



INSTYTUT
KULTURY
POLSKIEJ



View. Theories and Practices of Visual Culture

title:

Images Out of Sight: Some Remarks on Moderating Online Content

author:

Miłosz Markiewicz

source:

View. Theories and Practices of Visual Culture 21 (2018)

URL:

<http://pismowidok.org/index.php/one/article/view/557/1213>

publisher:

View. Foundation for Visual Culture

affiliation:

Institute of Literary Research of the Polish Academy of Sciences
Institute of Polish Culture, University of Warsaw

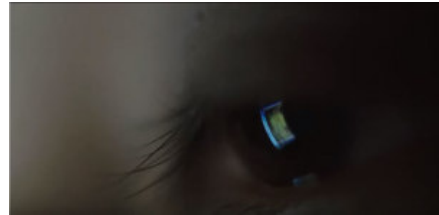
Miłosz Markiewicz

Images Out of Sight: Some Remarks on Moderating Online

Content

Translated by Jan Szelaqiewicz

When, in the fall of 2010, Safiya Umoja Noble set out to browse the Web for something that her stepdaughter might find interesting, she had no idea she was about to open Pandora's box. She recounted: "In just a few minutes while searching on the web, I experienced the perfect storm of insult and injury that I could not turn away from."¹ The first result returned by Google for the query "black girls" was the adult site HotBlackPussy.com. This encouraged the scholar to take a closer look at how the algorithms of the world's biggest search engine actually operate and why the avenues of representation offered by the Silicon Valley tech giant provide such dubious information about people of color. When, after two years of research, she realized that the results returned for the query still hadn't changed, Noble published her acclaimed piece in *Bitch* magazine.² It was only in the wake of its publication that the particular area of the Web related to the search results began to change, and although it's unlikely that the shift directly followed from her efforts, it's undeniable that she managed to draw public attention to a considerable problem and thus made it impossible for the corporation to continue its ambivalent stance towards it. Nicknamed Panda, the overhaul of Google's search algorithms effectively struck pornographic content from the first pages of search results for the phrase "black girls," but ultimately the change was superficial – similar websites continued to populate the search results returned for phrases such as "Latina girls" or "Asian girls." Furthermore, the image results for queries like "gorillas" continued to feature pictures of black people. Noble later argued that "on the Internet and in our everyday uses of technology, discrimination is also embedded in computer code



Hans Block, Moritz Rieseewieck, *The Cleaners*, 2018.

and, increasingly, in artificial intelligence technologies that we are reliant on, by choice or not.”³ Soon after, it was reported that Google’s facial recognition software automatically appended photographs featuring African Americans with tags such as “apes” or “animals.” Similar examples abound, with the most prominent of all occurring in 2015, when searching the keywords “nigga house” or “nigger king” in Google Maps returned the White House, occupied at the time by Barack Obama, as the top result.⁴

In her latest book, *Algorithms of Oppression: How Search Engines Reinforce Racism*, Noble argues for stricter control of the online world,⁵ bringing up a plethora of examples that reveal how racism, sexism, and other forms of discrimination are deeply entrenched in search engine algorithms. Unawareness of their existence, combined with our predilection for unquestioned obeisance to search engine interfaces, may ultimately contribute to the deepening of societal divisions. In this particular context, it bears mentioning the category of structural racism, used by Reni Eddo-Lodge, among others – which, at its core, is merely “the survival strategy of systemic power.”⁶ And that power is subordinate to the figure of the white male, the influence at its disposal inversely proportional to its visibility. Both Noble and Eddo-Lodge are academically focused on intersectionality – they seek linkages between different forms of discrimination and exclusion, then bring them to light in order to draw attention to the oppressive nature of the system we live in. Both also believe in the power of grassroots organizing. Thus, Noble’s postulate for more community control over the Web and throwing more weight behind non-commercial search engines, whose results are supposedly free from the financial considerations of interested commercial parties. But what seems particularly interesting to me in the context of the algorithms of oppression described by Noble is not the call for increased control over online content, as much as the public response from Google to the accusations leveled against it after the White House incident: “Our teams are working to fix this issue quickly.”⁷ The question remains – who, then, is responsible for “cleaning up messes” online?

One cannot but agree with Scott Galloway’s contention that “digital space needs rules.”⁸ Its widespread availability should entail some sense of security for those

who traverse its nearly endless expanse. The utopian vision of virtual reality holds no place for violence or racism. "The truth is, we wish platforms could moderate away the offensive and the cruel,"⁹ adds Tarleton Gillespie. But how does such moderation differ from censorship? Could we ourselves take on the task of policing online content? Who would decide what goes and what gets to stay? And finally: who would police those policing "online purity"? After all, people have differing views, hold different values, and come from specific socio-cultural contexts, and thus expecting uniform objectivity from people with such diverse backgrounds moderating online content is not only unfeasible, but can produce unpredictable political repercussions:

Content moderation is hard. This should be obvious, but it is easily forgotten. Moderation is hard because it is resource intensive and relentless; because it requires making difficult and often untenable distinctions; because it is wholly unclear what the standards should be; and because one failure can incur enough public outrage to overshadow a million quiet successes.¹⁰

Conversely, the algorithms the Internet now more or less runs on cannot bear responsibility for the results of their implementation. They cannot be punished, but only modified, as they were in Google's Panda update. Each subsequent update implements some new rules and patches up the underlying machine learning processes, allowing the algorithms (whose work affects us, although we're not necessarily aware of it) to improve their performance along pre-arranged trajectories. Algorithms gain insight into us, into themselves, and into each other – recognizing behavioral and decision-making patterns that transcend human perception, analyzing them, following decision trees down, and diving into the rulesets that the reality they're monitoring is based on.¹¹ Thus, they seem much better suited than us to the task of patrolling the vast expanses of the Internet. They're capable of processing myriad datasets and communicating with each other, but above all, they're equipped with clearly defined goals from which nothing can distract them. Their performance is based on so-called weak AI (also called narrow AI). This means that their perspective is truncated and tailored to the undertaking

of a specific task, which, in turn, makes them focused on performing that task to the best of the abilities their creators imbued them with (and, even if they use machine learning, their exploitation thereof is limited by their design).¹² As Yuval Noah Harari points out, “The seed algorithm may initially be developed by humans, but as it grows it follows its own path, going where no human has gone before – and where no human can follow.”¹³

It’s not hard to spot the issues that “ubiquitous” algorithms often get ensnared in.¹⁴ Most of them are probably rooted in human imperfection, which, by its very nature, is absent from machine code; the latter continuously strives for perfection, thanks at the very least to the aforementioned machine learning capabilities. Our conduct, however, is chaotic enough that it’s difficult to predict all of its vicissitudes. And since algorithms are designed to scrutinize and assimilate human behaviors, they also pick up the morally questionable ones. Their identification and repetition of patterns is not an act of judgment in and of itself, but rather an action they’re tasked with performing. This is one reason why those algorithms saddled with moderating content in social media ultimately trap us in our own informational echo chambers or facilitate the dissemination of fake news. As argued by Siva Vaidhyanathan, Facebook’s algorithms are deliberately designed to stack our newsfeed with content eliciting positive reactions, such as likes or shares:

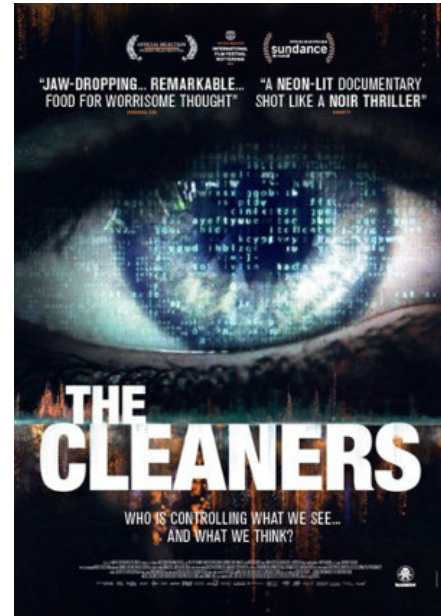
Every aspect of its design is meant to draw us back [...] Facebook has developed techniques that it hopes measure our relative happiness or satisfaction with aspects of the service. Facebook researchers have been trying to identify and thus maximize exposure to the things that push us to be happier and minimize exposure to things that cause us anxiety or unhappiness. Much of what we encounter, including many of our responses and interactions, leads us to be sad, frustrated, angry, and exhausted.¹⁵

The community control over digital spaces that Noble called for would necessarily entail the task of policing and hiding content that is deemed socially impermissible. Although the Internet should obviously be free of content that is prohibited in other public spaces, we need to remember that the policing will often be performed by

Thousands of people, chiefly in the Philippines, spend hours every day browsing visuals (images, video clips, memes, graphics, etc.) and deciding whether they will remain posted or be removed.

Their work – invisible, but producing tangible results – is explored by Eva and Franco Mattes in their installation *Dark Content*.²¹ After interviewing hundreds of content moderators, the duo fed their responses to digital avatars (with computers generating both their appearance and their voices). Within the installation space, the interviews were screened on displays placed on IKEA office furniture, but deliberately arranged to make viewing an uncomfortable experience, mentally and somatically.²² The toil of content moderators is also the subject of *The Cleaners*, a documentary by filmmakers Hans Block and Moritz Riesewieck.²³ Every

day, thousands of people are tasked with deciding whether a given piece of content remains on feeds across Facebook or Instagram, and the film tackles the back-end of the job as well as its consequences, which can affect high-ranking politicians, artists, and run-of-the-mill users. In the course of an eight-hour workday, moderators have to look over at least 25,000 visuals (their daily quota), which boils down to around one picture per second. They watch visuals that we'd never want featured on our newsfeeds, such as child pornography, rape, or summary executions performed by terrorists. As such, their daily work entails precisely what Safiya Noble or Tarleton Gillespie have been postulating – the community policing of online spaces. Most of us believe that visuals like these can only be found on what we've come to call the Deep Web or the darknet, and as such are concealed from the average user, accessible only to those "in the know" – and this is indeed so, mostly thanks to "the cleaners" and their tireless efforts to scrub such content before it makes its way into the feeds of the general public. Their labor thus remains simultaneously unseen (there is a reason why tech giants outsource most content moderation overseas,



Hans Block, Moritz Riesewieck, *The Cleaners*, 2018.

and, increasingly, in artificial intelligence technologies that we are reliant on, by choice or not.”³ Soon after, it was reported that Google’s facial recognition software automatically appended photographs featuring African Americans with tags such as “apes” or “animals.” Similar examples abound, with the most prominent of all occurring in 2015, when searching the keywords “nigga house” or “nigger king” in Google Maps returned the White House, occupied at the time by Barack Obama, as the top result.⁴

In her latest book, *Algorithms of Oppression: How Search Engines Reinforce Racism*, Noble argues for stricter control of the online world,⁵ bringing up a plethora of examples that reveal how racism, sexism, and other forms of discrimination are deeply entrenched in search engine algorithms. Unawareness of their existence, combined with our predilection for unquestioned obeisance to search engine interfaces, may ultimately contribute to the deepening of societal divisions. In this particular context, it bears mentioning the category of structural racism, used by Reni Eddo-Lodge, among others – which, at its core, is merely “the survival strategy of systemic power.”⁶ And that power is subordinate to the figure of the white male, the influence at its disposal inversely proportional to its visibility. Both Noble and Eddo-Lodge are academically focused on intersectionality – they seek linkages between different forms of discrimination and exclusion, then bring them to light in order to draw attention to the oppressive nature of the system we live in. Both also believe in the power of grassroots organizing. Thus, Noble’s postulate for more community control over the Web and throwing more weight behind non-commercial search engines, whose results are supposedly free from the financial considerations of interested commercial parties. But what seems particularly interesting to me in the context of the algorithms of oppression described by Noble is not the call for increased control over online content, as much as the public response from Google to the accusations leveled against it after the White House incident: “Our teams are working to fix this issue quickly.”⁷ The question remains – who, then, is responsible for “cleaning up messes” online?

One cannot but agree with Scott Galloway’s contention that “digital space needs rules.”⁸ Its widespread availability should entail some sense of security for those

while striving to preserve the appearance that it is algorithms that keep their websites free of offensive content)²⁴ and perfectly visible. And the effects of their work seem to be subject to the classic aporia of cleaning – juxtaposing the visible (the clean space left afterwards) and the invisible (that which was removed and hidden from view).²⁵ Visually, the Internet seems a relatively safe space; the chance of running into photos of dead bodies or footage of brutal violence is rather low, mostly because someone has already made the decision to keep such content away from our eyes.

In both *Dark Content* and *The Cleaners*, online content moderators talk in detail about their jobs, and offer us a closer look at the situations and images that had the most profound effect on them. We listen to people talk about watching footage depicting mass rapes, school molestation, teenagers livestreaming their suicides, or of seeing terrorists execute people. We also learn about governments “purging” online outlets of images of civil protests or photos depicting specific individuals. Anyone can approach online moderator shops with a job – and the firms end up cleaning up inconvenient facts, hiding unwanted individuals, and deleting leaked confidential documents.

But should the manipulation involved in managing online images be in any way surprising to us? Is it not an immanent characteristic of similar endeavors? To quote Piotr Zawojcki, “the ‘novelty’ of new images lies in doubts raised by radically novel ontology as to whether they can still be considered images, or merely symptoms and effects of algorithmic processes.”²⁶ As indicated by its very designation, the digital image is essentially code, but clad in a visual form. With the requisite tools and background, we too could see the digits underlying the image, akin to Neo in the iconic scene from *The Matrix* where he begins to see the physical world around him as cascades of green-tinted code. After all, as Mitchell argues, “the essential characteristic of digital information is that it can be manipulated easily and very rapidly by computer. It is simply a matter of substituting new digits for old.”²⁷ Thus, the essence of such an image would be its mutability – the ontological fluidity – resulting in a more relational than essentialist reading of the image itself. This transmutative capacity underpins the very existence of the digital

image. But can it also incorporate a peculiar suspension between existence and non-existence – or vanishing, to put it more precisely? What is the ontological status of the removed, and thus invisible, image?

Content moderators often say that they “take out the Internet’s trash.” The metaphor is particularly apt in the case of the story of one of the nameless characters featured in the abovementioned documentary, who admits that she always feared she would end up working at a waste dump. The audience – with the character’s comparison of their efforts to clean-up work still echoing in their minds – thus receives the suggestion that maybe the girl landed exactly where she had hoped not to. This, in turn, brings to mind the “poor images” described by Hito Steyerl, akin to the “contemporary Wretched of the Screen, the debris of audiovisual production, the trash that washes up on the digital economies’ shores.”²⁸ On the one hand, we can ask ourselves whether a discarded image is trash – after all, if an object isn’t there, does it become something? On the other hand, however, that seems to be the essence of refuse: that it is not there, it is invisible. Stripped of any aesthetic veneer, images do not exactly disappear, but are instead moved to our collective blind spot.²⁹ But in this particular context, the situation of the digital image is somewhat different – it is not made invisible so much as it disappears. It is beyond question, however, that, like the images mentioned by Steyerl, it is subject to the “alternative economy of images.”³⁰

But can we actually call an image expunged from the Web non-existent? It cannot, after all, be deleted completely; even if we scrub it from digital media, wipe the code, and replace the old digits with new ones – reprogramming the online source – then the data that “the cleaners” work to remove always gets stored somewhere, and that location cannot be “cleaned up.” As Hans Belting writes:

The human being is the natural *locus of images*, a living organ for images, as it were. Notwithstanding all the devices we use today to send and store images, it is within the human being, and only within the human being, that images are received and interpreted in a living sense; that is to say, in a sense that is ever changing and difficult to control no matter how

forcefully our machines might seek to enforce certain norms.³¹

Although the work of online content moderators mostly entails the removal of data, a considerable portion of that data may end up seeping into their own memories – which is why both *Dark Content* and *The Cleaners* feature Filipino moderators speaking at length about what they have seen and what we never will. This reveals, in a very tangible, even embodied way, the full extent of the digital image. Although expunged from online spaces, it persists in memory. The film's directors recount:

Theirs is a key job in the digital era. We were surprised that, rather than complain about their working conditions and the attendant emotional exhaustion, they continued to take pride in their work. One of the interviewees asked us to imagine what the Internet would look like if he stepped away from his computer for just two hours. "You'd probably end up seeing things you never wanted to see. And that's what people like me shield you from," he argued. And he's right. Imagine that a moderator deletes a video with child pornography, and a couple of weeks later their boss drops by, saying that the police caught the person who uploaded it. This sort of narrative, framing extremely draining labor as necessary and deeply meaningful, is what helps them survive. Regardless, most of these companies have a very high turnover rate, with moderators staying on for no longer than six months.³²

The work of content moderators mostly entails verifying whether what they're looking at will not be offensive to the general public. What they have to endure in order to do so, however, cannot simply be scrubbed from their heads – thus making them the last human recipients of "banned" content, keeping these illicit visuals under their eyelids, capable of recalling them in the blink of an eye, but unable to communicate their contents to third parties in ways other than through words (albeit even that is mostly *verboten*, as their work is covered by non-disclosure agreements).³³ When the aforementioned interviewee says "that's what people like me shield you from," we would do well to read it quite literally. Viewed through such a lens, the moderator becomes a bodyguard, using their own body to take the blow

or bullet intended for their charge. Although we will never suffer the violence or savagery of the expunged visual, someone will nevertheless be “injured” by it, their flesh forever marked with the torrent of savagery they had to endure.

One of the moderators featured in *The Cleaners* brings up a co-worker who killed himself because he could no longer bear the mental drain of having to watch the offensive images day in, day out. In this particular context, human moderators emerge as weaker or more vulnerable than their algorithmic counterparts; embodiment turns out to be more curse than blessing.



Hans Block, Moritz Riesebeck, *The Cleaners*, 2018.

From this perspective, man seems a crucial element in the ontology of the expunged digital image, as his body retains it, precluding it from disappearing altogether and demonstrating that no image ever exists only virtually, without physical consequences. It is rather an element of a part-virtual, part-physical mesh of relationships. Scrubbing it from one space will not erase it from existence, because it will continue to survive in another. Its existence and non-existence are inextricably entwined. The ontology of a deleted digital image transcends the dualisms that we usually think back to when dealing with networks, algorithms, the virtual, and the real. The long-established lines appear uncertain, vague. In the words of Karen Barad, such “indeterminacy is key not only to the existence of matter but also to its non-existence, or rather, it is the key to the play of non/existence.”³⁴

In this context it bears noting, as Siva Vaidhyanathan did in his analysis of the removal of Nick Ut’s *The Terror of War* from social media, that “lost in the uproar was the fact that Facebook did not remove the image from the world. It still existed in books, on hundreds of websites searchable by Google, and in newspapers and magazines.”³⁵ We’ve become used to seeing the ontology of a digital image – all ontologies, really – in a one-dimensional way, from the perspective of the entity and its existence, whereas its removal quickly reveals the complex nature of the relationships that lie beneath it.³⁶

Lambert Wiesing argues that the computer screen is supposed to “display and

present something.”³⁷ But what happens when said image is missing? Even the word “missing” seems a misnomer in this case. Deleted images are not missing, they are un-present. Their presence is founded upon visual absence – and vice versa. It is because they have been expunged that the online landscape looks as it does. Although images are supposed to present reality to us, it is precisely for that reason that some of them have to be neutralized, because, as it turns out, some of them should never be seen. The expunged images unmask the aporias underpinning the ontology of the digital image. Deleting something does not snuff it out of existence – it continues to function in myriad ways, both in the minds of content moderators and the people who created those images, as well as the memory banks of the computers and smartphones that they passed through. First and foremost, however, they are present in their absence – their very existence requires the policing of the spaces they function in. Thus, expunged images operate through lack. Their existence straddles the intersection of the past, the present, and the future – they are created and upload online, to be deleted by moderators (the past), scrubbed from the landscape of the Web (the present), and thus prevented from disseminating the messages ingrained in them – propagating brutality, violence, or any other behavior deemed socially unacceptable; as such, they play a significant role in driving social attitudes (the future). Thus, we ought to give credit to Wiesing, because even the absent image makes the computer screen present something, if only by stressing its absence.

The “vanishing” of the digital image can be considered an essential quality thereof, as the compression necessary for uploading, file format conversion, and even uploading itself often result in its degradation.³⁸ In this context, the labor of content moderators, entailing the expungement of such images, would not stand in contrast to the basic mode of the images’ existence online. Neither would it breach the ontology of the digital image – broad enough to hold noise, error, degradation, and even vanishing.

Is it possible, then, for the liquid ontology of the digital image, based on binary code and algorithmic processes, to incorporate the work of content moderators? If their work does not contravene the aforementioned ontology (on the contrary, the two

seem to align), but simultaneously involves tampering with the very existence of the images described herein, then the efforts of “the cleaners” should be given their proper place in the discourse on the digital ontology of visibility, as it unmask the relationship of power tied to visibility in online spaces.³⁹ Anonymous workers from Manila, relying on nothing but a manual supplied to them by their employer, are by no means the subjects of power, but merely its objects, tasked with enforcing a range of systemic preventive measures. After all, it is their work that reinforces the “secrecy and autonomy in the exercise of the power to punish,”⁴⁰ a notion emphasized in *Dark Content* by having the words of the moderators come from the mouths of digital avatars. The persons whose images were scrubbed from the Internet have no avenue to appeal the decision, and must reconcile themselves to the act of censorship. One particularly cautionary tale can be found in the story of Illma Gore, an LA-based artist who one day uploaded an illustration of a poorly endowed Donald Trump, titled *Make America Great Again*, to her Facebook profile. The picture was quickly flagged as offensive, and forwarded to a content moderator tasked with deciding whether it would remain on Gore’s feed or be removed. Without enough time to ponder the intricacies of freedom of speech in America, but apprised of the punishment that insulting the president of the Philippines might carry, she ultimately decided to remove the illustration, which then led to Gore’s entire account being shut down.

Situations like these are unavoidable when we saddle thousands of laborers with the task of sifting through at least 25,000 images for eight hours a day, working in shifts so that their employers can provide round-the-clock services. The custodians cannot be supervised in any way, nor would such oversight be welcomed. According to their work contracts, moderators making more than three mistakes per month are let go, but staff turnover caused by mental exhaustion is so high that most employers cannot allow themselves too many layoffs.

In a manner, therefore, the work of online moderators can be seen as a way to discipline the users of the Internet. It demonstrates that the relationship of power and incessant, clandestine surveillance is an inherent element of the ontology of the digital image – particularly the image uploaded online. Censorship is the disciplining

punishment designed to teach us what content is acceptable and what isn't. Borrowing from Roy Ascott, we might say that the Internet is a "site of interactivity and connectivity in which the viewer can play an active part in the transformation or affirmation of the images."⁴¹ The problem is, however, that broadcasters and their intended audience are separated nowadays by a particular filter, in the form of a "clandestinely viewing subject," tasked with the aforementioned affirmation of the image.

The labor of online cleaners, however, remains invisible to the majority of Internet users, probably because it is covered by the transparency of the interface, described at length by media theorist Lev Manovich.⁴² This invisibility shrouds the work of those custodians in another layer of anonymity. Not only do we not know who they are, but we should not be aware of their existence at all. That is why no one is really rectifying the misleading information about the extent of algorithm use for the purposes of content moderation, and why – aside from obvious economic reasons, of course – most of this "custodial" work is actually performed by Filipinos. They are supposed to remain not merely anonymous, but also invisible, leaving society at large unaware of the hidden mechanisms underpinning the Web (both the oppressive search algorithms described by Noble and the work of online content moderators). When using the Internet, we see neither the interface, the code, nor the underlying structure that disciplines our own movements and presence in online spaces. To quote Lambert Wiesing, we look "through a transparent plane without thematizing the transparent plane itself."⁴³ We should come to realize, therefore, that every image uploaded online actually contains much more information than is visible to the eye – and that those images concealed from our view usually have the most to say.

What does it mean to "keep the Web clean"? How can we genuinely judge which piece of content is clean and which isn't? Although value systems tend to depend on context, it seems obvious that content that is violent or calling for violence, brutal, or offensive, should be unambiguously treated as undesirable. Is it actually

possible to effectively police and cultivate online spaces? If so, who ultimately gets to decide what stays and what goes? The secrecy and systemic autonomy of the relationship of power that underpins content policing makes it relatively easy to put off any sort of reflection over its impact on the democratic nature of the Web, especially when, in return, it offers us a relatively safe space online, scrubbed of any violent or unpleasant content. But doesn't the content policing postulated by Safiya Noble, currently enforced by Silicon Valley tech giants via the hands of "online custodians," carry the risk of manipulating the image of reality? After all, the manipulation of the digital image often produces tangible results in reality. Scrubbing unwanted visuals from our collective field of view will not make the phenomena they either represent or stem from disappear; instead, they will become un-present, and thus be imbued with far greater power than they ever commanded while "out in full view."

At the same time, over-reliance on the ability to control online content – with the policing itself remaining beyond any control – ultimately implies acquiescence to a range of self-disciplining measures, most of which require the user to submit to whatever politics of visibility are currently being enforced. After all, how can we be sure whether the images we upload will not be flagged for impropriety and sent to a moderator? Such control is ubiquitous, and each image can ultimately be subject to the critical gaze of the online custodian. Thus, the Internet loses its democratic luster of a space where all voices have equal power. In this new regime, power rests in the hands of those who control the "delete" button.

Footnotes

1 Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018), 25.

2 See: Safiya Umoja Noble, "Missed Connections: What Search Engines Say About Women," *Bitch* 12, no. 4 (2012), 37-41.

- 3 Noble, *Algorithms of Oppression*, 21.
- 4 See: Andrew Griffin, "GOOGLE APOLOGISES FOR RACIST 'N*GGA HOUSE' SEARCH THAT TAKES USERS TO WHITE HOUSE," *The Independent*, May 21, 2015, www.independent.co.uk/life-style/gadgets-and-tech/news/nigga-house-racist-search-term-forces-google-to-apologise-for-taking-users-to-white-house-10266388.html (accessed March 7, 2019).
- 5 See: Noble, *Algorithms of Oppression*.
- 6 Reni Eddo-Lodge, *Why I'm No Longer Talking to White People About Race* (London: Bloomsbury, 2017), 64.
- 7 Noble, *Algorithms of Oppression*, 31.
- 8 Scott Galloway, *The Four: The Hidden DNA of Amazon, Apple, Facebook, and Google* (New York: Penguin, 2017), 119.
- 9 Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven and London: Yale University Press, 2018), 212.
- 10 Gillespie, *Custodians of the Internet*, 9-10.
- 11 Cf. John D. Kelleher, Brian Mac Namee, and Aoife D'Arcy, *Fundamentals of Machine Learning for Predictive Data Analytics* (Cambridge, MA and London: MIT Press, 2015).
- 12 Cf. Malcolm Frank, Paul Roehrig, and Ben Pring, *What To Do When Machines Do Everything: How to Get Ahead in a World of AI, Algorithms, Bots, and Big Data* (New Jersey: Wiley, 2017), 92-95.
- 13 Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (New York: HarperCollins, 2017), 398-399.
- 14 Terrence J. Sejnowski, *The Deep Learning Revolution* (Cambridge, MA and London: MIT Press, 2018), 207.

15 Siva Vaidhyathan, *Anti-Social Media: How Facebook Disconnects Us and Undermines Democracy* (New York: Oxford University Press, 2018), 33.

16 Cf. Alexandra Ma, "Facebook banned a user from posting a photo of a 30,000-year-old statue of a naked woman – and people are furious," *Business Insider*, March 1, 2018, www.businessinsider.com/facebook-bans-venus-of-willendorf-photos-over-nudity-policy-2018-3?IR=T (accessed March 7, 2019).

17 "Facebook usuwa obrazy Rubensa z powodu odsłoniętych piersi. Muzea protestują," *Polish Press Agency* online, www.polskieradio.pl/5/3/Artykul/2170631,Facebook-usuwa-obrazy-Rubensa-z-powodu-odslonietych-piersi-Muzea-protestuja (accessed March 7, 2019).

18 It bears noting that the human gaze works along similar lines: we also seek patterns or familiar shapes, and compare what we see against known visuals. The key difference, however, is that, as we grow our social and cultural faculties, we tend to rely more on context than anything else. Cf. Władysław Strzemiński, *Teoria widzenia* (Łódź: Muzeum Sztuki, 2016).

19 Galloway, *The Four*, 124.

20 Nicholas Mirzoeff, *How to See the World* (London: Pelican, 2015), 156-157.

21 The installation was part of the *First Person Plural: Empathy, Intimacy, Irony, and Anger* exhibition, held at the basis voor actuele kunst (BAK) in Utrecht, the Netherlands, from May 12 to July 22, 2018.

22 The universalization of office work, symbolized herein by the IKEA office furniture, is supposed to highlight the anonymity of content moderators (who thus become everymen), as well as their slave-like exploitation by the capitalist corporate machine. Just as control of the Web is based on a catalog of specific approaches, so office furniture that we use every day constrains our movements, limiting our range of motion to what its designers were able to envision. That is also why the recordings featured in *Dark Content* compel the audience to transcend their well-worn and practiced ways of using such furniture. Cf. Artur Szarecki,

Kapitalizm somatyczny. Ciało i władza w kulturze korporacyjnej (Warsaw: Wydawnictwa Drugie, 2017).

23 The film premiered in 2018 and was released in Poland at the *Millennium Docs Against Gravity Film Festival*. I'd like to thank the organizers for giving me access to the footage.

24 The filmmakers behind *The Cleaners* admitted that they, too, once believed that algorithms were responsible for content moderation: "In 2013, a video showing a woman being violent towards her child was uploaded to Facebook, and ultimately shared 60,000 times before it was scrubbed from the platform. Moritz and I were fascinated not just with the speed of its spread, but the swiftness of its deletion. When asked what sort of algorithm decided the removal and what its error rate is, Sara T. Roberts, a UCLA researcher specializing in social media, revealed to us that online content policing was performed not by machines, but primarily by hired hands hailing from the Global South"; Magda Roszkowska, "Czyszciciele internetu kontrolują to, co trafia do sieci. 'Oglądają wszelkiego rodzaju przemoc,'" *Gazeta.pl*, <http://weekend.gazeta.pl/weekend/1,152121,23538263,czyszciciele-internetu-kontroluja-to-co-trafia-do-sieci-ogladaja.html> (accessed March 7, 2019). The work done by Filipino content moderators was revealed for the first time in 2012 by Adrian Chen: see Adrian Chen, "Inside Facebook's Outsourced Anti-Porn and Gore Brigade, where 'Camel Toes' are More Offensive than 'Crushed Heads,'" *Gawker*, February 16, 2012, <https://gawker.com/5885714/inside-facebooks-outsourced-anti-porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads> (accessed March 7, 2019). See also: Gillespie, *Custodians of the Internet*.

25 This was the subject of *Guests*, Krzysztof Wodiczko's exhibition held at the Polish Pavilion at the 53rd Venice Biennale in 2009.

26 Piotr Zawojski, "Paradoksy obrazów w epoce nowych mediów," in *Sztuka obrazu i obrazowania w epoce nowych mediów* (Warsaw: Oficyna Naukowa, 2012), 11.

- 27 William John Mitchell, *The Reconfigured Eye: Visual Truth in the Post-Photographic Era* (Cambridge, MA and London: MIT Press, 1992), 6.
- 28 Hito Steyerl, "In Defense of the Poor Image," *e-flux* 10 (November 2009), <https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/> (accessed March 30, 2019).
- 29 Cf. Wolfgang Iser, "Ästhetik und Anästhetik," in *Ästhetisches Denken* (Stuttgart: Philipp Reclam jun.: 1993), 9–40.
- 30 Steyerl, "In Defense of the Poor Image."
- 31 Hans Belting, *An Anthropology of Images: Picture, Medium, Body* (Princeton: Princeton University Press, 2011), 37.
- 32 Roszkowska, "Czyściciele internetu."
- 33 This anonymity grants them the status traditionally reserved for state witnesses or secret operatives. The former narrative can be found in *Dark Content*, wherein moderators are given virtual avatars to hide behind, thus shielding them from the potentially harmful consequences of their custodial work. *The Cleaners*, on the other hand, brings up associations with covert operatives striving to save the world, or a secret brotherhood working behind the scenes to guard a sacred order. The smothering, secretive atmosphere of the narrative, blurry images, and darkened frames may bring to mind stories of people forced into hiding. Fragments of e-mails appearing on the screen, meanwhile, seem to be a trope taken straight from a spy movie.
- 34 Karen Barad, *What Is the Measure of Nothingness? Infinity, Virtuality, Justice* (Ostfildern: Hatje Cantz, 2012), 13.
- 35 Vaidyanathan, *Anti-Social Media*, 46.
- 36 Nick Ut's photo is also featured in *The Cleaners*. One of the interviewees looks at it, and says that although he's aware of the photo's history, the guidelines he has to adhere to would compel him to remove it, and that he would not try to take the

matter up with his superiors.

37 Lambert Wiesing, *Artificial Presence: Philosophical Studies in Image Theory* (Stanford: Stanford University Press, 2010), 23.

38 Jakub Dziewit wrote: “the JPEG file format [...] reduced the ‘weight’ of the photograph through means of compressing, which ultimately resulted in loss of quality. While developing the format, its creators exploited the fact that pictures displayed on a computer screen do not require a particularly high ‘density’ of points comprising them. Jakub Dziewit, *Aparaty i obrazy. W stronę kulturowej historii fotografii* (Katowice: grupakulturalna.pl, 2014), 212. Cf. Lev Manovich, “The Paradoxes of Digital Photography,” http://manovich.net/content/04-projects/004-paradoxes-of-digital-photography/02_article_1994.pdf (accessed March 31, 2019).

39 Images like these are always constituted collectively, and never exist “in and of themselves.” Their existence, and in this case un-presence, are the product of the intersection of different market, aesthetic, political, social, and communal webs (evinced, at the very least, in the notion of community control over the Internet). Although a more in-depth look at these issues lies outside the scope of this essay, it bears enquiring, in the spirit of Bruno Latour, whether digital images (and maybe even images in general) – particularly those functioning online – only come into existence as long as they answer the needs of – broadly defined – interests or some specific community concerns.

40 Michel Foucault, *Discipline and Punish: The Birth of the Prison* (New York: Vintage Books, 1995), 129.

41 Roy Ascott, “Photography at the Interface,” in *Telematic Embrace: Visionary Theories of Art, Technology, and Consciousness*, ed. Edward A. Shanken (Berkeley: University of California Press, 2003), 252.

42 See: Lev Manovich, *The Language of New Media* (Cambridge, MA: MIT Press, 2001).

43 Wiesing, *Artificial Presence*, 80.