

Research Article

Travis L. Wagner*, Ashley Blewer

“The Word Real Is No Longer Real”: Deepfakes, Gender, and the Challenges of AI-Altered Video

<https://doi.org/10.1515/opis-2019-0003>

Received November 14, 2018; accepted February 6, 2019

Abstract: It is near-impossible for casual consumers of images to authenticate digitally-altered images without a keen understanding of how to “read” the digital image. As Photoshop did for photographic alteration, so to have advances in artificial intelligence and computer graphics made seamless video alteration seem real to the untrained eye. The colloquialism used to describe these videos are “deepfakes”: a portmanteau of deep learning AI and faked imagery. The implications for these videos serving as authentic representations matters, especially in rhetorics around “fake news.” Yet, this alteration software, one deployable both through high-end editing software and free mobile apps, remains critically under examined. One troubling example of deepfakes is the superimposing of women’s faces into pornographic videos. The implication here is a reification of women’s bodies as a thing to be visually consumed, here circumventing consent. This use is confounding considering the very bodies used to perfect deepfakes were men. This paper explores how the emergence and distribution of deepfakes continues to enforce gendered disparities within visual information. This paper, however, rejects the inevitability of deepfakes arguing that feminist oriented approaches to artificial intelligence building and a critical approaches to visual information literacy can stifle the distribution of violently sexist deepfakes.

Keywords: gender, artificial intelligence, Information Literacy, media studies, programming

1 Introduction

The challenges of verifying authentic imagery of famous individuals is hardly a new phenomenon. A mythos revolves around whether or not most pictures of Abraham Lincoln were indeed authentic as he was purported to be ungainly by news reporters. Rumors exist of his campaign engaging in primitive forms of photograph manipulation to alter his popular image (Fadrid, 2006). This anecdote shows that the problem of authenticity is as old as the image itself and the emergence of Adobe Photoshop and other digital image editing techniques continue to make it near-impossible for casual consumers of images to know whether or not they are originals without a keen understanding of how to “read” the image, digitally. As Mila Fineman (2012) notes “digital photography and Photoshop have taught us to think about photographic images in a different way—as potentially manipulated pictures with a strong but always mediated resemblance to the things they depict” (p. 43). Skepticism around photographs is now common-practice, yet the visual still possesses some degree of presumed truth. In the modern era, photographic manipulation is becoming auto-generated and normalized, often without the knowledge and consent of the author or persons

Article note: Gender issues in Library and Information Science: Focusing on Visual Aspects, topical issue ed. by Lesley S. J. Farmer.

***Corresponding author: Travis L. Wagner**, The University of South Carolina, E-mail: traviswagner@gmail.com
Ashley Blewer, Independent Scholar

depicted. Technologist James Bridle (2018) gives an example of Google's AutoAwesome (now simply known as 'Assistant') software that tweaks uploaded images automatically and invisibly: "Smith uploaded two [photographs], to see which his wife preferred. In one, he was smiling, but his wife was not; in the other, his wife was smiling, but he was not. From these two images, taken seconds apart, Google's photo-sorting algorithms had conjured a third: a composite in which both subjects were smiling their 'best'. ... But in this case, the result was a photograph of a moment that had never happened: a false memory, a rewriting of history" (p. 152). Without pause of thought this could be seen as a *real* photograph.

Along this thread of visual evolution, the moving image became the promise of a purer truth one that could not be complicated by the threat to alteration made feasible by still images. Film theorist Andre Bazin (1967/2008) suggests individuals believe the moving image offers "a total and complete representation of reality," a "perfect illusion of the outside world in sound, color, and relief" (p. 235). Bazin, however, was more incredulous of film's potential for authenticity. Much like Photoshop did for changing the public's understanding of trustworthy photographs, evolutions in artificial intelligence training and advanced computer graphics have resulted in a moment wherein altered video can seamlessly replace authentic video. A current example of this is a technology that allows computers to swap the faces of individuals onto others seamlessly. The colloquialism used to describe these particular videos is a "deepfake": a portmanteau of deep learning AI and faked imagery. The term is one disconcertingly borrowed from a Reddit user who bragged about the technology's potential for making fake pornography using the faces of celebrities, specifically in this case using the face of *Wonder Woman* star Gal Gadot within preexisting adult film footage (Cole, 2017).

Though the augmentation and filtering of original video sources is not a new concept, the degree of apparent veracity is. The implications for these videos serving as authentic representations of people and activities matters, especially in a culture of rhetoric fixated on "fake news" and citizen recordings of institutional traumas that seem to stand at uncompromising odds with one another. To date, this image alteration software, deployable both through machine-learning automated processing, remains critically under examined within studies of information literacy.

A survey of popular uses of deep fakes reveal two common uses:

- 1.) an infusion of the bombastic actor Nicolas Cage into other movies;
- 2.) superimposing feminine:femme faces into pornographic videos. We use the terms feminine and femme here with an understanding that gendered ways of being exist beyond a binary of masculinity and femininity and are distinctly different from one's sex assigned at birth.

The second use remains particularly troubling, primarily for its reification of women's bodies as a thing to be visually consumed, here completely circumventing any potential for consent or agency on the part of the face (and bodies) of such altered images. Yet it also proves telling, especially considering that the very bodies used to test and train AI to see and alter video are those of able-bodied, cisgender white men (Munafò et al, 2017). With this acknowledged, this paper explores how the emergence and subsequent distribution of deepfakes serve as a profound example of one potential future that continues to enforce gendered disparities within visual information production. This paper rejects the inevitability of deepfakes to become 'normal' and argues instead for both an emboldening of feminist-oriented approaches to artificial intelligence-building and a commitment to critical approaches regarding historical biases in media production as they relate to visual information literacy. The belief here is that in doing so practitioners across multiple fields can deter (and potentially even cease) the distribution of potentially violent, exploitative, and sexist deepfakes. We argue that this must happen by not only negating their validity but by further interrogating the role of agency and identity in visual representations holistically. To do this necessitates a deeper exploration of what role ethics and identification play in visual information as a field more generally.

2 An Overview of Visual Information Ethics and Gendered Representations

Even as the challenges of locating a true visual information linger over how individuals trust photographs and moving images, the very act of interpreting information that mimics reality remains confounding. Cultural theorist Jean Baudrillard's (1981) *Simulacra and Simulation* attempts to make sense of this conundrum. Noting an increased shift towards versions of reality via simulation and mediated replication, Baudrillard warned of a potential moment wherein hyperreality would become indiscernible from the reality in which humans existed. In his discussion of this, Baudrillard locates the simulacra (or the replication of an image or individual) as being distinct from a mere simulation, which begins with a replica that is semi-real (clearly a replica), to one that is unreal (a copy of a copy of the replica). Eventually this image (copy of a copy) becomes so distinctively different that it becomes hyperreal. This hyperreal representation then becomes both distinctly a new version of reality that is different while simultaneously being indistinguishable from the moment of origin. Practically speaking, a version of this is hard to locate within contemporary media as any use of computer graphics retains a distinct difference from the depiction of real recorded.

Further, any attempts to get close to the replication of human perception run into the issue of diving into the uncanny valley. This concept refers to the degree to which a robot (or in this case graphic representation of a human) replicates human functionality too closely and thus becomes disconcerting (Mori, 1970). An infamous example (for the ridicule it receives for just this issue) would be Robert Zemeckis's 2004 work *Polar Express*. *Polar Express* was visually impressive, rendering lead actor Tom Hanks in an almost lifelike manner. The problem, however, was the proximity to realness was both too close, yet slightly too dissimilar, resulting in psychological distress for viewers who found this (and the other CGI renderings in the film) unsettling. This disconcerting reaction was so severe that the film gathered multiple negative reviews despite having state of the art technical effects (Eberle, 2009). To that end, even though computer scientists and philosophers alike have tried to imagine a way to get a CGI that is able to leap over the uncanny valley, no work entirely produced in CGI has yet to succeed. While production of visual images is one part of the equation of users and visual information literacy, the other role is one of perception.

Perceptibility remains contingent on what is observable only by a human viewer and those things that a computer can learn to see. These human images then reside on a divide between the things perceivable by a human and those capable of being ordered by a computer. However, before this distinction between what is knowable by a human and a computer with regards to visual image, we must contend with the very notion of meaning making within visual information. Art historian Erwin Panofsky (1955/1982) provides a germinal framework for the hierarchies of interpretation within visual images. These hierarchies (or levels), Panofsky asserts, afford viewers a means to distinguish between the more naturalistic components of a visual image and those tied to emotive and objective interpretations. Specifically, Panofsky identifies three layers of interpretation. First is the pre-iconographical level, which is used to represent the natural and factual components of the image. As an example, a pre-iconographical interpretation would note the hues of red in an image. Second is the iconographical analysis stage which represents the images and allegorical moments in a visual item. Here an example might be that of the aforementioned red color being present within a heart-shaped object. Finally, the third level is that of iconological interpretation, wherein meaning is added to an image. Concluding with an image their might be that of a red heart, one might apply the idea of 'love' to such an image (Panofsky, p. 40). It is within these frames that individuals can make meaning of objects in unison with one another. In conversation with Panofsky's ideas, Sarah Shatford-Layne's (1986) work provides an account of how latent cultural meanings complicate one's engagement with visual materials, here building directly off of the work of Panofsky. For Shatford-Layne, it is crucial to make marked distinctions between what she describes as an "ofness" and an "aboutness." Ofness serves to denote general descriptions such as a physical entity present (i.e. a sickle and a moon), whereas aboutness indicates potential cultural signifiers and ideologies tied to the entity (ie. communism) (1986, p. 44). Effectively, such distinctions provided a moment of pause, which still exists, regarding the imagined potential of both predicting and describing all possible interpretations of a given image. Asking about how

such divisions manifest themselves can further nuance discussions around literacies of visual information to the extent that it asks the role of ofness/aboutness as they pertain to validity. What remains absent in both Panofsky and Shatford-Layne’s discussions are the larger systemic roles of social discourses which privilege certain narratives over others. Here seeing is assumed to be culturally contextual, however, the role of privileged viewpoints receives little nuance. More directly, no mention is made of who consistently gets to do the seeing and who is consistently being seen within such encounters.

Visual perception and its gendered consumption are integral to discussions of understanding what is seen. In *Ways of Seeing* John Berger (1972/2008) argues that visual consumption historically exists to reify masculine roles of looking and obtaining the object upon which they are looking. Historically, this meant the feminine/femme body became the subject of this looking and could only serve functions by which masculine comfort in viewership was assured. Berger provided a useful dichotomy here of masculine subjects as being “surveyor” while feminine/femme subjects were “surveyed” (p. 46). We want to argue here that Berger’s model better operates within a discussion of deepfakes as a user/used dichotomy within the creation of deepfakes. As will be noted, the predominant users of deepfakes are likely cisgender, presumably straight men, and the images used are those of cisgender women. The consumption and pleasure tied to this process invite the use of Laura Mulvey’s (1975) foundational notion of the male gaze which asserts the presumed viewer of mediated content (specifically cinema for Mulvey) to be a straight male (by extension of the current historical moment we can also assume them to be white and cisgender). As a result of these critiques, the pervasive place of women and femme-presenting bodies within media become disconcertingly understandable. Again, while Mulvey is particularly attending to the information resource of film, Caetlin Benson-Allott (2013) tracks the continued objectification of women’s bodies from, as her text’s title suggests, VHS through file Sharing, thus cementing visual information as a means for consistently objectifying female and femme bodies via the views of a male gaze both in historical and contemporary contexts. Perhaps the most telling part here, however, is that even when mediated representations exist in purely digital formats the structuring around the objectification of women’s bodies remain. Lisa Nakamura (2013) makes a case for this occurring in online digital spaces, while James D. Ivory (2006) notes a prevalence of sexist visual information within video games. Again, in all of the aforementioned examples the role of male-dominated content creation bias the images to be those desired by straight male users and are thus constructed to please straight male users. As we will further argue, this helps illuminate why the use of deepfakes to produce pornography proves so rampant. If the creators presume the subjects to be a thing to be “used” it is clear that their consent within and opinions on the creation of deepfakes remain unacknowledged. Yet, seeing bodies and creating a corpus of what bodies can and do look like in increasingly virtual spaces proves more and more impacted by artificial intelligence warranting a further discussion about how image creation and retrieval factor into representations of gender within deepfakes, especially around the normalization of certain ways of looking and being looked at.

To define the challenge between what a human can perceive about a piece of visual information and what can be taught to a computer has resulted in what scientists term the “semantic gap.” Defined within image retrieval studies and technology as “the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation” the semantic gap ostensibly reflects the challenge of contextual interpretations by users (Smeulders et al., 2000, p. 1352). A simple way to divide this would be to use Shatford’s distinction between the aboutness and the ofness of an image, however, one might more accurately define this as a division between cognitive perceptions of an image (or in our case moving images) and the emotive responses such images evoke. It is the difference between looking at a picture Dorothea Lange’s *Migrant Mother* and seeing the colors black and white and looking at the picture and feeling abstract concepts like resilience and suffering. One should pause here to consider that the role played by the woman within Lange’s photograph is not one that adheres to the above dichotomy of user/used, at least not in the same notion of consumption. What does occur is a negotiation of Lange’s subject as being a woman who represents the burden of childcare and maternal labor for which she was likely not paid, while here her image was used as a means to leverage a cause by which men benefited. As Vince Leo (1996) argues, Lange’s subject became the image of a change in public notions of workers’ rights, but for the historical moment the woman pictured could not benefit directly,

nor did she receive financial compensation for her labor as a potential caretaker. As such, much like a computer cannot negotiate notions of suffering and exploitation within its interpretation of an image, it is also equally incapable of understanding how the information it is provided is biased towards certain views and ideologies.

Many images, regardless of computer-interaction, remain mired in the human response to evocative images. And with these evocations emerge potential biases. No contemporary work has done more to illuminate this issue than Safiya Umoja Noble's (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. In her work, she argues that a misalignment of technology as a neutral entity has led to systemic bias within the organization of information within search engines, wherein racist imagery (some of which is doctored imagery) emerges around certain search terms. Marked disconnects Noble observes include: searching for doctor and only seeing images of white men; searching for black teenagers and only seeing mugshots; or the stark difference between an image search for White Women and Asian Women. Yet, when such biases emerge companies plead that they are unanticipated glitches, to which Noble argues are not "occasional one-off moments" but indicative of "the ways that marginalized people have come to be represented by digital records" (2018, p. 6). This dissonance becomes clear when a computer is engaging in the disorientation and discombobulation of subjective and objective interpretation, something a well-done deepfake has the power to do. This is why we argue that it is paramount that information professionals and programmers understand how deepfakes emerged, what common examples look like, how they work, how they have evolved, and how they continue to reify a field of visual information that is both supremely sexist and meticulously misogynistic.

3 What is a Deepfake?

To understand this new form of media, necessitates first understanding what deepfakes are and how they are generated. As mentioned, the word is a portmanteau of "deep (learning)" and "fake (content)." Deep learning is the process of a computer system rapidly repeating procedures, identifying patterns of success within those procedures, and being able to generate new meanings from those patterns. In essence, it is a prototype of Artificial Intelligence. It is significant to note that a deepfake is more than just two videos that have been merged together to form one video by a person or group using advanced image-editing software (such as Adobe Premiere). Instead, the creation of deepfakes result from feeding information into a computer and allowing that computer to learn from this corpus over time and generate new content. The information required for deepfakes consists of low-quality cropped images of the faces of two people, processed separately. The computer then creates a comprehensive 3D model of these two people, as well as a mapping of characteristics between the two people's faces. The more time and the more processing speed a computer can utilize to learn from this data, the more developed and realistic the deepfake will be. To avoid the monotony of a human user going through and grabbing each frame of an individual's face available from a variety of data sources, programmers have coded scripts and designed programs available that make this kind of work relatively expedient.

Upon gathering this data, the computer then begins to work on understanding and recognizing these images without human intervention. Over time, by repeatedly running these simple tasks, the computer creates a model of each person's likeness. After enough time and training, the computer is then able to create new images that have never existed but can, with surprising validity, appear to accurately depict an entity. The computer is then able to take individual frames from a video and map the expression of a person's face in the video to the computer-generated replication of the generated model of a person's face matching that expression, ostensibly swapping in the new computer-generated image for each frame. This is similar to how visual effects studios operate for Hollywood films, but can be done without further human intervention and with only a collection of images of a person rather than themselves as performer of these actions. With a high-quality corpus of data and provided that the object and subject faces are similar (neither or both have facial hair, for example), this "face swap" transition can appear seamless. It is important to understand that new video content is created from a corpus of data containing the face of

one person, and, without a diligent lens for visual literacy and programming knowledge, such content can be presented to others as deceptively authentic. To make clear these moments of slippage from authentic to deceptively authentic and their particular relationship to gendered biases in the technology, we turn our attention to two prominent examples of the use of deepfakes, which track both their historical evolution and with the second example illuminate potentials for ethical, gendered exploitation.

4 What Does a Deepfake Look Like?

4.1 Nicolas Cage Example

Given the need for a large corpus of facial images to make deepfakes, it is of little surprise that Nicolas Cage became the point-person to exhaust the potentials of facial swapping. David McGowan (2016) argues that Cage’s inexplicably broad career from Oscar-winning performances to B-Movie leading roles make him ripe for meme generation. Jokes take particular aim at Cage’s prevalence for accepting seemingly any role offered to him. McGowan (2016) asserts that the memes centering around Cage thus “obscure the wider narrative explanation behind many of the choices” Cage makes as an actor (p. 216). Taking up Cage’s inclination to star in any movie, one prominent use of Cage’s visage is the use of deepfakes to place Cage into roles he did not perform. Myriad YouTube pages devote content to inserting the actor into pre-existing situations aptly titled Nicolas Cage Deepfakes. Many of the iterations of the video’s note them to be the “WORST” deepfake available, often placing Cage’s face into moments so iconographic or absurd that the viewer can easily distinguish the difference between the original and the fake (Nick Cage Deepfakes, 2018). These examples reside comfortably within Baudrillard’s notion of the semi-real. Yet, it is inside this slight fracturing of ambiguity wherein a startling actuality emerges.

Existing alongside these variant examples is the most telling of uses of the Cage’s deepfake within a clip titled “Really Bad(?) NIC CAGE deepfake (face off) (deepfakes)” which takes the trailer of John Woo’s 1997 film *Face/Off* and places the face of Nicolas Cage onto that of John Travolta (Nick Cage DeepFakes, 2018). To remind those adverse to excessive 90’s action movies, the plot of *Face/Off* is such that cop Sean Archer (John Travolta) seeks revenge upon international criminal Castor Troy (Nicolas Cage) by switching faces with him to infiltrate his criminal organization. Upon discovery of this, Troy then switches faces with Archer in an attempt to sabotage the already occurring sabotage. The film, instead of using computer-generated visual effects, opts to merely have Cage and Travolta switch roles. Yet in this clip, the creator actually switches out the face of Nicolas Cage’s Troy onto that of John Travolta’s Archer. The result is impressive (albeit of a poor quality). Serendipitously early theorizing around technological face-swapping possibilities through digitized puppetry and mimicry emerged from Weise et al. (2009), which they aptly called *Face/off*. What is less immediately clear in this discussion is the presence of amateur work in these examples. These are silly and frankly non-threatening examples and this has something to do with Cage being a male actor. Alternatively, gendered challenges emerge is when the bodies depicted are those of feminine/femme individuals being viewed in compromising situations without their consent.

4.2 Pornography Example

The most disconcertingly prolific deployment of the technology used to generate deepfakes devotes time to modifying existing pornographic material to swap out the face of a performer with that of non-adult-film actress, most infamously occurring with the insertion of Gal Gadot’s face onto another adult film performer. This act is now understood to be feasible not only within this example of pornography, but within the production of pornography more generally. Adult film company Naughty America has produced at least one video in which they show their ability to swap faces of two performers within the same clip (Roettgers, 2018). Again the technology is still limited and the images “look fake,” but a person with any desire to see them as real could believe them to be deceptively authentic, especially with an untrained eye and a lack

of knowledge around how they might have been made. It also works here to reify the user/used dichotomy discussed earlier, here with the reworked Gal Gadot pornographic deepfake being used by the user, who is likely a cisgender male (U.S. Reddit user, 2018). This illuminates a continued inconsistency in representation and engagement with media for non-cisgender, non-masculine bodies. To this end, it has direct impact on the agency and representative knowledges of those being stereotyped, erased, or eroticized. Any advance in visual imagery emerges with the desires of straight, cisgender male consumption at the forefront and these deepfakes are no different. The types of questionable content produced within such technology want to see those face-swapped individuals objectified (or used). Yet the dangers of use in pornography also evoke a larger concern around the continued mediation and objectification of feminine/femme bodies. Deepfakes show us this quite clearly displaying a desire to place women—famous or otherwise—in positions of non consensual exploitation, often of a sexual nature. We, however, want to be clear that this is not a stance on the ethics of pornography, but more directly a condemning of the failure for consent and willingness to be factors within its distribution. This concern for how deepfakes are distributed, especially the potentially violent and sexually-charged ones matters, as very little is understood about the legality around such content production.

As it stands an actress that did not perform the actions depicted in the content can claim defamation. The adult film star may claim copyright on their work, arguing that the deepfake work is illegally generated and distributed modification of their work. In theory, each would be dealt with in a legal system that would favor the person negatively subjected to such occurrences. The bigger issues at hand here, however, is the rapid widespread distribution of this content that is done through anonymous sources; the only entity available to attack are the hosts of the content, not the creators of the content. As adult film entrepreneur Christian Mann suggests, adult film content distributors (professional and amateur) are no longer “interested in why something is good or bad, or even the psychological basis for the appeal of certain content, prurient or whatnot, other than to the degree that they can use analytics to figure out buy rates and join rates” (Curtin et al., 125, 2014). To Mann and others, it is not even a question of the sensibilities around pornographic content any longer, but merely the trafficking in videos that yield high hit rates which can, in turn, be monetized. The problem with this shift to monetization of videos is that computers are engaging in this process, producing content at alarming rates that can emerge as quickly as any content provider can work to take it down.

Part of this rise in content comes via what is known as user-generated work or amateur work. This is where individuals can record and upload their own content, with the potential of monetizing their intellectual property. Examples of this can include Camgirls, SnapChat communication, direct uploads to adult film video services, and personal websites. Two components of this are crucial to understand. First, the emergence of this type of pornographic content has resulted in what Niels Van Doorn (2010) claims to be a fracturing of what constitutes real life sexual encounters. Despite, as Van Doorn notes, these videos “featur[ing] bodies that do not conform to conventional beauty standards,” thus shifting industrial production of adult content. However, it also produced a facturing of what real life sexual encounters look like (Van Doorn, 2010, p. 426). As a result, sexual expectations for both real encounters and high end encounters have slipped into a duality. The production of sexual content online even in its amateur forms is a simulacra of a replication of what sex is supposed to be. Simply, the idea that faces could be replaced on any piece of content (and specifically adult content) is disconcerting, but what is still more troubling is the reluctance to question even ‘authentic’ visual encounters as just that. Amateur porn can look real, but the mediated lens on which it is being viewed is not that and until this distinction is more directly made, the evocation of deepfake technology is but an overextension of this breach of reality. As we will note, it is this need to see bodies doing things, especially within a dichotomy of male pleasure and consumption that must be attended to within rhetorics around gender and visual information literacy. It is not so much a simple act of deterring the production and distribution of nonconsensual pornographic deepfakes, but a long overdue discussion on why individuals seek and desire to see this information in any capacity.

Patrick Keilty (2012) observes that pornography consumption is distinctly different than other engagements in online spaces, particularly because “the goal is to project into a moment of perfect satisfaction: obtaining the perfect image, one completely adequate to our desire” (p. 44). The desire for

new experiences is somewhat futile as, Keilty believes, the embodied information practice of consuming pornography is one that always presumes limitless possibilities. Fittingly, the proliferation of this type of face swapping adult entertainment lacks the external surveillance of other individuals, primarily concerned guardians. Here adults are ostensibly the same individuals producing and consuming the deepfakes so spaces ignore, circumvent, or avoid censorship entirely. Troublingly, even as this content is taken down the proliferation of such uploading occurs at exponential rates. As an example, adult video streaming provider Pornhub has joined with tech giants such as Twitter to work out ways to suppress and block the emergence of deepfake videos by putting in place official policies to remove deepfake content as it is in violation of their sites stance against non-consensual sex. Furthermore, the site deploys software that will track down any emergent examples of deepfake content and remove it and invites users on the site to report any content as it emerges. (Sharman 2018). Yet, even with these policies and practices in place, the deepfakes prevail. As anecdotal evidence, while we were looking for articles on Pornhub's policy with regards to deepfakes, links for Cameron Diaz deepfakes on Pornhub's site emerged noting that intended interventions failure (Figure 1).

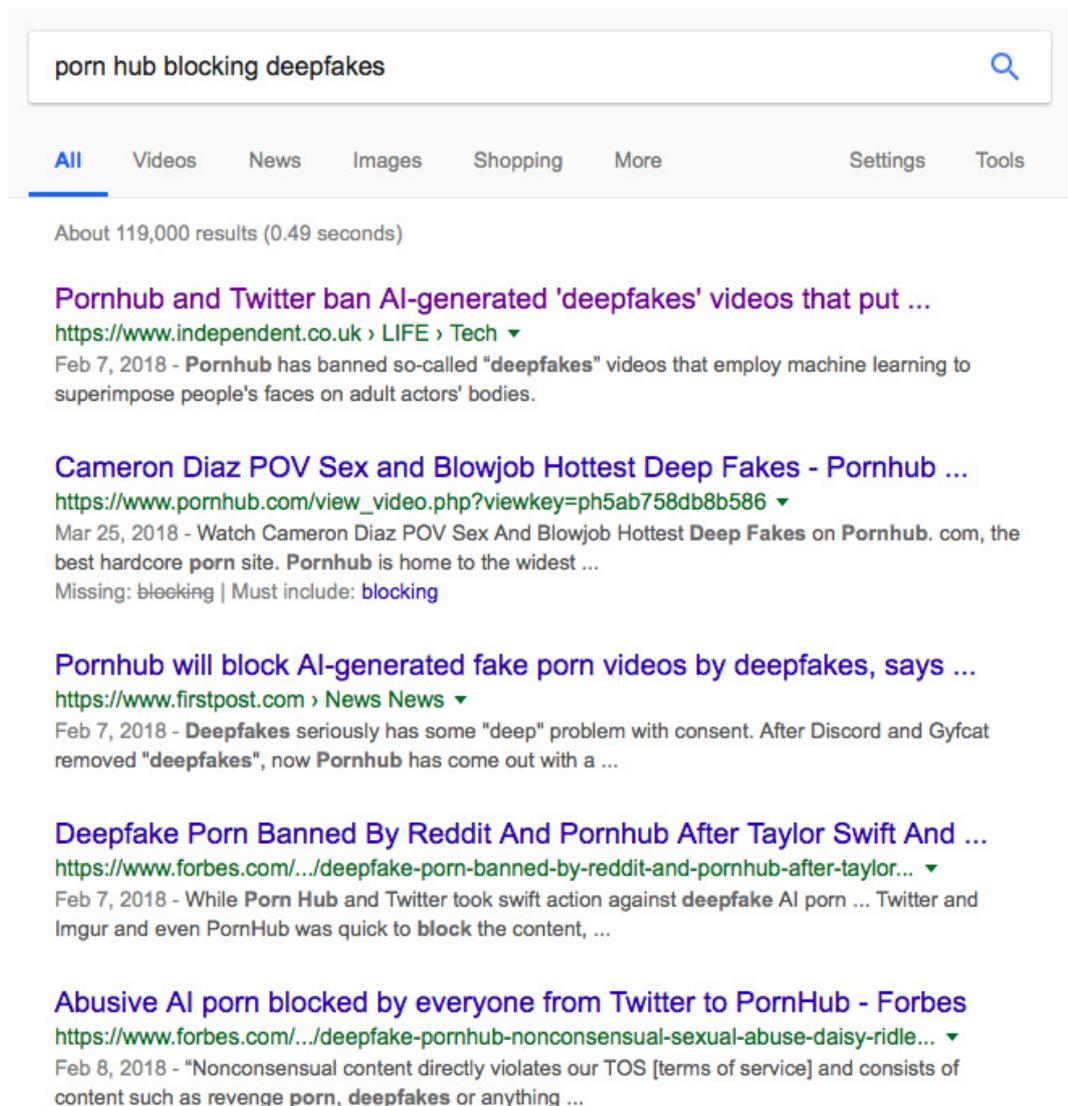


Figure 1: Narrative dissonance between assertion of concern for deepfakes and actualities of prominence of deepfakes. Screenshot from November 6, 2018

Part of the potential explanation of the oversight here is that a website like Pornhub is inundated with content at an uncontrollable rate. According to Pornhub's own data aggregation in the year 2017 they received 4,052,543 video uploads to their site, by both professional and amateur accounts which constituted over 595,482 hours of footage. This is roughly 68 years of total footage ("2017 Year in Review" 2018). Simply put, digital content is produced at an unstoppable rate.

Yet even as adult media production expands, Zabet Patterson (2004) rightly observes is that there remains a startling "blindspot" around the very content itself, instead singling in on the viewer. Patterson shows that the very concern of "cyberporn" and its ability to change sexual consumption has been placed directly onto the "featureless everyman" who consumes via technology "a near instantaneous mass mediation and dissemination of sexual representation" (p. 104-106). Thus, the impact of reading that Pornhub houses over six decades worth of video matters. It is not simply a question of the who that is consuming it, since Patterson asserts that very little has changed around our understanding of online pornography since a 1995 *Time* magazine expose. Alternatively, the question resides on how gender is shown, constructed, or negated within adult videos and crucially how deepfakes align with the inability to describe content. To this end, discussions around gendered representations within pornography are not necessary. Nor are they required as the above authors, as well Linda Williams (1999), explore this topic in depth. What we have to instead ask is how AI-produced mediations expand on this very failure to understand the significance of a content in context, then turn to understand why the viewer comes to this content. The user and used separation cannot be discussed in isolation of one another.

The meteoric rise of the production of deepfakes align with amateur production as it suggests transparency and simplicity regarding agency and control over one's distribution of content. However, the distribution of videos often go well beyond the original producer. As Soha and McDowell (2016) show, the secondary factor of digital media production is that others can compile remixes, mashups, and compilations. They show the prevalence of this with the many "best of" videos uploaded by YouTube accounts of meme content relating to the Harlem Shake which resulted in a "hybrid of robust amateur content alongside increasingly professionally produced video channels" (Soha and McDowell, 2016, p. 3). Part of the rise of this particular set of productions are the benefits of monetizing such content. To this end, the combination of a gendered consumption of women's bodies with a mass demand for "real" content and the ability of anybody to recreate and redistribute this content results in an ever rising amount of versions. As it stands, of these videos, most deepfakes are quite difficult to catch via regulatory software and practices, as they do not track for user generated content, do not know to look for compilation files, or, more crucially, cannot track content that exists outside their sites. Further this new mode of production means that each new manifestation of such content helps to build upon the very technology's ability to better create deepfakes.

5 Challenging the Potential Pervasiveness of Deepfakes

5.1 Programming as Advocacy

Much like our other discussions of gender so far, the problem is not how to cut the distribution of deepfakes off exclusively, but further explore what type of social space suggested this to be an acceptable practice in the first place. The answer to this question then becomes not what the technology does, but the ethical vantage points of the individuals who felt it necessary to produce such a technology. Tellingly, the technology industry consists of 25% women, in comparison with 57% of the general workforce, a gendered discrepancy on par with the larger population of male users within Reddit. Furthermore, the tech industry proves notoriously hostile to those identifying as women or non-binary, creating decreases in women graduating with computer science degrees. This leads to an inability to retain women within the field and causes distress and anguish to those who choose to stay in the field (Friedman, 2017). As technology evolves it is essential to develop ethical frameworks for mentally ingesting the complexity of these systems and their real ramifications on people: before, during, and after the development of new software. In 2018, Amazon (Conger, 2018), Google (Conger, 2018), and Microsoft (Frenkel, 2018) workers protested and resigned in

opposition to working on tools being developed for government surveillance, citing them as immoral. In June 2018, Google, arguably one of the leaders in AI-assisted technology, published a statement and guide on their principles with the stated goals which include: being socially beneficial, avoiding creating or reinforcing unfair bias, building and testing for safety, being accountable to people, incorporating privacy design principles, upholding high standards of scientific excellence, and being made available for uses that accord with these principles (Pichai, 2018). This is a good start if taken seriously, but ethical and human-based decision-making needs to be at the foundation of every software project.

The AI Now Institute, a research institute studying the social implications of artificial intelligence at New York University, breaks the development of AI-assisted technology into three large-scale categories of thinking: material resources, human labor, and data (Crawford, 2018). The ethical decision-making when developing software is sitting on top of a mountain of exploitation and oppression. The development of a Code of Ethics can only go as far as the people willing to follow the code, so there unfortunately must be work done towards advocating for legal precedents to be set for dealing with this new form of media manipulation. Some existing laws may apply currently in the United States, although they will have to be acted upon and played out in modern court systems to set an adequate legal precedent for this new form of work. The Electronic Frontier Foundation, an international non-profit dedicated to digital rights, recaps some of these legal justice paths in a blog post, arguing that extortion, harassment, defamation, and copyright infringement laws are adequate in the execution of justice against these types of videos (EFF, 2018). This is true at the personal level for those individuals with the time and money to take such claims through legal due process, but in many cases the damage may be irreparable if the faked content is used to misguide and shift opinions about the person’s image being used.

Technology has been and must continue to be developed to counteract and rapidly debunk manipulated media. Novice image forensics tools created without the expressed intent of working against deepfakes can be applied towards this effort, and in the future tools can be specifically created for use against deepfakes. A tool such as Forensic Magnifier can analyze the authenticity of a still image by lightening or darkening the image, or by increasing the “noise” (random variations of brightness in images) to determine inconsistent results (Wagner, 2015). Images can contain hidden metadata embedded within them, and free, open source tools such as Exiftool can help go beyond the metadata shown and reveal manually manipulated metadata using standard information tools (such as from right-clicking and getting more information about a file) (Friedl, 2006). Some social media sites will scrape this information upon upload, leaving the descriptive metadata fields as mostly blank (or replaced with site-specific metadata), while other sites do not perform this practice and leave the original metadata embedded within the image. Using the technology associated with deepfakes, new videos are being generated by computers based on a model of a person rather than edited by humans, so this detection is inherently different and the average person—at this moment—has less tools at their disposal to diagnose video validity. Additionally, “noise” generated by compression is common for internet video, so a lower-quality falsified video is capable of being presented as credible information (Heikal 2018). Because this material is likely to be considered illicit or removed from streaming services, viewers may be more likely to disregard other common forensic analyses like timestamps or upload times.

At this time but perhaps not for long, video can be used as evidence in court in many countries and internationally, such as in the International Criminal Tribunal for the Former Yugoslavia’s case against Slavko Dokmanović for crimes against humanity (“Filming Long After A Crime”, 2015) and International Criminal Court tribunal against Thomas Lubanga Dyilo (“The Role of Video”, 2016). SITU Research offers an online display of forensics research presented to Ukrainian courts in their work Euromaidan Event Reconstruction (Dykan and Iatsenko, 2018). Research and the studying of forensics related to video evidence has also resulted in art exhibitions, like the work of Forensic Architecture, which establishes this practice as an exhibition and public interrogation of video footage (“Project”, 2018). This is likely to shift with the increased ubiquity of video that has been tampered with through support of automated machine-learning algorithms. In the wake of this, these videos become exponentially more imperceptible and impossible to denounce as fake by average populations.

5.2 Teaching Visual Information Literacy

The immediacy of concern around increasing literacy against what seems like a potentially insurmountable problem is real and tangible. Nonetheless, the concerns brought forth by something like deepfakes are no different than the other issues plaguing contemporary discourses around information literacy. Literacy here must come from a position that notes latent biases in how technology and information represent the needs of diverse communities. Extending this, much like Kielty wants us to consider how persons engage in information with their bodies, it is also crucial to think through the material production of deepfakes. Part of the way information professionals can better prepare for the pervasiveness of deepfakes is to accept that materiality is part of how information is produced and exchanged. As Emily Drabinski (2018) argues “libraries are material, just as library workers and library patrons are. Each individual is “real, exist[s]” and “matter[s].” Understanding the material implications of the hyperreality of deepfakes might be how one gets over their becoming normalized. It is not so much that we promote paranoia around the content, but alternatively prepare users to engage with technology going forward. We must avoid asserting the products of technology exist in isolation and instead ask how the products got to us in the first place. Further, much like our earlier evocation of Noble suggests, we cannot think of glitches within technologies as being singular occurrences when it comes to the misrepresentation of bodies, instead we must attend to what Meyer and Wachter-Boettcher (2016) call “stress cases” or the “difficult use cases we don’t want to think about” (p. 42). So with the evocation of new advances in artificial intelligence, human computer interaction, or any other technological innovation, information professionals could serve a foundational role in this stress testing, offering up valuable information on how it might most negatively affect those stress cases and, in doing so, ask more of technology with regards to both users and how it will be used.

Rhetorics on technology often imagine it as exclusively liberatory, which is sometimes true, but hardly universal. In turns libraries are already spaces promoting ideas of universal design and its role in assuring for ease of access to even the most intense of stress cases, but this only comes through a deliberate attempt to meet community needs from a community oriented perspective, which “involves examining the power relations that have created and sustained the conditions in which we work” which results in “enabl[ing] some and not others” (Kumbier & Starkey, 2016 p. 488). We imagine this community-led approach as working within information spaces through some combination of community forums and training workshops. The reason for the tiered approach is the need for community input to take center stage in the influencing of how technologies are used, mobilized, and challenged within information spaces. Anne Goulding (2004) argues that community forums work to provide space to engage in topics that “enable members to meet and encounter one another” to discuss important and complex topics while “building a dynamic and active civic life” (p. 4). To this end, the community forums could introduce the challenge of deepfakes, perhaps in a larger set of discussions on fake news and misinformation. This open discussion could afford communities spaces to simultaneously understand the current trends in technology while raising concerns that may or may not be established within its current use. One can imagine here a generative discussion around the dangers of access to amateur deepfakes technology within their communities and within how they engage with visual information. Following this, and information space, library or otherwise, could then create a workshop to build guidelines on how to identify verifiable images, understand users of this technology, and implement a set of practices to give users agency over how they are represented within such technologies and their legal rights to their likeness in images.

The critical scholars mentioned earlier concerning gender and artificial intelligence (AI) make it clear that we cannot simply trust data to come to us without gendered (and, in turn, raced, sexed, embodied, aged, etc.) biases. Instead, we must prepare persons to see things with a lens that is not merely critical, but oriented towards a notion of social justice. Nicole Cooke (2017) notes this when confronting the perceived challenges around the deepfakes sibling problem of “fake news.” On it, Cooke states simply that “fake news is not new, and it is not going away,” however we can strive to create individuals who are “critical consumers and creators of information” and crucially serve in a “proactive” role to challenge misinformation when it emerges (p. 219). Individuals can push for companies to change their current policies and implement strategies that prevent the mass distribution of maliciously falsified content. A report from the Data &

Society Research Institute on content moderation and “fake news” identifies four strategies for mitigating information literacy: trust and verification, disrupting economic incentives, de-prioritizing content and banning accounts, and regulatory approaches. However, Data & Society warns that content moderation can be dangerous in itself as a tool that can be manipulated (Caplan et al, 2018). Data & Society’s approach also provides another way of engaging with not only deepfakes, but fake news more broadly, by providing taxonomic approaches to understanding the exact problems. By better naming types of fake news, we believe that we can potentially leverage control over the pervasiveness of deepfakes. More directly, making clear the difference between humorous and volatile content can allow information professionals to aid their users in avoiding their presence in and engagement with inflammatory deepfakes. The hurdle here that must be overcome is twofold.

First, the information professionals must make clear the limitations of a neutral stance on such technological innovations. While it is the role of information professionals to provide objective, authoritative research is equally crucial here to understanding that such information can appear to be all of these things, while merely mirroring what this type of information looks like. Again, this is not a unique challenge to larger discussions of fake news, wherein statistical information and peer-reviewed research can easily be replicated, but one of how this information is being used to make an argument or what one wants to see in the information. So, if an individual is seeking out imagery of a person doing something suspect, what might this mean for their intentions. Neutrality within information science historically asserts that a direct statement cannot be made about what is good versus bad with regards to information, but we want to implore a change for this, especially around visual information which evokes a distinctly different set of feelings some with embodied implications that cannot be easily alluded.

To this end, the second need of information professionals within this shift to acknowledging and attending to the pervasiveness of deepfakes is to note the role of gendered identity and sexuality within the production and perusal of information. For purposes of this, Melissa Adler (2018) posits that the history of libraries within their namings of sexuality tend to make quite clear a need to regulate sexualities and desires that are coded within social discourse as perverse, making these exceptions to inclusion within what is accessible to users. Moreover, the types of heteronormative sexualities which were then included become those which were deemed neutral (and normal). The non-normal sexualities then became what Adler calls “paraphillic” those desires which were perverse and might “cause distress or impairment to the individual or harm to others” (Adler, 2018, p.50). Following Adler’s line of thinking, we can understand discussions around sexuality and desire within information spaces to be presumed heterosexual and any actions around this identity marker remain regarded as neutral. In opposition, when an embodied desire differentiates from this it must be named, regulated, and heavily censored or surveilled. Common examples of this are LGBTQIA+ book displays remaining controversial or the continued placement of resources on transgender identity near information on pedophilia and necrophilia. Yet, what remains deeply confounding here is that part of the justification for these desires being regulated is that they might do “harm to others.” Information professionals should see the presence of sexually exploitative deepfakes as an example of this and consider ways to combat their prevalence and distribution within information spaces. A turn towards this allows information professionals to begin engaging in the user/used dichotomy mentioned earlier, but here with a keen sense of why certain gendered bodies are more likely to be doing the using of deepfakes, while others are likely to be used in the process.

So for information professionals to challenge the ubiquity of deepfakes they cannot simply adopt the mantra that “seeing is no longer believing,” but must understand that it was never the case to begin with. Instead, information professionals must adhere to a proactive critical informatics. As Cooke evokes, the critical information engagement (both of professionals and non-professionals) must challenge deepfakes when they emerge, especially those that emerge via malevolent means. If footage of a celebrity (especially those who are feminine/femme-bodied) in a compromising position emerges one must pause and ask about the validity of the content and the reason the person who shared it felt it necessary to do so. As Nancy Baym (2015) asserts, continued increases in technological interconnectivity and its resulting expanse of digital content requires “networked collectivism” wherein “mobile media and in-person communication” collide to construct a “distributed group identity” (p. 90-92). In acknowledging this, the role of information

literacy must attend to this networked era and its emphasis on a participatory culture and take up the task of confronting the problems of visual information (Jenkins et al., 2016). We want to suggest the potential for information professionals and programmers to work together here, as their collective work is part of this very networked collectivism. This can allow both groups to confront misinformation like deepfakes, while understanding how it works both with regards to lines of code and to individual embodied information needs. A major venue for these conversations to happen would be a space like the Code4Lib whose journal, listservs, and conferences aim to engage in these very topics in a way that advances the work of information professionals, while possessing a keen awareness of the need to build a “diverse and inclusive community” which “seek[s] to share ideas and build collaboration” (About, 2018).

6 Conclusion

To borrow from Baudrillard and the YouTube comment quoted in the title of this paper, we may exist in a moment in which non-reality is as real as reality. However, in either reality, some individuals face more inequality than others. Seeing these inequities, in turn, becomes the first step in changing visual literacy in a radical and meaningful way. This is a critical moment for authenticity and trust in media, and it is palpable in the news and academic writing about this issue. Technology has already advanced to the point of real-time media manipulation in which faces and speech can be applied to completely false actions/statements but be perceived as real by the public. Practical and easy-to-use applications are actively being developed and improved upon to generate false media and even without this technology, propaganda is already actively distributed to create confusion and deny real video. Very soon, all of these pieces will join together and we must collectively work quickly to be prepared for how to deal with this new reality. Yet, even as we move towards new reality, our very questions of mediated representation continue to remain the same. Paradigm shifts happen around consumption of images, but those most subject to their exploitation remain affected in profoundly negative ways. Like all issues of equity in representation that have come before, one must *say something* when they *see something*. More immediately, we must be keenly aware of the veracity of those being exploited for their traumas and for whom the potential for danger remains deeply real. Information professionals have plenty of individuals standing in line prepared to tackle the challenges of misinformation in a hypermediated age but the conversations remain divided on two sides of a the field: technology production and user engagement. This crossroad within misinformation can be dealt with by having information professionals do what they know best, providing clarity and understanding to a vast amount of information, here for the net good of individuals. Information literacy is no longer an expert knowledge that is taught from a position of knowledge that must be retained and regurgitated, instead it is something the collective of information users must actively resolve to make good. Information exists in a myriad of uses waiting to be done good with, but it is equally true that their misuse remains a potential. Thus advocacy falls on the hands of creators, consumers, redistributors, and even those caught in the crossfire. Amongst other things, transparency and advocacy become inextricably linked to the most finite lines of code that make up what we see and how we talk about it.

References

- Adler, M. (2017). *Cruising the library: Perversities in the organization of knowledge*. Fordham University Press.
- Baudrillard, J. (1981). *Simulacra and simulation*. The University of Michigan Press.
- Baym, N. K. (2015). *Personal connections in the digital age*. John Wiley & Sons.
- Bazin, A. (1967). *What is cinema? Volume 1*. University of California Press.
- Benson-Allott, C. (2013). *Killer tapes and shattered screens: Video spectatorship from VHS to file sharing*. University of California Press.
- Berger, J. (2008). *Ways of Seeing*. Penguin UK.
- Bridle, James. (2018). *New dark age: Technology and the end of the future*. Verso Press.
- Caplan, R., Hanson, L., and Donovan, J. (2018 Feb 21). Dead reckoning: Navigating content moderation after fake news. *Data & Society*. Retrieved from: <https://datasociety.net/output/dead-reckoning/>

- Code4lib (2018). About. Retrieved from: <https://code4lib.org/about/>
- Cole, S. (2017 December 11). AI-assisted fake porn is here and we're all fucked. *Vice*. Retrieved from: https://motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn
- Conger, K. (2018 June 21). Amazon workers demand Jeff Bezos cancel face recognition contracts with law enforcement. *Gizmodo* Retrieved from <https://gizmodo.com/amazon-workers-demand-jeff-bezos-cancel-face-recognition-1827037509>
- Conger, K. (2018 May 14). Google employees resign in protest against Pentagon contract. *Gizmodo* Retrieved from <https://gizmodo.com/google-employees-resign-in-protest-against-pentagon-con-1825729300>
- Cooke, N. A. (2017). Posttruth, truthiness, and alternative facts: Information behavior and critical information consumption for a new age. *The Library Quarterly*, 87(3), 211-221.
- Crawford, Kate and Joler, Vladan. Anatomy of an AI System (n.d.). [https://anatomyof.ai/Curtin, M., Holt, J. and Sanson, K. \(2014\). *Distribution revolution: Conversations about the digital future of film and television*. University of California Press. Department of Professional Employees AFL-CIO. \(2014\). *Women in STEM and the impact of guest worker visas: United States, 2014*. Retrieved from <https://dpeaflcio.org/programs-publications/issue-fact-sheets/women-in-stem/>](https://anatomyof.ai/Curtin, M., Holt, J. and Sanson, K. (2014). Distribution revolution: Conversations about the digital future of film and television. University of California Press. Department of Professional Employees AFL-CIO. (2014). Women in STEM and the impact of guest worker visas: United States, 2014. Retrieved from https://dpeaflcio.org/programs-publications/issue-fact-sheets/women-in-stem/)
- Drabinski, E. (2018, February 21). Are libraries neutral? [Blog post]. Retrieved from <http://www.emilydrabinski.com/are-libraries-neutral/>
- Dykan, P and Iatensko, A. (2018). Euromaidan Event Reconstruction. *SITU* Retrieved from <https://situ.nyc/research/projects/euromaidan-event-reconstruction>
- Eberle, S. G. (2009). Exploring the Uncanny Valley to Find the Edge of Play. *American Journal of Play*, 2(2), 167-194.
- Farid, H. (2006). Digital doctoring: How to tell the real from the fake. *Significance*, 3(4), 162-166.
- Fineman, M. (2012). *Faking it: Manipulated photography before Photoshop*. Metropolitan Museum of Art.
- Forensic Architecture (2018, September 28). Project. *Forensic Architecture* Retrieved from <https://www.forensic-architecture.org/project/>
- Frenkel, S. (2018, June 19). Microsoft employees protest work with ICE, as tech industry mobilizes over immigration. *The New York Times* Retrieved from <https://www.nytimes.com/2018/06/19/technology/tech-companies-immigration-border.html>
- Friedl, J. (2006, February 21). Online exif (image data) viewer [Blog post] Retrieved from <http://regex.info/blog/other-writings/online-exif-image-data-viewer>
- Friedman, A. (2014, October 1). Etsy's trying to fix tech's women problem. Why aren't you? *Medium* Retrieved from <https://medium.com/matter/this-is-the-last-thing-youll-ever-need-to-read-about-sexism-in-tech-56b9a3a77af0>
- Goulding, A. (2004). Libraries and social capital. *Journal of Library and Information Science*, 36(1), 3-6.
- Greene, D (2018, February 13). We don't need new laws for faked videos, we already have them. *Electronic Frontier Foundation* Retrieved from <https://www.eff.org/deeplinks/2018/02/we-dont-need-new-laws-faked-videos-we-already-have-them>
- Haraway, D. J. (1985). *A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s* (pp. 173-204). San Francisco, CA: Center for Social Research and Education.
- Heikal, W. (2018 February 2). How Syrian Archive use open source investigation to support accountability in Syria. *Medium*. Retrieved from <https://medium.com/meedan-updates/how-syrian-archive-use-open-source-investigation-to-support-accountability-in-syria-5a39ecfe3db8>
- Ivory, J. D. (2006). Still a man's game: Gender representation in online reviews of video games. *Mass Communication & Society*, 9(1), 103-114.
- Jenkins, H., Ito, M., & Boyd, D. (2016). *Participatory culture in a networked era*. New York: Polity.
- Keilty, P. (2012, February). Embodiment and desire in browsing online pornography. In *Proceedings of the 2012 iConference* (pp. 41-47). ACM.
- Kumbier, A., & Starkey, J. (2016). Access Is Not Problem Solving: Disability Justice and Libraries. *Library Trends*, 64(3), 468-491.
- Leo, V. (1996). Born under a visual sign. *Afterimage*, 23(4), 5.
- McGowan, D. (2017). Nicolas Cage—good or bad? Stardom, performance, and memes in the age of the Internet. *Celebrity Studies*, 8(2), 209-227.
- Meyer, E. & Wachter-Boettcher, S. (2016). *Design for real life*. New York: A Book Apart.
- Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33-35.
- Mulvey, L. (1975). Visual pleasure and narrative cinema. *Screen*, 16(3): 6-18.
- Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental brain research*, 235(3), 889-901.
- Nakamura, L. (2013). *Cybertypes: Race, ethnicity, and identity on the internet*. Routledge.
- Nick Cage Deep Fakes (2018, January 29). Nic Cage deepfakes mini compilation [Video file]. Retrieved from: <https://www.youtube.com/watch?v=2jp4M1cIJ5A&frags=pl%2Cwn>
- Nick Cage Deep Faeks (2018, January 29). Realy bad(?) NIC CAGE deepfake (face off) (deepfakes) [Video file]. Retrieved from: https://www.youtube.com/watch?v=_i8ecYD1V64
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Panofsky, E. (1982). *Meaning in the visual arts: Papers in and on art history*. The University of Chicago Press.
- Patterson, Z. (2004) Going on-line: Consuming pornography in the digital era. In L. Williams. (Ed.) *Porn studies*.(pp. 104-123) Duke University Press.

- Pichai (2018) <https://www.blog.google/technology/ai/ai-principles/>
- Pornhub Insights (2018, January 9). 2017 year in review. *Pornhub Insights*. Retrieved from <https://www.pornhub.com/insights/2017-year-in-review>
- Roettgers, J. (2018, August). Naughty america wants people monetize deepfake porn. *Variety*. Retrieved from <https://variety.com/2018/digital/news/deepfake-porn-custom-clips-naughty-america-1202910584/>
- Soha, M., & McDowell, Z. J. (2016). Monetizing a Meme: YouTube, Content ID, and the Harlem Shake. *Social Media+ Society*, 2(1), DOI: 2056305115623801.
- Shatford, S. (1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging & classification quarterly*, 6(3), 39-62.
- Sharman, J. (2018, February 7). Pornhub and twitter ban AI-generated 'deepfakes' videos that put female celebrities' faces on adult actresses' bodies. *The Independent*. Retrieved from <https://www.independent.co.uk/life-style/gadgets-and-tech/pornhub-twitter-deepfakes-ban-ai-celebrity-faces-porn-actress-bodies-emma-watson-jennifer-lawrence-a8199131.html>
- Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12), 1349-1380.
- U.S. Reddit user distribution 2016, by gender. In Statista - The Statistics Portal. Retrieved December 28, 2018 from <https://www.statista.com/statistics/517155/reddit-user-distribution-usa-gender/>
- Van Doorn, N. (2010). Keeping it real: User-generated pornography, gender reification, and visual pleasure. *Convergence*, 16(4), 411-430.
- Wagner, J. (2015, August 16) Forensically, photo forensics for the web [Blog post]. Retrieved from <https://29a.ch/2015/08/16/forensically-photo-forensics-for-the-web>
- Weise, T., Li, H., Van Gool, L., & Pauly, M. (2009, August). Face/off: Live facial puppetry. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer animation*(pp. 7-16). ACM.
- Williams, L. (1999). *Hard core: Power, pleasure, and the "frenzy of the visible"*. University of California Press.
- WITNESS. (2015) Filming long after a crime [Blog post]. Retrieved from https://vae.witness.org/portfolio_page/filming-long-after-a-crime/
- WITNESS. (2016) The role of video in the criminal justice process [Blog post]. Retrieved from https://vae.witness.org/portfolio_page/role-of-video-in-the-criminal-justice-process/