn't define what the role was – were they acting? Were they a body double? What were they? How are we going to pay them? There are different rates for different jobs and nobody could tell me whether they were acting, because it's not their voice but it is their movement.

Field found that actors' agencies were confused about the nature and status of the work that their clients were being asked to undertake, contributing to the loss of two potential candidates for the role. Agents who could understand how their clients might act in motion capture to create fantasy creatures found the deepfake process a more problematic concept. Unlike the complete transfiguration of Andy Serkis into Gollum in *The Hobbit* (Peter Jackson 2012), their clients would remain themselves onscreen except for their physiognomy. The problem demonstrated how the arrival of AI-generated synthesised performance into the mainstream requires a dramatic upskilling and knowledge transfer across the screen industries, a process that is in its naissance. Amongst actors, there is a broad lack of awareness about their rights when delivering a performance that will subsequently be adapted by Artificial Intelligence. In the UK the actors' union, Equity, is engaged in a campaign, *Stop AI Stealing the Show*, that responds to this challenge. In 2022, Equity reported that '79%

of performers who have undertaken AI work felt they did not have a full understanding of their performers' rights (as set out in the Copyright, Designs and Patents Act 1988) before signing the contract' (Equity 2022, 4). As a result of this confusion around the definition of deepfake performance and the rights of actors, a documentary producer must rely on their personal moral judgement to guide their actions during a film's production. Lacking industry self-regulation guidelines and with no relevant legal framework, it is only the individual sensitivity of today's mainstream filmmakers that restricts the use of a technology that has already been proven to be a destructive tool in the wrong hands.

Authors have considered how the deployment of responsible practice strategies is key to successful innovation businesses in ICT and creative AI (Flick and Worrall 2022; Stahl, Timmermans, and Flick 2016). For companies working on the creation of deepfakes in media content, the ethical challenges posed by the technology can create a risk to the success of their enterprise. The founders of the voice cloning company Respeecher understood from the outset in 2018 that trust was central to their business: if clients lacked confidence in their ethical practices, the commercial basis of the firm would be undermined. Co-founder Alex Serdiuk describes how,

> the ethics statement is something we started Respeecher with, that's the first thing we built in the company. And our ethics statement consists of several important things like having permission, not being involved in letting our technology be used for deceptive uses.

The company appointed a director of ethics, Anna Bulakh. She emphasises that when a business proposal is offered to Respeecher, an ethical evaluation leads to 80% being rejected. Importantly, the concerns discussed seldom relate to legal compliance: 'many questions are not strictly legal ones, the majority of them at this stage are ethical'. A further feature of the company's practice is that Respeecher has decoupled from a dominant trend in which developers share open source machine learning code. Serdiuk asserts that 'it's not just about protecting our technology, it's also about protecting society': the company does not see social benefit in allowing highly convincing voice cloning technology to become ubiquitously available.

In the context of the legal uncertainty and evolving regulatory frameworks for the use of deepfakes, the response of some stakeholders in the screen industries has been to draft proposed guidelines to be followed by responsible practitioners. 'Partnership on AI' (PAI) is an organisation working broadly on governance issues in Artificial Intelligence. In 2023, it launched its 'Framework for the ethical and responsible development, creation, and sharing of synthetic media' (Partnership on AI 2023). The document establishes strong principles governing the conduct of filmmakers using deepfakes, for instance to 'Disclose when the media you have created or introduced includes synthetic elements especially when failure to know about synthesis changes the way the content is perceived' (2023, 5). The organisation describes itself as representing stakeholders, but does not list these and says that it will not audit those associating with PAI. Its 'Framework' very accurately addresses multiple ethical issues pertaining to synthetic media technology, however its impact will be dependent on widespread uptake and responsible implementation – issues of concern in many forms of industry self-regulation.

Concluding discussions

Documentary deepfakes and docudrama

In the analysis of the two documentary films studied in this article, we have seen how both used deepfakes to manipulate archival material. This use of archive creates a strong link with docudrama, while the application of technology enables interesting discontinuities with this tradition. In docudrama, archival footage has been used to connect reconstructions of the past with audiovisually recorded history. Oliver Stone, for example, uses contemporary 1963 archive in his docudrama about the assassination of John F Kennedy, *JFK* (1991); the TV miniseries *Nuremberg* (TNT 2000), opens with black-and-white archive of Adolf Hitler's Nuremberg rallies and later inserts jagged cuts to archive shots of Auschwitz within the staging of its trial scenes. Neither of the two case studies in this paper fits closely into a definition of docudrama, but the use of archive reflects that tradition while the application of deepfakes allows the films to shift across forms. Aspects of the opening of *In Event of Moon Disaster* position it as a formal documentary, with archive of Armstrong and Aldrin boarding the space rocket at Cape Canaveral, grainy footage of Mission Control, and a waving President Nixon at the Apollo 11 launch. The use of archive invokes generic familiarity but, countering this, the 1969 footage is intercut with the film's disclaimer captions – 'What you are about to see is not real'. The effect is to transform the film into an alternative version of docudrama, a reenactment of a possible history, based on events that never occurred. It is AI technology that enables the film to inhabit this highly original space within the docudrama tradition. Panetta and Burgund's film is also linked to this current of filmmaking by its speculative character: similarly speculative was Oliver's Stone's *JFK*, with its proposal of an alternative to the accepted history of John F Kennedy's assassination.

The making of *Gerry Anderson: A Life Uncharted* demonstrated strong links to the production practices of docudrama. In preproduction, director Benjamin Field undertook the tasks of dramatic reconstruction, casting an actor to play Gerry Anderson and designing a set for the interview scene that was carefully designed to match the period of the animator's life. Such historical reconstruction sits securely within the traditions of docudrama, however instead of scripted dialogue, the deepfake technology enabled the filmmaker to use verbatim text for the staged scene. Field's film is thus linked to an established and growing practice of applying verbatim speech, described by Derek Paget as emerging since the 1990s 'in part due to a zeitgeist crisis in representation' (2011, 2), with examples across theatre, film and television. Creators of docudramas have inserted verbatim speech seamlessly into their reenactments of history. Craig Mazin, writer/showrunner of the docudrama *Chernobyl* (HBO/Sky 2019), told reporter Drew Schwartz that he used recorded dialogue of the nuclear reactor supervisor, Aleksandr Akimov, in his script:

> Akimov says, "We did everything right," and immediately following the explosion says, "Something strange has happened" – he said that. *That's what he said*: "Something strange has happened." I can't come up with a better line than that. (Schwartz 2019)

Deepfake technology provides Benjamin Field with an additional creative opportunity that delivers a further layer of authenticity: he can insert the verbatim dialogue of Gerry Anderson exactly as the historical character spoke his words.

Theoretical approaches to digital and AI manipulation

Studies of technologies in digital film have examined varied levels of manipulation, from what Ailish Wood called 'pixel-level micromanipulations' (2007, 92), which might include elements of digital colour grading, to the broader context of Lisa Purse's study of the digital composite (2018). The advent of synthetic media and deepfakes extends our consideration of the manipulation of digital images, the automation of the process through machine learning creating a new context of mass manipulation. The potential is for heightened levels of disruption of moving image cultures. Purse suggests a possible fragmentation caused by digital processes of altering the image, through the 'multiplying perspectives and orientations that digital media culture can provide' (2018, 167). While her attention is on the impact on mainstream feature films, this multiplication of perspectives represents an even more profound transformation within documentary culture. Creating deepfakes involves a layering of multiple images. In the example of *Gerry Anderson,* we watch two people give testimony, Gerry Anderson and the actor Roley Hyde who embodies the recorded voice. The role of the actor can never be neutral: Roley Hyde provides embodiment and physical performance that contribute to our understanding of Anderson's oral testimony. We are asked to trust that Hyde's performance is true to the intentions of Anderson in his original recording, but something niggles. The actor gave his performance after the death of his subject and the two men never met: here is an unavoidable multiplication of perspectives.

Despite the instability of the digital context, even in the era of AI manipulations, the documentary form may still retain much of its historical purpose. Jihoon Kim asserts that 'digitally manipulated images do not necessarily abandon documentary cinema's epistemological and aesthetic functions derived from its photochemical stage' (2022). Kim builds from Michael Renov's concept of 'documentary disavowal' (2004), the questioning of the representation of reality, yet he is optimistic about the impact of digital manipulation:

> digitally graphic and manipulated images do more than verify documentary disavowal inasmuch as they perform various rhetorical functions of documentary other than casting doubt on the truth value of a documentary image: they can be informative, persuasive, and expressive with regard to the chaotic and uncertain faces of reality. (2022)

Discussions of the impact of AI on documentary film extend the ongoing discourse on the digital within documentary practice and culture. Kapur and Ansari note how, 'The digital turn has had a decisive impact on how and what kind of events, data and experiences are (re)presented or (re)told via documentary media' (2022, 174). This is particularly pertinent to the case studies in this paper: we have seen how digital deepfake technology allowed filmmakers to extend the audiovisual representation of events and experiences. Panetta and Burgund were able to transform a written text into film, documenting not a historical event but a nearly-possible moment in US history. Ben Field was able to liberate an archival audio document and transform it into an audiovisual representation of Gerry Anderson's testimony. Each of these illustrates the impact of

deepfakes on the culture and possibilities of documentary film, expanding the form's capacity to articulate alternative histories. This illustration of the potential value of deepfakes to documentary culture reflects the statement of director Ben Field: 'I'm looking at deepfake as a way of telling the truth […] or at least making the truth accessible'.

Notes

1. Online responses to Chris Ume's highly convincing Tom Cruise deepfakes in 2021 included comments such as: Stokes Wait Bro he looks so real! I'm like wtf lol; Jennie Good Is this the real Tom? Or a fake? https://www.youtube.com/watch?v=nwOywe7xLhs&t=87s Accessed 4 Jan 2023
2. https://twitter.com/vince81viny/status/1623256000160534529; https://www.facebook.com/keith.gradwell.1/videos/3391975684453344?idorvanity=1339997533158747; https://www.tiktok.com/@weshmglckmt/video/7199597380491136262

Disclosure statement

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

Notes on contributor

Dominic Lees is Associate Professor in Filmmaking at the University of Reading. His research focuses on synthetic media, deepfakes and their impact on the screen industries, with published outputs in *Convergence*, the online journal *The Conversation* and the BFI magazine, *Sight and Sound.* He is convenor of the Synthetic Media Research Network. Dominic also writes on film and television aesthetics and practices: he co-authored the book *Seeing It On Television* (Bloomsbury, 2021) and has published in journals including *Critical Studies in Television, The Journal of Media Practice* and *Media Practice and Education.* His earlier career was in television and film production, working as a director in current affairs, TV drama, and as co-writer/director of the feature film, *Outlanders* (2008).

ORCID

Dominic Lees http://orcid.org/0000-0002-3787-3740

References

Ajder, Henry, and Joshua Glick. 2021. *Just Joking: Deepfakes, Satire and the Politics of Synthetic Media*. MIT Open Documentary Lab. https://cocreationstudio.mit.edu/just-joking/.

Ajder, Henry, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen. 2019. *The State of Deepfakes: Landscape, Threats, and Impact*. Amsterdam: Deeptrace. https://regmedia.co.uk/2019/10/08/deepfake_report.pdf.

The Alternative Christmas Message. 2020. Directed by William Bartlett. Channel Four Television. https://www.channel4.com/programmes/alternative-christmas-message.

Aneja, Shivangi, Cise Midoglu, Duc-Tien Dang-Nguyen, Michael Alexander Riegler, Pål Halvorsen, Matthias Nießner, Balu Adsumilli, and Chris Bregler. 2021. "MMSys'21 Grand Challenge on Detecting Cheapfakes." In *12th ACM Multimedia Systems Conference (MMSys'21)*, September 28–October 1, 2021.

Arnold, Richard. 2022. "Performers' Rights and Artificial Intelligence." In *Research Handbook on Intellectual Property and Artificial Intelligence*, edited by Ryan Abbott and David Geffen, 218–224. Cheltenham: Edward Elgar Publishing.

Ashok, M., R. Madan, A. Joha, and U. Sivarajah. 2022. "Ethical Framework for Artificial Intelligence and Digital Technologies." *International Journal of Information Management* 62. https://doi.org/10.1016/j.ijinfomgt.2021.102433.

The Bletchley Declaration By Countries Attending the AI Safety Summit. 1–2 November 2023. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

The Book of Boba Fett. 2022. US: Disney.

Captain Scarlet and the Mysterons. 1967–1968. Created by Gerry Anderson and Sylvia Anderson. UK: ITC Entertainment.

Chernobyl. 2019. US/UK: HBO/Sky.

Cole, Samantha. 2017. "AI-assisted Fake Porn is here and We're All Fucked." *Vice*. https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn.

Danry, V., J. Leong, P. Pataranutaporn, P. Tandon, Y. Liu, R. Shilkrot, P. Punpongsanon, T. Weissman, P. Maes, and M. Sra. 2022. "AI-Generated Characters: Putting Deepfakes to Good Use." In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (CHI '22 Extended Abstracts), April 29–May 5, 2022, New Orleans, LA, USA.

Deep Fake Neighbour Wars. 2023. Directed by Spencer Jones. UK:ITVX.

Deepfake porn: you could be next. 2022. Directed by Jess Davies. BBC 3.

De Ruiter, Adrienne. 2021. "The Distinct Wrong of Deepfakes." *Philosophy & Technology* 34 (4): 1311–1332. https://doi.org/10.1007/s13347-021-00459-2.

Equity. 2022. *Stop AI Stealing the Show*. https://www.equity.org.uk/media/6134/report-stop-ai-stealing-the-show.pdf.

Flick, C., and K. Worrall. 2022. "The Ethics of Creative AI." In *The Language of Creative AI*, edited by C. Vear and F. Poltronieri. Cham: Springer. https://doi.org/10.1007/978-3-031-10960-7_5.

Floridi, L. 2018. "Soft Ethics, the Governance of the Digital and the General Data Protection Regulation." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2133): 20180081. https://doi.org/10.1098/rsta.2018.0081.

Gaur, L., ed. 2022. *DeepFakes: Creation, Detection, and Impact*. 1st ed. Boca Raton: CRC Press.

Gerry Anderson: A Life Uncharted. 2022. Directed by Benjamin Field. UK: BritBox.

Here. 2024. Directed by Robert Zemeckis. US: Sony Pictures.

Hight, Craig. 2022. "Deepfakes and Documentary Practice in an age of Misinformation." *Continuum* 36 (3): 393–410. https://doi.org/10.1080/10304312.2021.2003756.

In Event of Moon Disaster. 2019. Directed by Francesca Panetta and Halsey Burgund. US: MIT. https://moondisaster.org.

Jackson, Peter, dir. 2012. *The Hobbit: An Unexpected Journey*. US: New Line Cinema and MGM.

Kapur, Anandana, and Nagma Sahi Ansari. 2022. "Coding Reality: Implications of AI for Documentary Media." *Studies in Documentary Film* 16 (2): 174–185. https://doi.org/10.1080/17503280.2022.2048235.

Kim, Jihoon. 2022. "Expanded Images." In *Documentary's Expanded Fields: New Media and the Twenty-first-century Documentary*, edited by Jihoon Kim, 24–63. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780197603819.003.0002.

Lees, Dominic, Thomas Bashford-Rogers, and Marcus Keppel-Palmer. 2021. "The Digital Resurrection of Margaret Thatcher: Creative, Technological and Legal Dilemmas in the Use of Deepfakes in Screen drama." *Convergence: The International Journal of Research Into New Media Technologies* 27 (4): 954–973. https://doi.org/10.1177/13548565211030452.

Lyons, James. 2020. *Documentary, Performance and Risk*. Abingdon: Routledge.

Lyu, Siwei. 2022. "DeepFake Detection." In *Multimedia Forensics*, edited by Husrev Taha Sencar, Luisa Verdilova, and Nasir Memon, 313–333. Online: Springer. https://doi.org/10.1007/978-981-16-7621-5.

Marquis, Elizabeth. 2013. "Conceptualizing Documentary Performance." *Studies in Documentary Film* 7:45–60. https://doi.org/10.1386/sdf.7.1.45_1.

Meikle, Graham. 2022. *Deepfakes*. Cambridge: Polity Press.

Mihailova, Mihaela. 2021. "To Dally with Dalí: Deepfake (Inter)Faces in the Art Museum." *Convergence: The International Journal of Research into New Media Technologies* 27 (4): 882–898. https://doi.org/10.1177/13548565211029401.

MIT Open Documentary Lab. 2020. *Deepfakery: A Livestream Talk Series*. http://opendoclab.mit.edu/?s=deepfakery.

Morley, J., L. Floridi, L. Kinsey, and A. Elhalal. 2020. "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices." *Science and Engineering Ethics* 26 (4): 2141–2168. https://doi.org/10.1007/s11948-019-00165-5.

Mustak, Mekhail, Joni Salminen, Matti Mäntymäki, Arafat Rahman, and Yogesh K. Dwivedi. 2023. "Deepfakes: Deceptions, Mitigations, and Opportunities." *Journal of Business Research* 154:113368. https://doi.org/10.1016/j.jbusres.2022.113368.

Naruniec, J., L. Helminger, C. Schroers, and R. M. Weber. 2020. "High-resolution Neural Face Swapping for Visual Effects." *Computer Graphics Forum* 39 (4): 173–184. https://doi.org/10.1111/cgf.14062.

Nuremberg. 2000. US: TNT.

Navarro, Vinicius. 2020. "Time, Digital Environments, and the Documentary Experience." *JCMS: Journal of Cinema and Media Studies* 60 (1): 92–110. https://doi.org/10.1353/cj.2020.0064.

Paget, Derek. 2011. *No Other Way to Tell it: Docudrama on Film and Television*. 2nd ed. Manchester: Manchester University Press.

Paris, Britt, and Joan Donovan. 2019. *Deepfakes and Cheap Fakes: The Manipulation of Audio and Video Evidence*. https://www.archives.gov/files/presidential-libraries/events/centennials/nixon/images/exhibit/rn100-6-1-2.pdf.

Partnership on AI. 2023. *PAI's Responsible Practices for Synthetic Media: A Framework for Collective Action*. Online: PAI. https://partnershiponai.org/wp-content/uploads/2023/02/PAI_synthetic_media_framework.pdf.

Pavis, M. 2021. "Rebalancing our Regulatory Response to Deepfakes with Performers' Rights." *Convergence: The International Journal of Research into New Media Technologies* 27 (4): 974–998. https://doi.org/10.1177/13548565211033418.

Peele, Jordan. 2018. "A Public Service Announcement from Monkeypaw Productions and BuzzFeed." YouTube. https://www.youtube.com/watch?v=cQ54GDm1eL0.

Pietrobon, Agnese. 2020. 'Are We not Going to be Able to Ever Know What the Truth is, Again?' – Halsey Burgund (in the Event of Moon Disaster)." Online: *XRMust*. https://www.xrmust.com/xrmagazine/halsey-burgund-event-moon-disaster/.

Purse, Lisa. 2018. "Layered Encounters: Mainstream Cinema and the Disaggregate Digital composite." *Film-Philosophy* XXII (2): 148–167. https://doi.org/10.3366/film.2018.0070

Rana, M. S., M. N. Nobi, B. Murali, and A. H. Sung. 2022. "Deepfake Detection: A Systematic Literature Review." *IEEE Access* 10:25494–25513. https://doi.org/10.1109/ACCESS.2022.3154404.

Renov, Michael. 2004. *The Subject of Documentary*. Minneapolis: University of Minnesota Press.

Rini, Regina. 2020. "Deepfakes and the Epistemic Backstop." *Philosophers' Imprint* 20 (24): 1–16. http://hdl.handle.net/2027/spo.3521354.0020.024.

Rothkopf, Joshua. 2020. "Deepfake Technology Enters the Documentary World." *The New York Times*, July 6. https://www.proquest.com/newspapers/deepfake-technology-enters-documentary-world/docview/2420903688/se-2.

Safire, Bill. 1969. *Memo to H.R. Haldeman In Event of Moon Disaster*. US: Presidential Libraries. https://www.archives.gov/files/presidential-libraries/events/centennials/nixon/images/exhibit/rn100-6-1-2.pdf.

Schwartz, Drew. 2019. "Craig Mazin's Year-Long Obsession with Making 'Chernobyl' Terrifyingly Accurate." *Vice*, June 3. https://www.vice.com/en/article/j5wbq4/craig-mazin-interview-about-chernobyl-hbo-miniseries-on-how-accurate-and-what-really-happened.

Seow, J. W., M. K. Lim, R. C. W. Phan, and J. K. Liu. 2022. "A Comprehensive Overview of Deepfake: Generation, Detection, Datasets, and Opportunities." *Neurocomputing* 513:351–371. https://doi.org/10.1016/j.neucom.2022.09.135.

Stahl, B. C., J. Timmermans, and C. Flick. 2016. "Ethics of Emerging Information and Communication Technologies – On the Implementation of Responsible Research and Innovation." *Science and Public Policy* 44 (3): 369–381.

The Capture. 2021–2022. UK: BBC.

Thunderbirds. 1965–1966. Created by Gerry Anderson and Sylvia Anderson. UK: ITC Entertainment.

Tolosana, Ruben, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia. 2020. "Deepfakes and Beyond: A Survey of Face Manipulation and Fake Detection." *Information Fusion* 64:131–148. https://doi.org/10.1016/j.inffus.2020.06.014.

Ume, Chris. 2021. *The Chronicles of Deepfake Tom Cruise (Viral TikTok Videos)*. YouTube https://www.youtube.com/watch?v=nwOywe7xLhs.

Vaccari, Cristian, and Andrew Chadwick. 2020. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media + Society* 6 (1): 2056305120903408. https://doi.org/10.1177/2056305120903408.

Welcome to Chechnya. 2020. Directed by David France. US: Public Square Films/HBO. https://www.welcometochechnya.com/.

Wood, Aylish. 2007. "Pixel Visions: Digital Intermediates and Micromanipulations of the image." *Film Criticism* XXXII:72–94.

Yu, Peipeng, Zhihua Zha, Jianwei Fei, and Yujiang Lu. 2021. "A Survey on Deepfake Video Detection." *IET Biometrics* 10 (6): 607–624. https://doi.org/10.1049/bme2.12031.