

# Configuring Fakes: Digitized Bodies, the Politics of Evidence, and Agency

Britt Paris 

## Abstract

This comparative case study analysis used more than 200 examples of audiovisual manipulation collected from 2016 to 2021 to understand manipulated audiovisual and visual content produced by artificial intelligence, machine learning, and unsophisticated methods. This article includes a chart that categorizes the methods used to produce and disseminate audiovisual content featuring false personation as well as the harms that result. The article and the findings therein answer questions surrounding the broad issues of politics of evidence and harm related to audiovisual manipulation, harassment, privacy, and silencing to offer suggestions towards reconfiguring the public's agency over technical systems and envisioning ways forward that meaningfully promote justice.

## Keywords

audiovisual content, cheap fakes, configuration, deepfakes, politics of evidence

## Introduction

For some time, computer graphics systems fed by data systems have been able to parse and generate images, even realistic images, from an existing video. Prior to 2017, this was the province of the major motion picture studios, which used the technology to entertain mass audiences, and computer science research laboratories, which used it to develop computer vision (Bansal et al., 2018; Diakopoulos & Johnson, 2021; Paris & Donovan, 2019; Suwajanakorn et al., 2017; Vaccari & Chadwick, 2020). In 2017, consumer-grade and sometimes free, image manipulation software using machine learning gained public attention as pornographic videos with the faces of famous women grafted onto pornographic actors' bodies appeared on Reddit (Cole, 2017; Maddocks, 2020). Since then, an application (app) for creating deep nudes of anyone's picture was developed, made widely accessible, and almost immediately shuttered when the developers suddenly understood the harm that could result. However, just 1 year later, manipulated nude images, many featuring teenaged girls, generated by the app mysteriously appeared on encrypted Telegram conversations (Adjer et al., 2020; Harwell, 2020). Hologram technology for generating performances by deceased artists for mass audiences has employed machine learning. Kanye West's 2020 gift to his wife, Kim Kardashian, a hologram of her late father (Kneese, 2020), was the first time the wider public became aware of the use of this technology to produce a personalized performance.

The abovementioned examples illustrate the personal nature of these false performances for the viewers and the viewed. They also highlight the increasing prevalence of these audiovisual objects, generated for expression, play, or scientific experimentation and the need to consider the attendant ethical and policy questions. For anyone with an online persona, images posted to a public social media account are vulnerable to impersonation attacks. Women, people of color, and lesbian, gay, bisexual, and transgender (LGBT) individuals are at higher risk for attacks (Citron, 2016; Franks, 2017; McGlynn et al., 2017). This article discusses the state of audiovisual manipulation through artificial intelligence (AI) and unsophisticated techniques to identify the methods used to produce false audiovisual content. The consequences, particularly the harm, of these productions are also discussed answers questions about the broad issues surrounding the harm of algorithmic technology used to push false, misleading, and misrepresented data as fact. It also offers suggestions for mitigating this harm through political rather than technical solutions.

---

Rutgers University–New Brunswick, USA

**Corresponding Author:**

Britt Paris, School of Communication & Information, Department of Library and Information Science, Rutgers, the State University of New Jersey, 4 Huntington St. New Brunswick, NJ 08901, USA.  
 Email: britt.paris@rutgers.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Social Media + Society  
 October–December 2021: 1–13  
 © The Author(s) 2021  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](http://sagepub.com/journals-permissions)  
 DOI: 10.1177/20563051211062919  
[journals.sagepub.com/home/sms](http://journals.sagepub.com/home/sms)



Studies on algorithmic harm have shown that technologies that rely on mass data collection and classification to make decisions about the ordering of the world have reified and reinforced the existing social structures (Benjamin, 2019a, 2019b; S. U. Noble, 2018). Thus, the present discussion includes the epistemic and policy concerns around information disorder (Bennett & Livingston, 2018; Wardle & Derakhshan, 2017) and post-truth epistemic crises (Bimber & Gil de Zúñiga, 2020; Waisbord, 2018). They are examined in terms of the social construction of truth through the technologies of evidence, expertise, and structural power that are situated within the context of audiovisual manipulation.

Regarding these epistemic concerns, these images, which have become prevalent, are expressed and accepted as truth. They configure how individuals interpret and live in the world that which, in turn, shapes social structures and sociotechnical programs for the upkeep of those structures (Barad, 2003; Haraway, 1991; Suchman, 2000). This concept of configuration from feminist STS contributes to the discussion of how audiovisual material takes on the character of evidence as it proliferates across social media. Second, configuration helps us understand the misdirected hype about the possible negative effects of deepfakes on electoral politics and the insufficient attention to the consequences for individuals who belong to minoritized groups. These technologies have been developed to determine the spectrum of possible individual action in the interest of maintaining the social order (Paris & Donovan, 2019; Vaccari & Chadwick, 2020). In turn, the status quo is promoted in a never-ending mutually-constitutive handshake. However, feminist STS has understood the configuration of society and technology as ever-changing. This mutability is a juncture for human agency in the reconfiguration and reconditioning of human possibility, the enactment of a possibly more just social order, and the development and proliferation of new technologies to promote alternative social arrangements that might include justice (Barad, 2003).

To understand how different manifestations of audiovisual manipulation of false personation both strengthen and are shaped by structural power, the study drew from a wide variety of sources to situate instances of audiovisual impersonation in their sociotechnical context. More than 200 examples of audiovisual manipulation that disproportionately affects women, people of color, and lesbian, gay, bisexual, transgender, and queer and/or questioning, intersex, and asexual and/or ally (LGBTQIA) individuals from 2016 to 2021 were collected and annotated with their social, political, evidentiary, forensic, and aesthetic dimensions. This article presents examples as a comparative case study using a previously developed audiovisual manipulation spectrum (Paris & Donovan, 2019) that considers technical methods and production modes. This article argues for understanding these examples in their respective contexts to identify the ethical and political implications as well as possible solutions.

## Literature Review

Even before the advent of deepfakes, social media, false personation, and subsequent harassment, manipulation, and misinformation were rampant. Those who lacked power were affected. The advent of amateur image manipulation software, smartphone cameras, and social media has given anyone with an email address the ability to generate and disseminate content. Thus, an increasingly volatile information sphere has emerged (Feigenson, 2011; Paris & Donovan, 2019). The present study was guided by media and information studies theories and methods. The discourse around video and image manipulation has fomented panic about the use of technology in national and state politics, for example, elections (Diakopoulos & Johnson, 2021). The present study suggests that the more common risks have disproportionately affected women, people of color, LGBTQIA individuals, and those fighting established power (Maddocks, 2020; Paris & Donovan, 2019).

### *The Politics of Audiovisual Evidence*

Evidence, like truth, is a way of discretizing and codifying lived experience into “objective” phenomena, or indicators of truth, that people can agree on to serve as a basis for action (Daston & Galison, 2010; Golan, 2007). The selection of the truths that become evidence is the result of social processes. Historian of Science Tal Golan (2007) offered the frame of the introduction of photographic equipment in the courtroom. The courts were careful to explain photography and X-ray technologies to juries. In some cases, they were forced to offer demonstrations to prove that the photographs were admissible evidence. But in practice, this process of determining what is evidence and how it can be demonstrated is rife with problems that fall along the lines of structural power (Fiske, 1996). A poignant example is the recorded video of documenting members of the Los Angeles Police Department (LAPD) beating Rodney King in 1991 that the LAPD defense slowed down for the trial demonstration. The use of this slowed-down footage and the transfer of the trial to a predominantly White district with a predominantly White jury “made all the difference,” said one juror who, along with the 11 others, acquitted the police officers caught beating King on camera. Courts require evidence to be vetted and displayed for the “right to a fair trial”; however, evidence is often biased and tampered with. Furthermore, the process of constructing evidence as well as its interpretation depends on how those doing the work of constructing or interpreting evidence are situated in structures of power.

While “evidence” is often used as the basis for constructing legal arguments, in archival studies, “evidence” is the way that documents, artifacts, and data within their contexts provide insights into the processes, activities, and events that led to their creation for certain purposes (Gilliland-Swetland, 2000, p. 10). Understanding evidence from this perspective,

archival theorist Tonia Sutherland (2017) and visual culture theorist La Charles Ward (2018) argue that visual evidence around police murder of Black people was once and continues to be used as a form of memorialization and refusal of White supremacist structures undergirding this violence, but at the same time, the digital sphere's drive toward commodification reinscribes gratuitous White supremacist renderings of Black suffering as evidentiary, as a "natural" or "immutable" fact through its repetition, as it is continually served up in search results and spread via recommendation algorithms (Sutherland, 2017). Similarly, Safiya Noble (2013) has shown that sexualized Google search returns for "Black girls" proliferate and are commonly mistaken as evidence of how Black girls are, instead of being understood as evidence of what makes money for Google, as it serves the dominant white supremacist, patriarchal culture's tastes, opinions, and attitudes to users.

This subsection has shown how the politics of evidence, a concept in STS that explains how evidence is shaped by the powerful to suit their interests of cultivating and strengthening their power, often at the expense of those who are most vulnerable in structures of systemic inequality, works in present and historical cases of audiovisual evidence (Daston & Galison, 2010; Golan, 2007; Murphy, 2006). Understanding the politics of evidence around audiovisual material suggests we ask: who gets to make decisions about what is and what is not evidence, whose expertise is valued and devalued in this decision-making process, what are acceptable and unacceptable forms of expression and evidence, and who is harmed by expression and/or evidence and who benefits? (Paris & Donovan, 2019). The politics of evidence relates closely to the next section on expertise that highlights whose expertise is valued and devalued in the discourse around emergent technology.

### *The Politics of Expertise and Amateurism*

The public confers authority to experts because of the experts' training and proximity to evidence and information. Experts must often establish relationships with powerful entities during training, and this relationship to powerful entities often facilitates access to evidence and information (Marvin, 1990). At present, Paris and Donovan (2019, 22) stated, "When the expert interpreter works in service of the powerful, existing social structures are reified, regardless of what 'truth' is captured through or by any given media technology." However, amateurism, like expertise, is related to structural power. It is often evoked by technology companies to outsource research and development for free. Amateur tinkering of all types, but specifically tinkering with technical apparatuses, has historically provided an outlet for people, typically White middle- to upper-class men, who feel disaffected by changing social norms and material economic realities that they endeavor to master technology as a way to control their world (Douglas, 1987; Irwin & Wynne, 1996;

Marvin, 1990). However, it is crucial not to uncritically accept narratives that position online amateurs as necessarily politically revolutionary or just.

With the increasing popularity of social media platforms and smartphone cameras, investments in these technologies soared in the mid-2000s as did amateur content production. The combination of social media and smartphone cameras allowed more people to create and distribute audiovisual content (Jenkins, 2004). Audiovisual media, with the totality of their evidentiary, expressive, and interpretative power, have never before achieved the speed and scale that are now possible. As Ward and Sutherland note, this enables those from minoritized groups to use social media to promote oppositional readings (Hall, 1973/2001) of events by sharing evidence and calling for justice. Jenkins (2004) also makes this connection arguing that amateur expression online could be a revolutionary political force. However, set within the legal system and related structures of inequality, there has been little substantive change based on this audiovisual evidence wielded by nonexperts (Currie et al., 2016; Paris & Pierre, 2017; Wood, 2017). Moreover, the market-based incentives and libertarian ideals that undergird online expression privilege amateurs who espouse ideas that reify the status quo.

### *Digitized Bodies and Harm*

The proliferation of pornographic expression online has been linked to existing social power relations (Attwood, 2007; Coopersmith, 1998; McGlynn et al., 2017; S. Noble, 2013). The internet provides access to pornographic content made with and without consent. Unmoderated and unregulated sites "normalize misogyny, hurtful sexuality, racism, and even seeking revenge on female ex-partners" (McGlynn et al., 2017, p. 33).

Patricia Hill Collins' "matrix of oppression" (Collins, 1990) can help explain this phenomenon. Collins used this conceptual framework to discuss how people are oppressed differently by hegemonic discourse based on the minoritized identities they inhabit. The matrix of oppression explains how and why existing social structures or existing structural power relations are configured provide and foreclose opportunities for different groups of people, for example, to privilege White upper-class men over poor White women who are, in turn, privileged over poor Black women. Crenshaw's (1989) notion of "intersectionality" charts this same process of allocating privilege and oppression through law and policy. In these schema, privilege and oppression are not binaries but should be understood in a matrix or comparative system that contextualizes a person's existence in processes of social stratification.

Instances of pornographic and objectifying content are not simply disembodied doubles that have no effect on the people who are recreated and faked. These avatars stand-in for the people they represent. In virtual reality, they are

embodied; thus, users can feel in-game or in-system physical harassment (Franks, 2019). More broadly, these false personations can cause psychological trauma and real-life, material difficulties. The audience's association of a false image with someone can cause harm, such as lost job opportunities, employment termination, and child custody revocation (Citron, 2016).

Feminist legal scholars Claire McGlynn et al. (2017, p. 33) referred to the nonconsensual superimposition of an individual's face or other identifying feature onto pornographic depictions of sexual acts as "sexualized photoshopping," which is often wielded in ways that are tantamount to gendered, racist, and sexualized "image-based sexual abuse." McGlynn et al. (2017) noted that images that are manipulated and miscontextualized without the featured individual's consent are deemed permissible forms of expression. The originals are public; however, at the moment of manipulation, they strip the individual of their rights to the depiction of their sexuality. Image-based sexual violence has been used for many purposes, including harassment, suppression of the press, civil society, and political opposition. For decades around the world, audiovisual content has been manufactured and manipulated to silence opposition (Bevins, 2020; Coopersmith, 1998). Today faked sex videos are disseminated on social media with dire, and sometimes violent, consequences. Primary targets have been and continue to be LGBT politicians, women politicians or activists, especially women of color, and those questioning power (Citron, 2016). Sophisticated technology is not required to create these fakes. A semi-convincing double to depict the target in a sexualized or pornographic image or video is sufficient to drive discourse, and sometimes, public opinion.

While major platforms have moderated pornographic material so that it no longer looms as the threat it once was, they have claimed that they cannot moderate all problematic content at this speed and scale of social media and are loath to appropriately support human moderators to do this work (Caplan, 2018; Roberts, 2019). They are ill-equipped and uninterested in managing audiovisual manipulation performed with inexpensive editing software that allows the cutting, speeding up, slowing down, or simply recontextualizing images (Dan et al., 2021). These are the same techniques advertisers and content creators use to generate content that drives engagement with the platform.

### *Configuring Agency in Sociotechnical Systems*

Social media and the ecology of information and communication technologies have opened up socialization opportunities while introducing countervailing constraints. As with older technologies, individuals use images and videos to "identify" with another as they construct themselves and their personal identities (Floridi, 2013; S. U. Noble, 2018). This is similar to the concept of configuration (Barad, 2003) in STS, which suggests that individuation is a social process that includes

inanimate objects, such as technology. Moreover, this process of configuration or individuation is shaped by social structures and extant technologies. Political economists (Braverman, 1974/1998; Harvey, 2003; D. F. Noble, 1984/2013) and critical race and technology scholars (Benjamin, 2019b; Brock, 2018; S. U. Noble, 2016, 2018) have argued that because the technologies of individuation are built to reify the status quo, individuals are shaped to see efficiency, market-driven ideals, and the misogynist and White supremacist structures that permeate technology as normal, natural, and inevitable. The public has been tricked into believing that there is no way out. The historical perspectives in feminist STS, along with critical race theorists of technology, suggest that technology individuation can be rerouted and reconfigured (Barad, 2003; Benjamin, 2019b). Ruptures, in the form of the many problems we see with our current information and communication landscape, can be useful in shaking uncritical adherence to these individuation engines to help individuals to realize that they are in control. At the core, technology is nothing without people; thus, it can be changed.

This section has contextualized the issues surrounding audiovisual manipulation, the digitization of bodies. The scale of social media has amplified the volume and velocity of visual falsehoods that users encounter through automated information and communication technologies to the point that these manipulated images shape the political economy of our sociotechnical systems. Deepfakes and other AI-generated information objects have provided new means of generating copies with no originals (Paris & Donovan, 2019). However, a more serious problem is the countless new channels for distributing not these technically impressive AI fakes that are easy to detect, but fakes that have long proliferated across social media that are created with unsophisticated means that occurs at the levels of the personal and the public; the political and the banal. It can be seen in the endless fan remixes of popular television, films, and music videos as well as the public and private figures silenced by faked sex videos. An understanding of these examples modes of sociotechnical configuration indicates that the current problems do not emanate from the existence of the new technologies but, rather, how these technologies are built and deployed by a few, then used and interpreted by the wider public.

### **Research Question, Method, and Analysis**

This article centers on the primary research question and two subquestions related to the production and dissemination of false audiovisual personation online. The primary research question speaks to the ultimate goal of this work: to engage in an ongoing discussion of how structural power is embedded in and results from the uncritical use of technologies in which identity is the product and the engine. This illustrates the relationship between these false performances and questions of identity and agency over expression and evidence. The

subquestions address structural power in technical systems. The significance of these questions is their contribution to the growing expertise in critical informatics, critical technology policy, and technical practice. The overall goal is to produce research that encourages active reflection on technology and society as institutional, discursive, and practice-based formations that reify and extend structural power in harmful ways, that once identified, can be dismantled, reconceptualized, and reoriented to more egalitarian ends as identified in Question 1. This article attempts to reach this goal by addressing the sub-questions below:

1. How does contemporary audiovisual manipulation for false personation coexist with structural power?
  - (a) What is the role of expertise in the production of manipulated content?
  - (b) What is the harm or benefit of such content, and to whom does this accrue?

The abovementioned theoretical perspectives undergirded this comparative case study of the role of false personation in reifying structural power. A comparative case study is well suited to the research questions at hand because it allows the analyses of patterns across multiple cases. Comparative case studies employ both qualitative and quantitative methods in iterative phases of data gathering, case selection, and comparative analysis among cases. This method is particularly useful for understanding how the context of any particular case influences patterns that are found (Goodrick et al., 2020).

This work began with collecting 200 video and image examples tracked from 2016 to 2021 through Google Alerts that returned news stories and examples of these types of videos on social media and pornography video-hosting platforms. Acker and Donovan's (2019) data craft facilitated the selection of examples of well-known forms of audiovisual manipulation and the assessment of their production and dissemination. Critical technical discourse analysis (Brock, 2018) was used to analyze the social, political, evidentiary, forensic, and aesthetic dimensions of the video examples and to place them on an audiovisual manipulation spectrum developed by Paris and Donovan (2019) to better understand the ethical and political implications. The work of feminist legal scholars Clare McGlynn et al. (2017), Mary Ann Franks (2017, 2018, 2019), and Citron and Franks (2014) informed the differentiation between harmful and harmless uses of audiovisual manipulation not by their intent but by their effects, specifically on those who have not traditionally been included in technological development. Case selection and analysis occurred in three overlapping steps, which are discussed subsequently.

### ***Step 1: Collecting and Annotating Content***

Google Alerts notifications were set up in 2016 before the onset of the deepfake phenomenon. The first three categories were “fake video,” “audiovisual manipulation,” and “fake images.” After the emergence of the term “deepfake” in 2017,

the search terms “deepfake,” “deepfake porn,” and “fake video” were added. In 2018, “cheapfakes” and shallowfakes” were also included. The alert yielded daily results, each of which was traced through headlines to the original source where possible. Source links to the videos, screenshots to retain the context and metadata where possible, and source video images were saved in AirTables interactive spreadsheet, allowing easy annotation of how videos were discussed in the news. In accordance with Acker and Donovan's (2019) data craft, the metadata related to the creator, production technology, and dissemination of the videos were recorded along with the videos in the AirTables spreadsheet. Where available, the location of the original video usually gave a pretty clear indication of the creator. Metadata on the original time stamps on the videos or images, the content sharer, the timing and number of shares or likes, and the influential nodes the content passed through revealed the modes of dissemination.

### ***Step 2: Understanding Discourse and Harms***

To make sense of the broad corpus of data collected, the author performed a critical discourse analysis of news sources and contextual content around the videos, such as comments, and response videos with relation to questions of harm and expertise highlighted in the sections on the politics of evidence and expertise (Daston & Galison, 2010; Golan, 2007; Murphy, 2006). The critical discourse analysis entailed the following questions in these instances of audiovisual manipulation: who gets to make decisions about what is and what is not evidence, whose expertise is valued and devalued in this decision-making process, what are acceptable and unacceptable forms of expression and evidence, and who is harmed by expression and/or evidence and who benefits? The online context around the instances and surrounding news sources in the Step 1 annotations were used in the critical discourse analysis (Brock, 2018; Fairclough, 2013; van Dijk, 2005; Wodak & Meyer, 2009), which considered the harm and structural power themes identified by the feminist scholars (Citron, 2016; Franks, 2019; McGlynn et al., 2017; S. Noble, 2013; Sutherland, 2017) cited in the literature review. These analyses of discourse around the videos collected attended to sexualized, racialized, gendered, harassment, hate speech, political speech, misinformation, rumors, and the individuals or processes that were discussed as producing and/or being harmed by the video. Videos coded as not harmful were excluded from the analysis. In Step 2, 126 unique videos and images were analyzed. In all, 94 fit into more than one category of harm. This process of analysis led to the development of ideal types for each category, which further facilitated the analysis and categorization of the video and image content.

### ***Step 3: Selecting Examples for the Comparative Case Study***

The notion of harm and its manifestation in the discourse surrounding the videos guided the selection of examples for

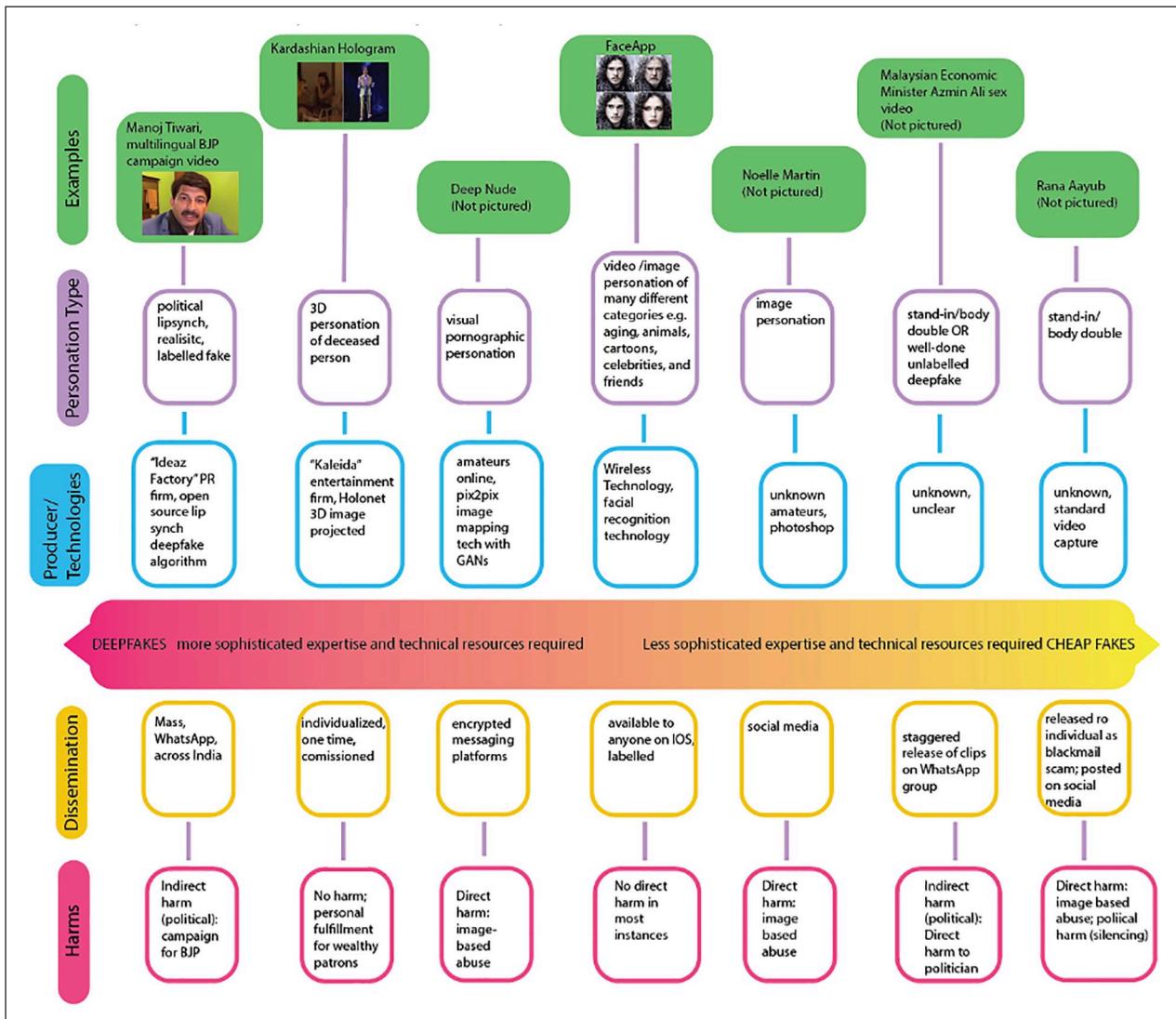


Figure 1. The false personation deepfakes/cheap fakes spectrum.

the comparative case study. The cases discussed subsequently were relatively recent high-profile examples that best illustrate harm. This includes the individuals who were harmed and the mechanism and the manifestation of the harm. A previously published table by Paris and Donovan (2019) was adapted to reflect the analytical dimensions of interest in the present study (Figure 1). It illustrates the harm caused by amateurs leveraging the audiovisual or visual likeness of a person for reasons such as play, expression, revenge, and harassment.

The three steps overlapped temporally. Step 1 lasted from 2016 to 2021. However, collecting examples intensified after the first instance of deepfake pornography in 2017. In many cases, Steps 1 and 2 were concurrent. The harm categories in Step 2 were refined in 2018 during the generation of the first research output (Paris & Donovan, 2019). Step 3, the selection of examples for the case, occurred in late 2020 and early 2021 as the article took shape.

## Findings: Sociotechnical Configuration

The data collected led to the creation of the spectrum subsequently that organizes information into a typology of audio-visual manipulation examples that illustrate the differences between deepfakes and cheap fakes regarding technical sophistication, techniques, dissemination modes, and harm. A review of the spectrum shows that as the technical resources and expertise required for the production of fakes decreases (toward the right of the spectrum), the ability for amateurs to produce them increases. Deepfakes, which require extensive technical resources and sophisticated machine learning techniques, are toward the left of the spectrum. These are the least common and most computationally difficult to accomplish. They have caused the least harm; however, they are not completely harmless as is exemplified in the case of Manoj Tiwari's Bhartiya Janata Party (BJP) deepfake to target his message to voters. Other forms of

audiovisual manipulation are performed with easy-to-use downloadable software, some of which is free. In the more common examples, simpler techniques, such as mislabeled footage or lookalike stand-ins, were used, and harm was caused in the ways described with the examples from the cases.

As each example collected in the broader corpus included contextual discursive information found in news stories, comments, or captions where available, these discursive instances were analyzed with relation to questions of harm and expertise to suggest what these examples can tell us about the politics of evidence. The following critical discourse analysis around the topics of expertise and harm are meant to shed light onto how the politics of evidence surfaced in the examples.

Over the past decade, computer scientists have experimented with deep learning neural networks to generate image, audio, and audiovisual clips that appear realistic (Suwajanakorn et al., 2017). In addition, adversarial networks have been used to distinguish fakes from authentic content (Bansal et al., 2018). This technology is available to artists, advertisers, and hobbyists who congregate online to reproduce these resource-intensive machine learning techniques with open source, free, and consumer-grade technology. From the rows in the spectrum above labeled “producer/technologies” in blue, and “dissemination” in yellow, we see that the most technically sophisticated and realistic examples are those produced in academia, entertainment, and PR firms, and are released widely with a label or disclaimer, while the examples requiring less technical expertise and computational resources were released anonymously to specific individuals, small groups or over encrypted platforms.

Prior to 2017, the most common, banal forms of this technology exist in the hands of mobile users across the world as users have enjoyed the viral appeal of apps like Face App, that are easily downloaded onto an individual’s phone, and with the app turned on and video function enabled, individuals can, in real time, age themselves, make themselves smile, all the while that application stores user video data at the app creator’s whims and, generally, their benefit (Tiffany, 2019). Artists, including those employed by advertisers, and PR firms have produced public service announcements about the dangers of deepfakes (Paris & Donovan, 2019). However, these artists employed at corporate firms are no longer confined to raising awareness about the implications of deepfakes. In the example of Manoj Tiwari, a candidate for the BJP, expert technicians at PR firm Ideaz Factory produced a set of deepfake videos in different languages to quickly and cheaply getting out Tiwari’s message encouraging voters not to vote for his rival to targeted language demographics. The BJP, India’s right-wing party known for its adherence to Hindu nationalism, noted that the Tiwari videos reached 15 million people in 5,800 WhatsApp groups (Christopher, 2020). Kanye West’s gift of a holographic revivification of Kim Kardashian’s father is a sophisticated technology has

also been used in the entertainment industry (Kneese, 2020). Holographic revivification is costly and generally meant for a wide audience; however, when used to revivify loved ones, that is, in the Kardashian instance, it is meant to revive a personal connection for specific audiences with exorbitant amounts of money.

The harms that result from these are seemingly innocuous in intent, and largely symbolic, if not slight norm-pushing behavior around socially accepted practices. FaceApp and SnapChat filters represent another way to pull users in and collect their facial and video data by the technology companies that own these properties. In the Kardashian example, we see unfettered conspicuous consumption of a fleeting, intensely personal luxury. The Tiwari example is seemingly innocuous in its intent—to produce political ads cheaply and quickly in multiple languages to target different demographics, its brazenness in the face of accepted Indian norms for producing and targeting political advertising suggests the party’s drive to win coupled with an unwillingness to play by the socially-accepted rules.

The discursive context around these examples suggests that machine learning applications are understood as creative learning technologies that encourage sharing and expressive play or for cutting costs and making money; there is no guarantee of responsible behavior. That the first notable amateur output of the new technology was the nonconsensual depiction of the face of a famous woman was grafted onto the original female actor engaged in a sex act (Cole, 2017, n.p.; Paris & Donovan, 2019),<sup>1</sup> while the original actor was erased in the video, and could not gain compensation for redistribution, foreshadowed the subsequent problems. AI generated videos of celebrities presaged the faked videos and images of regular people. By 2019, DeepNude had allowed people to insert an image or video of anyone, such as a classmate, coworker, or acquaintance, to undress the person. The creators of the app did not discern any problems during its development. Just days after its deployment, they suddenly saw the error of their ways and shuttered their app but not before someone had copied the open source GitHub code. Amateurs are still able to use DeepNude’s GitHub processing to “pornify” anyone’s image to distribute via encrypted platforms, such as Telegram (Adjer et al., 2020). Deepfake pornography is on the continuum of image-based sexual abuse because, currently, social media and face swapping rely on the consent that is barely granted. To participate, users must agree to the vaguely worded terms of service. This soft consent justifies data collection that goes beyond user traffic and ad clicks. It involves the capture of image, audio, and users’ face and body data (McGlynn et al., 2017, p. 33). Anyone with images of themselves online risks their likeness being faked. Women, LGBTQ individuals, and people of color are at higher risk.

A plethora of technologically unsophisticated examples precede the deepfake phenomenon and continue unchecked illustrate who is at higher risk of being faked. In 2016,

Australian teenager Noelle Martin discovered that her face had been superimposed on pornographic images with hateful captions and shared online by her classmates than by scores of amateur harassers online (Sturmer & Barney, 2016). These hateful pornographic images constituted the first few pages of search results for Noelle Martin who was graduating that year. On top of the trauma, Martin experienced from this image-based sexual abuse directed toward her a minor at the time; this online harassment pushing these pornographic results to the top of search results for her name caused her to worry about her future, particularly her job prospects.

As is the case with Noelle Martin and countless others (Maddocks, 2020), this practice has become even more pervasive because technology has made it easy to edit original images in ways that are often impossible to detect. But images need not be heavily edited to cause harm. Also in 2016, supporters of the BJP used a body double to produce a sex video of Rana Ayyub, a female Indian journalist who criticized the Indian government, ruled by the BJP at the time, for jailing accused sex offenders without trials (Ayyub, 2016, 2018). Blackmailers and harassers sent her the video over social media, and via mail to extort her for this ostensible sex tape, and to intimidate her into silence. This tactic was also used against Malaysian Economic Minister Azmin Ali. Images purportedly of him involved in gay sex were disseminated not to him but to other government officials, in timed, staggered releases to their WhatsApp groups in a country where homosexuality is illegal (Walden, 2019). It was debated in the press whether this was an instance of a deepfake, and in recent months, it came out that it was a body double.

The examples demonstrate how expertise over audiovisual manipulation for false personation is wielded to sustain and shape practices that benefit the existing structure of society. These are examples of how these videos are used to harm women, specifically women of color, LGBTQIA individuals, and those questioning power, and alternately to benefit those who are already powerful, like right-wing political parties, entertainment companies, PR firms, social media companies, and wealthy individuals. Analyzing the contextual elements of these videos around harm and expertise demonstrates how the politics of evidence—who gets to define the terms of acceptable use, official policy, whose creative play is painted as harmless and who that affects, and who has the time, know-how, and resources to refute or redress harms—are shaped by the already powerful and negatively impacts those individuals who silenced, coerced, and made most vulnerable in systems of structural inequality.

## Discussion: Configurations of Evidence, Expertise, and Harm

The Martin, Ayyub, and Ali examples of image-based sexual abuse and the recontextualization of images to incite violence and bring about authoritarian regimes around the world

suggest these manipulated images and videos are examples of configuration (Barad, 2003): They become so widespread that they are interpreted as evidence or markers of truth to justify action and can lead to image-based harm on a broader continuum. These examples also prompt two crucial considerations of manipulated images and audiovisual content. First, technical expertise is not a prerequisite for creating harmful audiovisual fakes. The consequences range from the harmful depictions of individuals to the undermining of political processes. Second, we see that those who are harmed fall within unique, minoritized positions within the existing social structure.

Nonexpert experiments in machine learning technology that *are not* mobilized to engage in nonconsensual objectification and harassment were present in the findings as were expert instances that were comparatively harmless as seen in the examples of artists at public relations and entertainment companies operating sophisticated technical expertise widely share their process and intent. Conversely, information as to who produced and spread these more unsophisticated attempts to malign these individuals' character is unclear in many of the cases toward the right side of the spectrum mentioned earlier. Based on the aesthetics of the videos and images, news stories, and comments surrounding the videos and how they were distributed, it is possible to surmise the amateur techniques of production.

Currently, the manipulation of images or videos is protected as expression and allowed to proliferate. But many argue that even when pornography and objectifying content is created or consumed, it is not inherently nefarious (Attwood, 2007). In these cases, amateur “expression” and “playing with” image and audiovisual technology is positioned as normal and natural: an extension of previous practices (Douglas, 1987; Irwin & Wynne, 1996; Marvin, 1990). However, in addition to performing unpaid research and development in image-based machine learning and deepfake technology for the benefit of corporate technology owners, amateur experimentation in this arena is more problematically an outlet for individuals who subscribe to the dominant culture or similar worldviews to express their libidinal desires, dissatisfaction with society, and rages despite the infliction of harm on those who are often unable to enact meaningful protections for themselves (Citron, 2016; Franks, 2018, 2019; S. Noble, 2013).

Targets of harmful false personation in the findings exist at varying levels of vulnerability and privilege in existing social structures in their countries. Noelle Martin was a young woman who had not yet graduated secondary school when targeted as are many of the young women faked with DeepNude. Rana Ayyub was an older, established journalist speaking against the party in power. Azmin Ali is relatively powerful as a member of the government, but his majority status in Malaysia is brought into question with the videos that were produced that try to provide evidence of his homosexuality. But not everyone targeted would be harmed in

exactly the same way because they inhabit differently in what Patricia Hill Collins' "matrix of oppression" (Collins, 1990). For example, Noelle Martin's photoshopped images appearing at the top of a prospective employer's Google Search for her name would enact a different and more direct type of harm than that extended to Rana Ayuub who was a seasoned journalist or Ali who is part of the Malaysian government.

### What Can Be Done?

Anyone whose images are online can be faked; however, time and resources are needed to address false impersonation through the legal system. Currently, the First Amendment to the United States Constitution protects expressive content, such as the creation of pornography with novel technology and to post falsehoods and harassment on platforms. In the United States, laws to punish people who make or distribute a malicious deepfake with prison time were introduced by Republican senator Ben Sasse (2018) in December 2018. The law would limit audiovisual fakes that attempt to intervene in democratic processes. However, it is not clear that everyday victims of harassment, nonconsensual use of images of one's likeness, or other kinds of harmful fakery would be protected. Revenge pornography laws have been adopted across the country (Cyber Civil Rights Initiative, 2020), providing a ray of hope for those who fall prey to image-based abuse. However, there are differences in the state laws; the internet is not state based. Even if the court orders an injunction of the content, it cannot do so for all sites. Under Section 230 of Title 47 of the Communications Decency Act of 1996, often referred to as Section 230, platforms decide which videos stay up and which are taken down (Federal Communications Commission, 1996). Thus, pornographic videos hosted on pornographic sites are likely to stay up (Citron & Wittes, 2017).

Outside of the United States, there are legal frameworks more hospitable to redress. The European Union's (EU) General Data Privacy Regulation's Article 17 includes the "right to be forgotten" which allows individuals the right to request that "the data controller erase his or her personal data, cease further dissemination of the data, and potentially have third parties halt processing the data," (European Parliament and Council of the European Union, 2016, p. 44) when the data is no longer relevant, has not been consented to, has been obtained unlawfully, or, must be erased in accordance with laws in EU member countries. South American and Asian countries have instituted policies that have followed suit with the GDPR (Voss & Castets-Renard, 2016) and the U.S. state of California initiating the California Consumer Privacy Act in 2018 (California State Legislature, 2018), all of which entail some notion of the right to be forgotten. Indeed, feminist legal scholars have noted that the right to be forgotten can be immeasurably useful in the fight against nonconsensual pornography (Citron & Franks, 2014;

Cyber Civil Rights Initiative, 2020; Franks, 2015; Hartzog & Selinger, 2015). In 2015, Google instituted a policy to remove "revenge porn" from search results, showing that the company, and other companies in turn, can remove links to socially relevant information that can harm autonomy, reputation, and emotional well-being (Hartzog & Selinger, 2015). Google's decision provides an example to ground U.S. policy discussion around "the right to be forgotten" and how it might be instituted in the United States, in keeping with First Amendment concerns around the expression. But even this hopeful example demonstrates how regulatory contexts vary by country and are difficult to enforce across geopolitical borders. Furthermore, relying on corporate benevolence is a slippery slope, as the primary value of these companies is the economic bottom line, not the social good or the public interest (McChesney, 2008; S. Noble, 2013; Schiller, 1995).

The most harmful instances of audiovisual false impersonation found in this study were not technologically sophisticated. However, the response to this new hyperreality seemingly brought by deepfakes has included calls for technical and legal limitations, such as regulations, design features, and cultural norms, to finalize the role of the technology in society to promote the goals of people who are already powerful (Pinch & Bijker, 1984). The alarmist discourse around AI-generated images, audio, and audiovisual clips has led to economic opportunities for those seeking to mitigate the superficial risks that maintain the status quo and individualize the risk of being faked and being tricked by a fake. Many popular press articles have proposed large-scale information and media literacy programs for users (Bernal, 2018; Sunne, 2018) so that they can learn to detect fakes, to install special browser plug-ins, and to employ tools to avoid being faked or believing a fake. Technologists see economic openings in the development of technical solutions to address security vulnerabilities. Technical approaches usually fall into one or both of the following categories: automated detection of manipulated content and "distributed verification technologies to verify all online and even offline interaction" (Paris & Donovan, 2019, p. 22). The U.S. Department of Defense and social media companies, such as Facebook, have been developing automated systems to detect AI-generated manipulations (Facebook, 2018; Turek, 2018). The start-up company Truepic (n.d.) has touted blockchain-verified press-captured images and videos. Other verification proposals require users to verify themselves and to label all content (Pornhub, 2020).

The media, technology industry, and other knowledge institutions have overwhelmingly positioned technical solutions as objective and neutral, that is, devoid of the messy constellations of power in society. This enthusiasm for technical solutions for fake videos presages a future in which the speed and opacity of the solution threaten a democratic internet even more than the problem itself. Specifically, these technical solutions promote a future in which formal and informal governance is decided by an increasingly few of the

already powerful who have a vested interest in maintaining their power. Individualized literacy solutions also miss the mark because they put the onus on individuals to protect themselves in a harmful information environment but do nothing to hold those who profit from the harm accountable.

These technologies of audiovisual capture and manipulation cannot be uninvented, but they can be reconfigured. As have the scholars in STS (Barad, 2003), political economy of technology (Braverman, 1974/1998; Harvey, 2003; D. F. Noble, 1984/2013), and critical race theory and technology (Benjamin, 2019b; S. U. Noble, 2016, 2018), future studies must meaningfully address structural power as it is intertwined with technological deployment and use. To create more just sociotechnical systems, we must reconfigure these systems. Regulatory entities might re-imagine ownership, control, guidance, and develop deeper, more nuanced considerations of the politics of evidence. For example, these powerful technical systems that cause so much harm could be reformed in keeping with the EU General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) regulation around the right to be forgotten. Or they might be dismantled and replaced by a new cooperatively-owned and -governed internet infrastructure (MacLellan, 2021; Paris, 2021; Dan et al., 2021; Scholz, 2016; Tarnoff, 2019) where users or collectives of users, not corporations, are more meaningfully in control of creating and enforcing accountable and meaningful processes through which individuals who are abused, targeted, and extorted by using technologies can address these abuses (Citron, 2016; S. U. Noble, 2018; Roberts, 2019). A regulatory and legal policy can press technology companies for effective refutation strategies for those who are most disadvantaged in the online environment. Addressing contemporary sociotechnical problems like that of audiovisual misinformation and harassment online must take many forms, but at the core, meaningful redress requires a deliberate reconfiguration of technologies to mitigate the harm that allows the public to retain power over the use of these technologies.

## Conclusion

This qualitative study's goals were to provide a critical reflection on technologies of audiovisual impersonation, not an to supply or analyze every case of audiovisual manipulation that exists. The examples analyzed suggest that manifestations of audiovisual manipulation of false personation both strengthen and are shaped by structural power—Deepfakes were used by BJP in India to target specific language demographics, and cheapfakes were used in Australia, India, and the Philippines to silence, coerce, and harass a young woman, a woman journalist, and a closeted gay economic minister. In the U.S., effects experts and technology companies can make enormous amounts of money from producing applications that can produce viral video filters, deepfakes, and holograms for entertainment or enjoyment; amateurs can use

rudimentary deepfake technology to “pornify” their classmates, coworkers, and anyone they choose. These examples show how stakeholders situated differently within existing social structures participate in the politics of evidence online—who gets to make these decisions and enjoy the benefits that flow from these decisions and who is left out of these discussions and exploited. Addressing this phenomenon requires more than simply promoting systems that work for a generalized subject, which is always coded as White, male, upper-middle class, heterosexual subjects. Closer attention must be paid to those who are harmed and are unable to advocate against that harm. The effects of power differentials should be at the forefront of the development of social practices, particularly those related to the regulation and moderation of the technologies for the dissemination of expressive and evidentiary content.

This article has shown the complementary qualitative methods and considerations that can improve the understanding of the social, political, and historical contexts and consequences of images or audiovisual content. The approach can be replicated in future comparative qualitative studies. The insights gleaned from this study could provide justifications for rejecting the uncritical development and deployment of algorithmic technology, especially that related to “computer vision” and visuality. This demonstration of a critical informatics approach is one of the few such studies in this field. With the ongoing crises and positioning of technology as the only expedient or most logical solution, these types of critical analyses are increasingly necessary.

## Acknowledgements

The research presented here expands on the author's previous work on audiovisual manipulation in “Deepfakes and Cheapfakes: The manipulation of audio and visual infomraiton online” co-authored with Joan Donovan and published by the Data & Society Research Institute.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Britt Paris  <https://orcid.org/0000-0003-1527-7953>

## Note

1. This is generally true in online pornographic communities. However, pornographic practice is not monolithic. The production of counter-hegemonic pornographic content by and for women, LGBTQIA individuals, and people of color suggests that greater visibility and access to sexual culture for nondominant groups might be described as liberatory (Attwood, 2007).

## References

- Acker, A., & Donovan, J. (2019). Data craft: A theory/methods package for critical internet studies. *Information, Communication & Society*, 22(11), 1590–1609. <https://doi.org/10.1080/1369118X.2019.1645194>
- Adjer, H., Patrini, G., & Cavalli, F. (2020). *Automating image abuse: Deepfake bots on Telegram*. Sensity.ai. <https://sensity.ai/reports/>
- Attwood, F. (2007). No money shot? Commerce, pornography and new sex taste cultures. *Sexualities*, 10(4), 441–456. <https://doi.org/10.1177/1363460707080982>
- Ayyub, R. (2016). *Gujarat files: Anatomy of a cover up*. CreateSpace Independent Publishing Platform.
- Ayyub, R. (2018, May 22). In India, journalists face slut-shaming and rape threats. *The New York Times*, p. sec. Opinion. <https://www.nytimes.com/2018/05/22/opinion/india-journalists-slut-shaming-rape.html>
- Bansal, A., Ma, S., Ramanan, D., & Sheikh, Y. (2018, August). *Recycle-GAN: Unsupervised video retargeting* (arXiv:1808.05174 [Cs]). <http://arxiv.org/abs/1808.05174>
- Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs*, 28(3), 801–831. <https://doi.org/10.1086/345321>
- Benjamin, R. (Ed.). (2019a). *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life*. Duke University Press.
- Benjamin, R. (Ed.). (2019b). *Race after technology: Abolitionist tools for the new Jim code* (1st ed.). Polity Press.
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bernal, N. (2018, November 15). Journalists at Wall Street Journal to be taught to identify “deepfakes.” *The Telegraph*. <https://www.telegraph.co.uk/technology/2018/11/15/journalists-wall-street-journal-taught-identify-deepfakes/>
- Bevins, V. (2020). *The Jakarta method: Washington’s anticommunist crusade and the mass murder program that shaped our world*. Public Affairs Books.
- Bimber, B., Gil de Zúñiga, H. (2020). The unedited public sphere. *New Media & Society*, 22(4), 700–715. <https://doi.org/10.1177/1461444819893980>
- Braverman, H. (1998). *Labor and monopoly capital: The degradation of work in the twentieth century* (25th Anniversary ed.). Monthly Review Press. (Original work published 1974)
- Brock, A. (2018). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- California State Legislature. (2018). *California Consumer Privacy Act of 2018 [1798.100 - 1798.199.100]*. California Legislative Information. [https://leginfo.legislature.ca.gov/faces/codes\\_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5](https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5)
- Caplan, R. (2018). *Content or context moderation?* Data & Society. <https://datasociety.net/output/content-or-context-moderation/>
- Christopher, N. (2020, February 18). Deepfakes by BJP in Indian Delhi election campaign. *VICE Motherboard*. <https://www.vice.com/en/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp>
- Citron, D. K. (2016). *Hate crimes in cyberspace* (Reprint ed.). Harvard University Press.
- Citron, D. K., & Franks, M. A. (2014). *Criminalizing revenge porn* (SSRN Scholarly Paper ID 2368946). Social Science Research Network. <https://papers.ssrn.com/abstract=2368946>
- Citron, D. K., & Wittes, B. (2017). The Internet will not break: Denying bad samaritans § 230 immunity. *Fordham Law Review*, 86(2), 401–423.
- Cole, S. (2017, December 11). AI-assisted fake porn is here and we’re all fucked. *VICE Motherboard*. [https://motherboard.vice.com/en\\_us/article/gydydm/gal-gadot-fake-ai-porn](https://motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn)
- Collins, P. H. (1990). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Unwin Hyman.
- Coopersmith, J. (1998). Pornography, technology and progress. *ICON*, 4, 94–125.
- Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *The University of Chicago Legal Forum*, 140, 139–167.
- Currie, M., Paris, B. S., Pasquetti, I., & Pierre, J. (2016). The conundrum of police officer-involved homicides: Counter-data in Los Angeles County. *Big Data & Society*, 3(2), 1–14. <https://doi.org/10.1177/2053951716663566>
- Cyber Civil Rights Initiative. (2020). *48 states + DC + one territory now have revenge porn laws*. <https://www.cybercivilrights.org/revenge-porn-laws/>
- Daston, L., & Galison, P. (2010). *Objectivity*. Zone Books.
- Dan, V., Paris, B., Donovan, J., Hammeleers, M., Roozenbeek, J., van der Linden, S., & von Sikorski, C. (2021). Visual mis- and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly*, 98(3), 641–664. <https://doi.org/10.1177/10776990211035395>
- Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, 23, 2072–2098. <https://doi.org/10.1177/1461444820925811>
- Douglas, S. J. (1987). *Inventing American broadcasting, 1899-1922*. Johns Hopkins University Press.
- European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679: General data protection regulation*. Official Journal of the European Union. <https://gdpr-info.eu/art-17-gdpr/>
- Facebook. (2018, September 13). Expanding fact-checking to photos and videos. *Facebook*. <https://newsroom.fb.com/news/2018/09/expanding-fact-checking/>
- Fairclough, N. (2013). Critical discourse analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 9–20). Routledge.
- Federal Communications Commission. (1996). *47 USC 230: Protection for private blocking and screening of offensive material*. U.S. House of Representatives. [https://uscode.house.gov/view.xhtml?req=\(title:47%20section:230%20edition:prelim\)](https://uscode.house.gov/view.xhtml?req=(title:47%20section:230%20edition:prelim))
- Feigensohn, N. (2011). *Law on display: The digital transformation of legal persuasion and judgment* (Reprint ed.). New York University Press.
- Fiske, J. (1996). *Media matters: Race and gender in U.S. politics*. University of Minnesota Press.
- Floridi, L. (2013). *The ethics of information*. Oxford University Press.
- Franks, M. A. (2015). *Drafting an effective “revenge porn” law: A guide for legislators* (SSRN Scholarly Paper ID 2468823).

- Social Science Research Network. <https://papers.ssrn.com/abstract=2468823>
- Franks, M. A. (2017). The desert of the unreal: Inequality in virtual and augmented reality. *UC Davis Law Review*, 51, 499–538.
- Franks, M. A. (2018). *Beyond “free speech for the White man”: Feminism and the first amendment* (SSRN Scholarly Paper ID 3206392). Social Science Research Network. <https://papers.ssrn.com/abstract=3206392>
- Franks, M. A. (2019). Unwilling avatars: Idealism and discrimination in cyberspace. *Columbia Journal of Gender and Law*, 20. [https://repository.law.miami.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1304&context=fac\\_articles](https://repository.law.miami.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1304&context=fac_articles)
- Gilliland-Swetland, A. J. (2000). Enduring paradigm, new opportunities: The value of the archival perspective in the digital environment. *Library Hi Tech*, 18, 383–386. <https://doi.org/10.1108/lht.2000.18.4.383.4>
- Golan, T. (2007). *Laws of men and laws of nature: The history of scientific expert testimony in England and America*. Harvard University Press.
- Goodrick, D., Atkinson, P., Delamont, S., Cernat, A., Skashaug, J., & Williams, R. (2020). *Comparative case studies*. SAGE. <https://doi.org/10.4135/9781526421036849021>
- Hall, S. (2001). Encoding/decoding. In M. G. Durham & D. Kellner (Eds.), *Media and cultural studies: Keyworks*. Blackwell. <https://blog.richmond.edu/watchingthewire/files/2015/08/Encoding-Decoding.pdf>. (Original work published 1973)
- Haraway, D. (1991). A cyborg manifesto: Science, technology and socialist-feminism in the late 20th century. In D. J. Haraway (Ed.), *Simians, cyborgs, and women: The reinvention of nature* (pp. 149–181). Routledge.
- Hartzog, W., & Selinger, E. (2015, June 25). Google’s action on revenge porn opens the door on right to be forgotten in U.S. *Center for Internet and Society, Stanford Law School*. <http://cyberlaw.stanford.edu/publications/googles-action-revenge-porn-opens-door-right-be-forgotten-us>
- Harvey, D. (2003). The fetish of technology: Causes and consequences. *Macalester International*, 13, 7.
- Harwell, D. (2020, October 20). A shadowy AI service has transformed thousands of women’s photos into fake nudes: “Make fantasy a reality.” *The Washington Post*. <https://www.washingtonpost.com/technology/2020/10/20/deep-fake-nudes/>
- Irwin, A., & Wynne, B. (Eds.). (1996). *Misunderstanding science? The public reconstruction of science and technology*. Cambridge University Press.
- Jenkins, H. (2004, June 4). Photoshop for democracy. *MIT Technology Review*. <https://www.technologyreview.com/s/402820/photoshop-for-democracy/>
- Kneese, T. (2020, November 2). How data can create full-on apparitions of the dead. *Slate Magazine*. <https://slate.com/technology/2020/11/robert-kardashian-joaquin-oliver-deep-fakes-death.html>
- MacLellan, L. (2021, January 12). For better or worse, web infrastructure is not a public good. *Quartz*. <https://qz.com/work/1956070/internet-infrastructure-companies-should-be-public-utilities/>
- Maddock, S. (2020). “A deepfake porn plot intended to silence me”: Exploring continuities between pornographic and “political” deep fakes. *Porn Studies*, 7(5), 1–9. <https://doi.org/10.1080/023268743.2020.1757499>
- Marvin, C. (1990). *When old technologies were new: Thinking about electric communication in the late nineteenth century*. Oxford University Press.
- McChesney, R. W. (2008). Rich media, poor democracy. In *The political economy of media: Enduring issues, emerging dilemmas*. Monthly Review Press. [http://sites.psu.edu/comm292/wp-content/uploads/sites/5180/2014/08/McChesney-RichMedia\\_PoorDemocracy.pdf](http://sites.psu.edu/comm292/wp-content/uploads/sites/5180/2014/08/McChesney-RichMedia_PoorDemocracy.pdf)
- McGlynn, C., Rackley, E., & Houghton, R. (2017). Beyond “revenge porn”: The continuum of image-based sexual abuse. *Feminist Legal Studies*, 25(1), 25–46. <https://doi.org/10.1007/s10691-017-9343-2>
- Murphy, M. (2006). *Sick building syndrome and the problem of uncertainty: Environmental politics, technoscience, and women workers* (1st ed.). Duke University Press.
- Noble, D. F. (2013). *Forces of production*. Knopf Doubleday Publishing Group. (Original work published 1984)
- Noble, S. (2013, October 29). Google search: Hyper-visibility as a means of rendering black women and girls invisible. *InVisible Culture*, 19. <http://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>
- Noble, S. U. (2016). A future for intersectional Black feminist technology studies. *Scholar & Feminist Online*, 13(3). <https://sfonline.barnard.edu/traversing-technologies/safiya-umoja-noble-a-future-for-intersectional-black-feminist-technology-studies/>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism* (1st ed.). New York University Press.
- Paris, B. S. (2021). The internet of futures past: Values trajectories of networking protocol projects. *Science, Technology, & Human Values*, 46, 1021–1047. <https://doi.org/10.1177/0162243920974083>
- Paris, B. S., & Donovan, J. (2019). *Deepfakes and cheap fakes*. <https://datasociety.net/output/deepfakes-and-cheap-fakes/>
- Paris, B. S., & Pierre, J. (2017). Naming experience: Registering resistance and mobilizing change with qualitative tools. *InterActions: UCLA Journal of Education and Information Studies*, 13(1). <https://escholarship.org/uc/item/02d9w4qd>
- Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, 14(3), 399–441.
- Pornhub. (2020, December). The latest on our commitment to trust and safety blog. *Pornhub*. <https://www.pornhub.com/blog/11422>
- Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.
- Sasse, B. (2018, December 21). *Text—S.3805—115th congress (2017-2018), Malicious Deep Fake Prohibition Act of 2018*. <https://www.congress.gov/bill/115th-congress/senate-bill/3805/text>
- Schiller, H. (1995). *Information inequality* (1st ed.). Routledge.
- Scholz, T. (2016). *Platform cooperativism: Challenging the corporate sharing economy*. Rosa Luxemburg Stiftung.
- Sturmer, J., & Barney, A. (2016, October 12). Noelle Martin fights to have harmless selfie removed from “parasite” porn sites. *Australian Broadcasting Corporation*. <https://www.abc.net.au/news/2016-10-12/womans-fight-to-have-harmless-selfie-removed-from-porn-site/7924948>

- Suchman, L. (2000). Embodied practices of engineering work. *Mind, Culture, and Activity*, 7(1–2), 4–18. <https://doi.org/10.1080/10749039.2000.9677645>
- Sunne, S. (2018, May 28). What to watch for in the coming wave of “deep fake” videos. *Global Investigative Journalism Network*. <https://gijn.org/2018/05/28/what-to-watch-for-in-the-coming-wave-of-deep-fake-videos/>
- Sutherland, T. (2017). Making a killing: On race, ritual, and (re)membrance in digital culture. *Preservation, Digital Technology & Culture*, 46(1), 32–40. <https://doi.org/10.1515/pdtc-2017-0025>
- Suwajanakorn, S., Seitz, S., & Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama: Learning lip sync from audio. *ACM Transactions on Graphics*, 36(4), 95. <http://grail.cs.washington.edu/projects/AudioToObama/>
- Tarnoff, B. (2019, November 30). A socialist plan to fix the internet. *Jacobin*. <https://jacobinmag.com/2019/11/tech-companies-antitrust-monopolies-socialist>
- Tiffany, K. (2019, July 17). Probably don’t download FaceApp but also definitely don’t panic about it. *Vox*. <https://www.vox.com/the-goods/2019/7/17/20698271/faceapp-privacy-panic-russia-old-face-filter-app>
- Truepic. (n.d.). *Photos and videos you can trust*. <https://truepic.com/about/>
- Turek, M. (2018). *Media Forensics (MediFor)*. Defense Advanced Research Projects Agency. <https://www.darpa.mil/program/media-forensics>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- van Dijk, T. A. (2005). Critical discourse analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 349–371). Blackwell Publishers. <https://doi.org/10.1002/9780470753460.ch19>
- Voss, W. G., & Castets-Renard, C. (2016). Proposal for an international taxonomy on the various forms of the “right to be forgotten”: A study on the convergence of norms. *Colorado Technology Law Journal*, 14(2), 281–344.
- Waisbord, S. (2018). The elective affinity between post-truth communication and populist politics. *Communication Research and Practice*, 4(1), 17–34. <https://doi.org/10.1080/22041451.2018.1428928>
- Walden, M. (2019, June 12). “Dirty political games”: Gay sex video sparks scandal for Malaysia’s government. *ABC News Australia*. <https://www.abc.net.au/news/2019-06-12/gay-sex-video-sparks-scandal-for-malaysian-government/11201834>
- Ward, L. (2018). “Keep runnin’ bro”: Carrie Mae Weems and the visual act of refusal. *Black Camera*, 9(2), 82–109.
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Counsel of Europe. <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>
- Wodak, R., & Meyer, M. (2009). *Methods for critical discourse analysis*. SAGE.
- Wood, S. E. (2017). Police body cameras and professional responsibility: Public records and private evidence. *Preservation, Digital Technology & Culture*, 46(1), 41–51. <https://doi.org/10.1515/pdtc-2016-0030>

## Author Biography

Britt Paris (PhD University of California, Los Angeles, M.A. The New School) is an Assistant Professor of Library and Information Science at Rutgers, the State University of New Jersey, and an affiliate scholar at Data & Society Research Institute. Her research interests include audiovisual evidence, epistemic burden, and alternative internet infrastructure.