Napalm Girl as Data Set. On the Reappraisal of Iconic Images by Automated Filtering Technologies

Katja Müller-Helle

This text was automatically generated on March 24, 2025.

Napalm Girl as Data Set.
On the Reappraisal of Iconic Images by Automated
Filtering Technologies

Katja Müller-Helle

Prologue

1　In "Kamera in Aufsicht," a review of Michael Klier's film *The Giant* (Federal Republic of Germany, 1983), Harun Farocki opens with a description of a split-screen image: the left-hand side shows a laboratory situation, a person sleeping peacefully with "ruffled blanket" and eyes closed (Fig. 1).[1] The right-hand side is entirely taken up with "readings from a sleep study"—i.e., with polysomnography; the sleeping subject's data are being transformed into visual information on a paper readout. This split-screen image shows two materializations of the same thing: one is an iconic visualization of the sleeping situation in video stills from a surveillance camera, the other an abstract visualization of internal physiological processes: heartbeat, respiration, and breathing effort as measured by a sleep electrocephalograph and other recording instruments. According to Farocki, this datafication of sleep represents the realization of one of mankind's childhood dreams: "the ability to stick a plug in someone's head and extract all the words, stories, and dreams."[2] A related childhood dream is the mass harvesting, transfer, and filtering of this personal image data as guidance for systems of control.

1. Photogram from Michael Klier's film *The Giant,* 1983
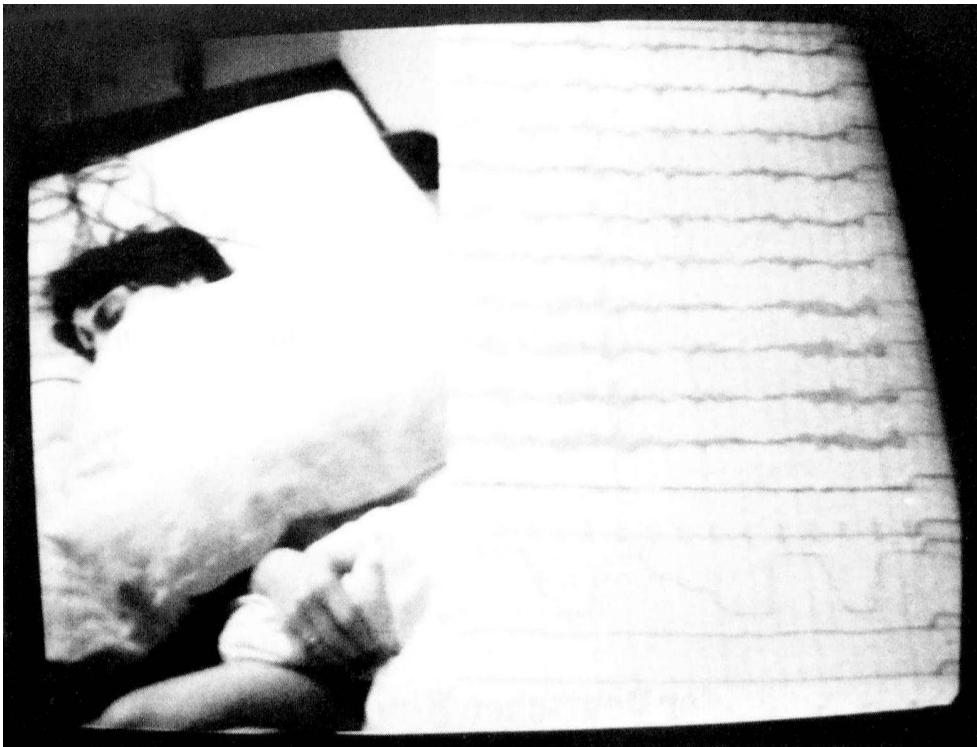
© Michael Klier

2   For this sort of picture, Harun Farocki coined the term "operative image"[3]—an image amenable to mathematical expression and computational processing. As Volker Pantenburg has shown, Farocki's use of the term "operative image" goes back to his review of Roland Barthes's *Mythen des Alltags*, the first German edition of a selection from his *Mythologies*.[4] In terms of the history of ideas, this surprising connection makes perfect sense: Barthes's *Mythologies* was dedicated to the perpetual confusion of myth and history in everyday life and thus exposed an ideological critique that Farocki relates, in 1965, to an operative approach to the world. In an introduction recorded for the radio station Sender Freies Berlin (SFB), Farocki quotes Barthes at length and uncovers these two different approaches to the world (mythical and historical), here still conceived of in linguistic terms:[5] "If I am a woodcutter and I am led to name the tree which I am felling, whatever the form of my sentence, I 'speak the tree', I do not speak about it. This means that my language is operative, transitively linked to its object."[6] Operative language is contrasted here with mythical language. It is a "language resistant to ideological appropriation by the meta-language of myth."[7] But this promise of the autonomy of the pure operation contains an inherent ambivalence: "On the one hand, 'operative images' are the realization of the utopia requiring an autonomy of the image, an autonomy that doesn't depend on mediation, translation, and commentary. These are images that do that work themselves, and they operate entirely at the level of objects and processes. On the other hand, this operationalism also entails, at the same time, the self-erasure of the image as image; it is now a mathematical, technical operation and is no longer required as an image."[8]

3   This article is based on the assumption that the relationship between operative images, which appear to have come into their own in their technicity, and their ideological

appropriation, as related to everyday phenomena and their linguistic naturalization in Barthes's *Mythologies*, has taken on a new form with new actors in today's digital infrastructures. Here, I propose to explore how images operate in the specific field of automatic filtering techniques in contemporary platform capitalism, where their operation is embedded in the myth of the neutrality of major online service providers, from Meta to YouTube.[9] Writing from the perspective of the history of ideas in *What Tech Calls Thinking*, Adrian Daub has observed that the ideology of the Silicon Valley tech companies, which channel, direct, and delete our data streams, is based squarely on the concealment of their own historical development, their own history.[10] Their CEOs, from Mark Zuckerberg to Elon Musk, hide behind the rhetoric of technological progress, as though their location had no history—the old trick, deciphered by Barthes, of rhetorically transforming historical development into the universal nature of images and signs.

Loss of Context

4   In 2016, an image that had featured in the history books for decades was roundly deleted from social media; the anti-war icon *Napalm Girl*, a press photograph by photographer Nick Ut, originally published as *The Terror of War*, had been automatically deleted by Facebook's filtering system on the grounds that it included inappropriate content. More recently, the Digital Services Act (DSA), which came into force in Europe on November 16, 2022, and regulates the deletion practices of the global online platforms, has promised greater transparency in the content moderation of the many billions of visual datasets that are saved, tagged, shared, and deleted on the major online platforms. And yet users are still confronted with arbitrary and nontransparent decisions on image deletions and blocked accounts. Under European law, the enactment of the DSA ought to bring stronger regulation to the indexing, deletion, and moderation of content, processes that use convolutional neural networks (CNNs) to siphon off vast quantities of user data. The new legal powers promise to bring fundamentally improved transparency to the many filtering techniques—from metadata filtering to hashing[11]—that are used to process large datasets (e.g., Image Net) in real time, often applying problematic labels to audiovisual data and "purging" them from social media.[12] Previously, iconic images such as Ut's *Terror of War* were fed into the deep-learning architectures of the social media concerns as operative images, where they were subjected to the new assessment criteria without the binding forces of traditional contexts such as the press, the museum, or the judicial system. Of course, platforms like Meta and YouTube have no interest in historiographical questions or in using new technologies to reassess historical images. By nature, they are built on viable business models and the optimization of profit. But critical research on the history of images should be interested in the political decisions taken by the platforms and the significant changes these entail for our knowledge of the world of images.

5   According to my thesis, since the year 2000 there has been a merger of the technical invisibilization of image processes (blackboxing) and the institutional invisibilization of the criteria by which visual content is evaluated, indexed, and deleted by the major platforms, and a subsequent transformation of these processes and criteria into algorithmic micromanagement and new evaluation scenarios.[13] Both forms of

invisibilization—of production processes and institutional publication processes—have converged in the discourse on operative images.

Algorithmic Governance

6   When Bruce Sterling wrote his brief history of the internet in February 1993, it was still plausible to talk about the new technological infrastructure of the digital age in terms of its absolute freedom: "There is no 'Internet Inc.' There are no official censors, no bosses, no board of directors, no stockholders."[14] The network was "strictly technical" and not socially or politically constituted. Around the turn of the millennium, the great promise of "a true, modern, functional anarchy" was supplemented by the utopias of connectivity and "liveness" on social media platforms. On Flickr, Twitter founder Jack Dorsey described his idea, which he came up with one night in July May 31, 2000, for a highly mobile app that would enable status updates anytime and anywhere: "Real-time, up-to-date, from the road. Akin to updating your AIM status from wherever you are, and sharing it."[15] The particular charm of the mythical origin story contained in this "twttr sketch" derives from its localization in sunny California ("I was living in Sunshine Biscuit Factory"), where tech companies had been working to revolutionize the social design of our world through technology ever since the 1970s. Nowadays, we're increasingly concerned with the flip side of this technology—namely, the question of how the excessive connectivity of the formerly anarchic and seemingly boundless possibilities of the internet could produce a politics and praxis of regulation, deletion, and censorship that pervade every aspect of digital culture.[16] At the same time, the idea of a free internet was confronted with its first issues in the early 1990s: a glut of copyright violations and pornographic material in the form of "piracy" and "cyberporn" drew attention to decentralized networks such as the music-sharing platform Napster in an early test for the utopian notions of sharing and connectivity.[17] When it came to online image circulation, photo historian André Gunthert, writing as recently as 2009, could still make the sweeping assertion that the fundamental value of images on Web 2.0 lay in their shareability.[18]

7   There are currently more images being shared on social media, sharing platforms, and search machines than ever before—and more of them are being deleted and blocked. The technical upscaling of digital infrastructure poses a fundamentally new problem for the regulation of images. This has important consequences both for image-making practices and for the way we look at images today. There's been a shift from looking at individual images and series of images to the infrastructures of images, and this has led to new practices and technologies of image deletion.

8   As a result of social media, the classic scenarios of image regulation have been extended to include new stakeholders and decision-makers who operate within completely new parameters. In an age of "infodemics" and image sharing, content regulation, formerly monopolized by individual nation states, is increasingly becoming the responsibility of US American companies. And so, with increasing regularity, their instruments of domination are also being regarded as techniques of governance. In light of this transformation of global power structures, cultural scientist Joseph Vogl has described players that make capital from information channels as "parastate" social infrastructures: "The major platforms are impinging on state and parastate territory. . . . And the creation of value is sealed and secured by taking responsibility

for what used to be the tasks of the state."[19] Kate Crawford and Tarleton Gillespie have referred to these new techniques of government as "algorithmic governance."[20] Algorithmic governance is bound up with a marked shift that is to be understood as the systematic socio-technical control of media content on social networks; previously, it was given the soft designations of moderation, regulation, and "editing." "All platforms moderate content," writes Friedrich Balke, "and they all attempt to distinguish between that activity and censorship, but what's striking is that their attempts to implement moderation as a complex and extremely labour-intensive, high-level process (both in terms of technology and personnel) are accompanied by so little willingness to expose the principles and techniques of moderation—in short, the apparatus of moderation—to public scrutiny."[21]

9 Instead, novel "media and network effects . . . and the opacity of the new gatekeepers, which act as largely invisible instances of prefiltering, selection, weighting, and potentially viral circulation"[22] have led to the consolidation of three main tendencies: a loss of power on the part of traditional gatekeepers such as newspapers and editorial boards, an increase in the importance of the technical infrastructure that directs operative images, and a concomitant increase in censorious black box effects, which render the control systems of image data invisible. As a way of describing this development towards invisuality on the level of operative images, the paradigm of blackboxing has been gaining in significance since the year 2000. According to Bruno Latour, successful technologies—besides fulfilling their functions—work toward their own invisibility: "[Blackboxing is] the way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become."[23] Since 2010 (the year Instagram was founded), the opaque processes of content regulation have been able to continue consolidating their positions without triggering widespread public debate because, thanks to a narrative that's remained firmly in place ever since the 1990s, the platforms have acted as neutral facilitators and have thereby obfuscated their own role as gatekeepers. Facebook, Twitter, and Co have worked ceaselessly to maintain an image of neutrality. As Simon Rothöler puts it: "The myth of neutrality also appears fundamentally implausible in light of moderation practices that effectively operate as censorship."[24] After the millennium, the new agents of platform capitalism—which continue to expand on a position that straddles social media, information machinery, and commercial enterprise despite the persistent danger of social media escalation—had sought to absolve themselves of any responsibility for the content posted to their platforms. In legal terms, the "safe harbor privilege" and the Communications Decency Act of 1996 meant that these companies were de facto neutral "intermediaries" or "distributors" and could guarantee distribution of content without having to bear the responsibilities that usually fall to a publisher. Sarah T. Roberts gives two reasons why the platforms are obliged to maintain this master narrative of neutrality: protection against copyright claims and protection against reputational damage.[25] In the European context the companies are coming under increasing legal pressure as telemedia service providers owing to legislative efforts such as the Netzwerkdurchsetzungsgesetz (Network Enforcement Act, NetzDG) of 2017 and the abovementioned Digital Services Act of 2022, which are essentially aimed at stronger media regulation and increased transparency of criteria. But neither the national nor the European legislation has

produced any resounding success so far. So the question of whether deletions and blockings by platforms should be regarded as censorship or legitimate content regulation (e.g., in the field of copyright) really is bound up with the question of what sort of new actors we're dealing with when we as "produsers" post, comment on, and like content on their platforms.[26]

Filtering Techniques

10　If we accept that the regulatory techniques of the social networks are defined by the lack of transparency of their internal mechanisms, the arbitrariness of deletion processes, and the rise of the responsible companies to powerful parastate infrastructures in which operative images are now embedded, then another problem can be added to that of indecent imagery, which has always set the boundaries of what can be shown. Judging these boundaries has invariably depended on types of use and social and historical contextualization, which are now "judged" in turn by the moderation apparatus of social media platforms. The automatic deletion of difficult images also has consequences for the sphere of iconic imagery and can lead to abstruse decontextualizations. Given the many billions of bits of image data that exist in the cloud, filtering techniques are now automatically integrated into the infrastructures of online services on the basis of machine learning.

11　One of the most important images in the history of violent conflict in the twentieth century has given new currency to this question—which was always relevant to analogue media—as a result of its recontextualization on social media. This press photograph, Nick Ut's Terror of War (1972), shows a group of children in the foreground—framed by armed soldiers and being forced away from an ominous cloud of smoke—fleeing along an open road (Fig. 2). The pained faces, the wailing mouths, and the nakedness of the girl in the middle ground, who runs toward the photographer with outstretched arms, depict with drastic urgency the horrors of the Vietnam War in 1972. The girl at the center of the image, later identified as Phan Thị Kim Phúc, was naked because her clothes had been burned off by combustibles from the napalm bombs that were dropped on the village of Tràng Bàng. Nick Ut, who was just twenty-one years old at the time, had driven to Tràng Bàng as a war correspondent for the Associated Press (AP); starting at eight in the morning, he shot photographs for several hours. Only this one became iconic.

2. Nick Ut (Associated Press), *The Terror of War*, Vietnam, June 8, 1972

Analog print, 15.2 × 24.1 cm, Washington DC, National Gallery of Art.

12   The image that went down in the history of photography as *Napalm Girl* was selected as World Press Photo of the year and won the Pulitzer Prize in 1973. But prior to publication, it only escaped censorship by a whisker. During the Vietnam War—which was misleadingly called the "uncensored war"[27]—film footage and photographs from war correspondents became enmeshed with the visual politics of American ideology because of the fear of anti-war protests back home: the US government claimed that the Vietnamese communists had started the war and that American military policy was strengthening democracy in Vietnam, and so many images of atrocities were withheld.[28] In the American occupied zones, there were special regulations governing the editorial selection of images: all photographs had to pass inspection by a press officer before they were sent to the press. Despite these constraints, *Napalm Girl* first got past the press officers, then the AP picture editors in Saigon and New York, and appeared in *The New York Times* the day after it was taken. Almost fifty years on, the image was still firmly anchored in cultural memory as an anti-war icon. It served as an example of how war photographs could influence momentous political decisions: along with the Watergate scandal, the worldwide publication of *The Terror of War* and other photographs of the violent conflict in Vietnam forced the resignation of Richard Nixon in 1974.

13   As we've already heard, *The Terror of War* was republished in 2016, but this time it was deleted. As a result of algorithmic and human actors working in concert within the Facebook moderation apparatus, the historic image was removed from newsfeeds.

14   Owing to the increasing demand around the world for moderation of the 1.8 billion websites currently on the internet, there is no alternative to technologically facilitated automatic content moderation.[29] With the global data situation being what it is, only statistical data analyses, not nuanced discussions among image historians, reach beyond the bounds of what can be displayed, and Napalm Girl becomes a case of statistical data filtering. The future trajectory is one of increasingly refined automation.[30] This automation of deletion processes presupposes the conversion of

images into data or datasets that exist as data points or metadata with specific characteristics (categories, spatial coordinates, forms, trajectories, time stamps, etc.)[31] —operative images, in other words, which are normally saved as data and which, in certain specific situations, are materialized as iconic images (e.g., on a screen). So in digital infrastructures, the processing and deletion of images is increasingly transposed into the realm of the invisual, the non-visible.[32] The three main categories in the realm of audiovisual filtering and deletion here are, firstly, metadata filtering through the capture of characteristics such as authorship, title, or standardized identifiers (DOIs); secondly, hashing, where a hashing algorithm calculates a value from bit-level data contained in an image, video, or audio file and compares it to a pre-existing database; and thirdly, fingerprinting, which, like hashing, compares image data to a pre-existing database, though unlike hashing it also recognizes content that has been modified (e.g., colors, dimensions, format, etc.).[33] These filtering techniques organize and redefine the categories that determine the parameters for the use of digital images. In the past, social processes for the negotiation of borderline cases of what can be shown were regulated by legal, religious, editorial, or artistic norms. If image data are converted, indexed, and classified by fingerprinting or a hashing system that uses automatic image-recognition technology on the internet, the media or institutional contexts of images are unraveled and new criteria are introduced by an algorithmic bias.

15 Hence, the appearance of operative images represents a sudden conversion from mathematically operable data to visual structure, albeit one that's already been prestructured by automatic filtering. All materialized imagery is surrounded by the non-visuality of its underlying data. The interesting thing here is that—aside from discussions about spam and fake accounts, which make up the vast majority of deleted content on the major platforms—public debate is dominated by discussion of online bullying, drugs, and terrorism or organized hate speech. According to the statistics, though, all these areas make up just a fragment of the content deleted by, for instance, Meta. The main targets of automatic filtering systems are images of nudity and violence.[34]

The Moderation Apparatus

16 This strong focus on violence and representations of nudity on the part of the filtering systems in turn has consequences for the debate about our present example, the historic anti-war icon. Even in 1972, the newspaper editors and press officers had their doubts about whether this shocking image of violence and nudity could be shown to the public. Under the conditions of publication on social media, though, such acts of deletion become a fundamentally different matter. The digital revenant of the analogue photograph from 1972 is fed into social media, where, as one image among billions, it is then detected by a filtering system and made to appear for inspection on the screen of a content moderator.[35]

17 Although content deletion on sharing platforms and social media such as Facebook/ Meta relies on specific deletion architectures, according to Tarleton Gillespie, the moderation apparatus of the online platforms can be roughly divided into four fields in which both human and non-human agents operate: "editorial review," "crowdworking," "automatic detection," and "community flagging."[36] Due to the volume of image data, the entire development process is currently shifting away from

human content moderators toward technical automatic filtering. When our example of *The Terror of War* is shown alongside photographs of decapitations and child pornography, it is effectively taken out of the context in which it was originally made and published, then subjected to Facebook's "community guidelines" on the regulation of content according to the criteria of pornography, obscenity, violence, illegal activity, abuse, sexual harassment, and hate speech. In 1972, when the newspaper editors were evaluating the press photograph, they too had to decide whether it was a historical document about the war or an unacceptable representation of "frontal nudity," which wasn't allowed to be shown in the US American media. On this basis, the deputy picture editor from the AP in Saigon, having been through all the negatives, initially rejected *Napalm Girl.* It was only by chance that the editor-in-chief, Horst Faas, happened to go through the checked images again later that day and decided to revoke the earlier decision and send the photograph to the head office of the AP in New York. The heated and difficult debate that followed—conducted via telex between the editor in Saigon and picture editor Hal Buell in New York—turned on whether an exception needed to be made. The argument in favor of historical necessity would ultimately change the course of history. In the face of image filtering and content moderation, this considered decision on the part of the newspaper editors, who were also acting as gatekeepers of politically sensitive information, has given way to the deletion processes in neural networks. Moreover, the moderation apparatus of the social networks is marked by a structural invisibility that also features at the level of the working conditions of those who implement the regulation of image content from behind the scenes.

18   Hans Block and Moritz Riesewieck's documentary film *The Cleaners* (2018) follows the case of the *Napalm Girl* through the activity of the Facebook content moderators in the Philippines, who spend ten hours a day checking around 25,000 images for violations of company guidelines on depictions of violence and nudity. "So the number one grave error is approving a nude photo. Boobs or male genitals—totally unacceptable."[37] These are the words of one of the moderators interviewed in the film, who has nightmares about the many thousands of genitalia she deletes every day. The algorithms of the US American company tend to preselect images of nudity and violence. In this way, the guidelines drawn up in Silicon Valley become globalized and the moral conceptions that underlie them are reproduced by omissions in the international fabric of images. According to the Facebook guidelines, questions about historical images and press freedom aren't even posed. The content manager ultimately deletes Nick Ut's image because it violates Facebook's rules for curbing child pornography.[38] But the content moderator only has a few seconds to assess each image; she earns one to three dollars a day, and if more than 3 percent of her decisions are wrong, she loses her job.[39] Although nudity had been a problematic issue in the historical publication of this anti-war image, its decontextualization by social media had moved it into the proximity of child pornography. In the longer term, this sort of confusion of criteria in the internal mechanisms of platform regulation presents a real risk that certain iconic images will disappear from cultural memory.

19   Despite all this, the self-descriptions of the conscientious content moderators are pervaded by metaphors of cleansing, which they use to protect themselves from mental breakdowns: their work makes the internet "healthier," "safer," and "cleaner." They "clean up the dirt." In this respect, *Napalm Girl* was a clear-cut case. As Reuters reports, it was even used in content moderation training sessions, as an example of the sort of picture that must always be deleted. "Ut's photo is a test image, literally and

figuratively, a proxy with which an industry and a society draws parameters of what is acceptable."[40]

20   Luckily, the deletion came to the attention of informed users of the social network and thus triggered a scandal anyway. Norwegian author Tom Egeland wrote a piece for the *Aftenposten* newspaper about how images have influenced the history of armed conflict, and he posted the article to Facebook along with the image in question. In line with company guidelines, content moderators deleted the post. After it was removed, Egeland reposted the article with a critique of Facebook's deletion practices. His account was then blocked twice, first for twenty-four hours, then for three days. The editor-in-chief of *Aftenposten*, Espen Egil Hansen, then wrote an open letter to Facebook's CEO Mark Zuckerberg: "Dear Mark, I am writing this to inform you that I shall not comply with your requirement to remove this picture."[41] Besides expressing incomprehension at the limitations imposed on his editorial team, Hansen made it clear that the image belonged to the genre of historical documentary photography and therefore played an important part in the media history of armed conflict in the twentieth century. The picture was only reinstated when the case came to the attention of prominent politicians, and Facebook was flooded with images of the *Napalm Girl.* Norwegian Prime Minister Erna Solberg created the campaign with the widest reach when her post of the *Napalm Girl* and other gems from the global history of photography got 18,203 likes and almost 5,000 retweets (Fig. 3).

"While I was on a plane from Oslo to Trondheim, Facebook deleted a post from my Facebook page"
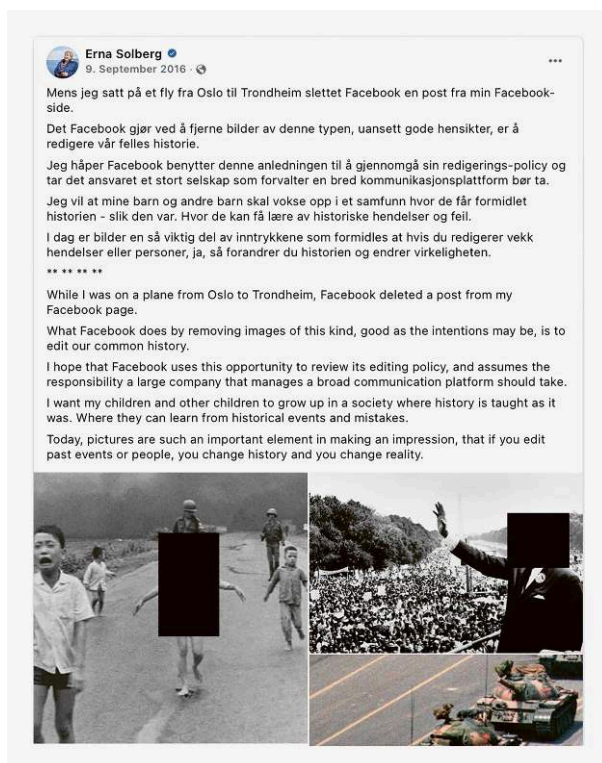
Screenshot of a Facebook post published by Erna Solberg, September 9, 2016.

21   When her post of the *Napalm Girl* was deleted after just three hours, she pasted black boxes into the historical photographs, visually combining Facebook's deletion practices with classic image censorship techniques, some of which have been in use since the

days of analogue print media: blacking, pixelation, and whitewashing. In stark contrast to the underlying photographic images, the black boxes obscure the offending details and remove the objectionable content from visibility while simultaneously highlighting its significance. The ambivalent simultaneity of the "Look away!" and "Over here!" achieved by the integration of these abstract black pictorial elements becomes a visual identifier of picture censorship. With reference to the manipulated photographs—an unrecognizable Winston Churchill, the face of Martin Luther King obscured by a black box—Solberg speaks of the unacceptable editing of collective historical images by Facebook and, on the visual level itself, places it squarely in the realm of censorship (Fig. 4).[42] In the comments on the manipulated image of Martin Luther King, one user has posted a picture of Mark Zuckerberg hidden behind a black box.

4. Photograph of Martin Luther King retouched by Erna Solberg, screenshot of her Facebook post, September 9, 2016

22   William J. T. Mitchell has previously drawn attention to the instability and context-dependency of images: "Offending images are radically unstable entities whose capacity for harm depends on complex social contexts. . . . The offensive character of an image is not written in stone but arises out of social interaction between a specific thing and communities."[43] With its emphasis on the three fundamental strategies of image negation, which Mitchell defines as the basic iconoclastic operations of destruction, defacement, and concealment, his image theory is concerned with manual practices and individual images. What he identified as the social character of the process of negotiating their significance now needs to be expanded to include the socio-technical ensemble of digital platform regulation. For the social element of the social media platforms is now intrinsically linked to the operational logics of distribution.

23   Given the circulation of images on the internet, the profound effects of platform regulation will always raise the question of what we actually want to see as individuals. In a universally networked world, status updates, Instagram stories, and tweets on X (formerly Twitter) populate our newsfeeds with images that once again ask the old question about the limits of what can be shown to each individual. In the face of so many graphic depictions of violence, hate, and nudity, the answer is as ambivalent as the question itself. And the situation is complicated by the fact that automated deletion has become more prevalent since the coronavirus pandemic. Whereas previously the

evaluation of images in the context of content moderation still depended on human judgment, now it is increasingly being replaced by automated deletion practices in order to deal with the sheer volume of global image data. In March 2020, when the coronavirus attained the status of a global pandemic, the following information was circulated via online news platforms and social networks: owing to Covid-19 conditions, YouTube would be shifting more responsibility for the curation and deletion of content from its human content moderators to AI-based systems. The social distancing requirements for limiting the spread of the pandemic could not be met in their open-plan offices, so they would instead be relying on the "virus-free" technology of automatic deletion. Once Google's 100,000 or so employees had been sent home to work —including those responsible for the moderation of YouTube content—there would be more mis-categorizations and more overblocking than usual.[44] The practice of moderating visual content, previously based on a combination of algorithmic recognition and human verification, would "temporarily start relying more on technology to help with some of the work normally done by reviewers."[45] Other sharing platforms such as Facebook, Twitter, and TikTok also gave advance warning that users' content and accounts would be "autonomously deleted by automated systems in future."[46] Although the pandemic is now behind us, increasing automation of deletion practices is unfortunately here to stay; it has changed the structure of what we can know about the world of images, and of human judgment in the face of vast numbers of images.[47]

NOTES

1. Harun Farocki, "Kamera in Aufsicht," *Filmkritik* 27, no. 9 (September 1983): 416–19, here: 416.

2. Farocki, "Kamera in Aufsicht": ". . . einen Stecker in den Kopf stecken und somit alle Wörter, Geschichten, Träume abzapfen zu können."

3. Harun Farocki introduced the term "operative images" to describe the military and industrial uses of digital technologies. It refers to images that "do not represent an object, but rather are part of an operation." Harun Farocki, "Phantom Images," trans. Brian Poole, *Public* 29 (2004): 12–22, here: 17.

4. Daniel Eschkötter and Volker Pantenburg, "Was Farocki lehrt," *Zeitschrift für Medienwissenschaft* 11 (February 2014): 207–11.

5. Harun Farocki, "Einführung: 'Mythen des Alltags' von Roland Barthes," part of the radio series *Thema*, edited by Lore Ditzen for Sender Freies Berlin, June 26, 1965 [manuscript].

6. Roland Barthes, *Mythologies*, trans. Annette Lavers (1972; New York: Noonday, 1991), 146. I'm grateful to Tom Holert for this reference to Farocki's early reception of Barthes.

7. Eschkötter and Pantenburg, "Was Farocki lehrt," 207: ". . . gegen die ideologische Vereinnahmung durch die Metasprache des Mythos resistente Sprache."

8. Eschkötter and Pantenburg, 207–8: "Denn einerseits verwirklicht sich in den 'operativen Bildern' die Utopie, dass es eine Eigengesetzlichkeit und Autonomie des Bildes geben muss, die nicht auf sprachliche Vermittlung, Übersetzung und Kommentierung angewiesen ist. Es sind Bilder, die selbst arbeiten und vollständig auf der Ebene der Objekte und prozessualen Vollzüge agieren. Andererseits jedoch löscht sich das Bild als Bild in dieser Operativität zugleich aus; es ist

nicht mehr als Bild, sondern nur noch als mathematisch-technische Operation gefragt." For a general study of the operative image, see Jussi Parikka, *Operational Images: From the Visual to the Invisual* (Minneapolis: University of Minnesota Press, 2023).

9. For the US American discourse, see Nick Srnicek, *Platform Capitalism* (Cambridge: Polity Press, 2017). The term "platform capitalism" was introduced into German-speaking discourse by web activist Sascha Lobo. It designates a new economic order in which digital platforms such as Google, Apple, Facebook, and Amazon (collectively, GAFA) dominate the market by "setting and controlling standards in the sector," orchestrating "each and every economic transaction as an auction." See Sascha Lobo, "Auf dem Weg in die Dumpinghölle," *Spiegel Online*, September 3, 2014, https://www.spiegel.de/netzwelt/netzpolitik/sascha-lobo-sharing-economy-wie-bei-uber-ist-plattform-kapitalismus-a-989584.html.

10. Adrian Daub, *What Tech Calls Thinking: An Inquiry into the Intellectual Bedrock of Silicon Valley* (New York: Farrar, Straus and Giroux, 2020).

11. An overview of the various filtering techniques used in neural networks and investigated in the context of the Digital Services Act can be found in a paper published by the Policy Department for Citizens' Rights and Constitutional Affairs, Directorate-General for Internal Policies of the European Union, "The Impact of Algorithms for Online Content Filtering or Moderation" (September 2020), 39ff., https://www.europarl.europa.eu/RegData/etudes/STUD/2020/657101/IPOL_STU(2020)657101_EN.pdf.

12. See the terminology of Facebook content moderation in the documentary film *The Cleaners* (Hans Block and Moritz Riesewieck, Germany, 2018), https://www.bpb.de/mediathek/video/273199/the-cleaners/, at 35:40–35:53; see also Moritz Riesewieck, *Digitale Drecksarbeit: Wie uns Facebook & Co. von dem Bösen erlösen* (Munich: dtv, 2017).

13. See Joseph Vogl, *Kapital und Ressentiment: Eine kurze Theorie der Gegenwart* (Munich: Beck, 2021).

14. Bruce Sterling, "A Short History of the Internet" (February 1993), 3, http://sodacity.net/system/files/Bruce_Sterling_A_Short_History_of_the_Internet.pdf.

15. Jack Dorsey, "twttr sketch: On May 31st, 2000, I signed up with a new service called LiveJournal," flickr, July 5, 2006, https://www.flickr.com/photos/jackdorsey/182613360.

16. For a discussion of disconnecting or "unplugging" as a result of hyper-connectivity, where flight into the analogue world becomes a matter of urgency, see Urs Stäheli, *Soziologie der Entnetzung* (Berlin: Suhrkamp, 2021).

17. Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven, CT: Yale University Press, 2018), 29.

18. André Gunthert, "Shared Images: How the Internet Has Transformed the Image Economy," trans. James Gussen, *Études photographiques* 24 (November 2009), http://etudesphotographiques.revues.org/3436.

19. Julia Encke and Harald Staun in conversation with Joseph Vogl, "Die Nutzer spielen mit," *Frankfurter Allgemeine Sonntagszeitung*, March 14, 2021, https://www.faz.net/aktuell/feuilleton/debatten/plattformkapitalismus-joseph-vogl-ueber-kapital-und-ressentiment-17241098.html: "Die großen Plattformen drängen gewissermaßen in staatliche und parastaatliche Aufgabengebiete. ... Und die Wertschöpfung wird durch die Besetzung ehemaliger Regierungsaufgaben komplettiert und gesichert."

20. Kate Crawford and Tarleton Gillespie, "What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint," *New Media & Society* (July 2014): 411–28, here: 414.

21. Friedrich Balke, "Zensieren," in *Historisches Wörterbuch des Mediengebrauchs*, ed. Heiko Christians, Matthias Bickenbach, and Nikolaus Wegmann, vol. 3 (Cologne: Böhlau, 2022), 695–724, here: 695–96: "Alle Plattformen moderieren – und sind bemüht, diese Aktivität grundsätzlich von Zensur zu unterscheiden, wobei auffällig ist, dass ihre Bemühungen, die Moderation als einen extrem aufwendig und hochgradig (technisch wie personell) ausdifferenzierten Vorgang zu

implementieren, mit wenig Bereitschaft verbunden ist, die Prinzipien und Techniken der Moderation, kurz: den Moderationsapparat einer öffentlichen Inspektion auszusetzen.”

22. Bernhard Pörksen, *Die große Gereiztheit: Wege aus der kollektiven Erregung* (Munich: Hanser, 2018), 64: “Medien und Netzwerkeffekte … und neue intransparent agierende Gatekeeper [wirken] als weitgehend unsichtbare Instanzen der Vorfilterung, der Auswahl und Gewichtung sowie der potentiell epidemischen Verbreitung.”

23. Bruno Latour, *Pandora's Hope: Essays on the Reality of Science Studies* (Cambridge, MA: Harvard University Press, 1999), 304.

24. Simon Rothöler, “Informationen, die Bilder haben: Zur Moderierbarkeit von visuellem Content,” *Zeitschrift für Medienwissenschaft* 19, no. 2 (2018): 85–94, here: 87: “Prinzipiell fragwürdig erscheint der Neutralitätsmythos zugleich mit Blick auf Moderationspraktiken, die effektiv als Zensur wirksam werden.” For the current definition of the relationship between content moderation and censorship, see Katja Müller-Helle, *Bildzensur: Infrastrukturen der Löschung* (Berlin: Wagenbach, 2022).

25. Sarah T. Roberts, “Digital Detritus, ‘Error’ and the Logic of Opacity in Social Media Content-Moderation,” *First Monday* 23, nos. 3–5 (March 2018): “This operating logic of opacity serves to render platforms as objective in the public imagination, driven by machine/machine-like rote behavior.”

26. Joseph Vogl employs the term “produser” to designate the conflation of user and producer when engaging with these platforms. Early instances of the word occur in works by artist Daniel Kellers around 2010; see Vogl, *Kapital und Ressentiment*, 78ff.

27. See Daniel C. Hallin, *The “Uncensored War”: The Media and Vietnam* (Oxford: Oxford University Press, 1986), 114.

28. Just 10 per cent of the 4,100 articles on this conflict (1955–75) were actually about warfare—the vast majority of reports were about corruption in Saigon and US American soldiers who had been killed; see Lars Klein, “Größter Erfolg und schwerstes Trauma,” in *Augenzeugen: Kriegsberichterstattung vom 18. zum 21. Jahrhundert*, ed. Daniel Ute (Göttingen: Vandenheck & Ruprecht, 2006), 197–202. See also Marc Frey, *Geschichte des Vietnamkriegs: Die Tragödie in Asien und das Ende des amerikanischen Traums* (Munich: Beck, 1999), 67–68.

29. See Policy Department for Citizens' Rights and Constitutional Affairs, “The Impact of Algorithms,” 22.

30. Lev Manovich, “Data Science and Digital Art History,” *DAH-Journal* 1 (2015): 22–23.

31. Manovich, “Data Science.”

32. Adrian MacKenzie and Anna Munster, “Platform Seeing: Image Ensembles and Their Invisualities,” *Theory, Culture & Society* 36, no. 5 (2019): 3–22.

33. See Policy Department for Citizens' Rights and Constitutional Affairs, “The Impact of Algorithms,” 39–40.

34. See Paul M. Barrett, *Who Moderates the Social Media Giants? A Call to End Outsourcing* (NYU Stern Center for Business and Human Rights, 2020), 10: https://bhr.stern.nyu.edu/wp-content/uploads/2020/06/NYUContentModerationReport_FINALVERSION.pdf.

35. On the term and the figure of the commercial content moderator (CCM)—here abbreviated as content moderator—and their function on social media platforms, see Sarah Roberts, “Your AI Is a Human,” in *Your Computer Is on Fire*, ed. Thomas S. Mullaney, Benjamin Peters, Mar Hicks, and Kavita Philip (Cambridge, MA: MIT Press, 2021), 51–70.

36. Gillespie, *Custodians of the Internet*.

37. See *The Cleaners* (Hans Block and Moritz Riesewieck, Germany, 2018), https://www.bpb.de/mediathek/video/273199/the-cleaners/, at 21:34–21:48.

38. Ibid., 35:40–35:53.

39. Moritz Riesewieck, *Digitale Drecksarbeit: Wie uns Facebook & Co. von dem Bösen erlösen* (Munich: dtv, 2017).

40. Gillespie, *Custodians of the Internet*, 4.

41. Espen Egil Hansen, "Dear Mark," *Aftenposten*, September 8, 2016, aftenposten.no/meninger/kommentar/i/G892Q/dear-mark-i-am-writing-this-to-inform-you-that-ishall-not-comply-with-your-requirement-to-remove-this-picture.

42. Erna Solberg, Facebook post, September 9, 2016, facebook.com/permalink.php?story_fbid=pfbid02P79i1NJgHnRfLXbNvexiaWHQE2R4akDJ8t2tMx7u86H4aWh2qYS8B35FbdJhjkeol&id=56800396831.

43. W. J. T. Mitchell, *What Do Pictures Want? The Lives and Loves of Images* (Chicago: University of Chicago Press, 2005), 131.

44. See YouTube blog, March 16, 2020: https://blog.youtube/news-and-events/protecting-our-extended-workforce-and/. Overblocking, also known as overtargeting or overcensoring, is the technical blocking of sensitive content by excessively sensitive filters, outdated databases, and misguided pattern-recognition software, which subsumes individual cases under generalized guidelines. But overblocking also occurs where national investigative agencies block multiple IP addresses or entire IP subnets and where content is filtered for copyright or child protection purposes.

45. Ibid.

46. Simon Hurtz, "Warum das Virus die Redefreiheit im Netz bedroht," *Süddeutsche Zeitung*, March 27, 2020, https://www.sueddeutsche.de/digital/coronavirus-facebook-google-content-moderation-1.4859147.

47. See Marion Kliesch, *Ästhetik der Zensur* (Salenstein: Niggli, 2017), 126.

ABSTRACTS

With the billionfold increase in the amount of image data that is now saved, indexed, and sent via neural networks every day, there has, since the year 2000, been an ever-increasing demand for filtering technologies and content moderation on global platforms. Major service providers such as Google and Meta (Facebook, Instagram) employ automatic deletion technologies in conjunction with filtering systems such as metadata filtering, hashing, and fingerprinting to stem the tide of borderline images. With reference to an iconic press photograph, Nick Ut's *Terror of War* (1972), this text investigates how, since the turn of the millennium, there has been a merger of the technological invisibilization of image processing (blackboxing) and the institutional invisibilization of the criteria by which visual content is evaluated, indexed, and deleted from major online platforms. Iconic images have also been subject to algorithmic micromanagement and thus transposed into new evaluation scenarios that have replaced traditional gatekeepers such as the state, the museum, and newspaper editors. The two forms of invisibilization—of both technological production and institutional publication processes—converge in the discourse of operative imagery. Convolutional neural networks and the social media platforms that depend on them now do not simply produce new operative images, the filter architectures that control them actually have a retroactive effect on pre-existing visual traditions, which in some cases produces radical decontextualizations of historical images.

Avec la multiplication par un milliard de la quantité de données d'images enregistrées, indexées et transférées chaque jour au moyen de réseaux de neurones, le besoin en matière de technologies de filtrage et de modération de contenu sur les plateformes mondiales n'a cessé de croître depuis l'an 2000. Pour contenir le flot d'images « limites », les principaux fournisseurs de

services tels que Google et Meta (Facebook, Instagram) font appel à des technologies de suppression automatique ainsi qu'à des systèmes de filtrage, notamment au filtrage des métadonnées, au hachage et aux empreintes numériques. En prenant pour référence une photographie de presse emblématique, *La Terreur de la guerre* (*The Terror of War*, 1972) de Nick Ut, cet article examine comment, depuis le tournant du millénaire, une fusion s'est opérée entre l'invisibilisation technologique du traitement des images *(blackboxing)* et l'invisibilisation institutionnelle des critères sur la base desquels les contenus visuels sont évalués, indexés et supprimés des principales plateformes en ligne. Les images ont également fait l'objet d'une microgestion algorithmique et ont pu être intégrées à de nouveaux modèles d'évaluation qui ont supplanté les responsables traditionnels du contrôle des images, tels que l'État, les musées et les rédacteurs en chef des journaux.

INDEX

AUTHOR

KATJA MÜLLER-HELLE

Dr. Katja Müller-Helle is director of the research center The Technical Image, a joint department of the Institut für Kunst- und Bildgeschichte and the Zentrum für Kulturtechnik at Humboldt-Universität zu Berlin. From 2013 to 2019, she was a postdoctoral fellow of the Center for Advanced Studies BildEvidenz. History and Aesthetics at Freie Universität Berlin. From 2014 to 2015 and in 2018, she was a Volkswagen fellow at the Getty Research Institute in Los Angeles. Together with Claudia Blümle, Katja Müller-Helle is co-editor of *Bildwelten des Wissens. Kunsthistorisches Jahrbuch für Bildkritik.*
Katja Müller-Helle est directrice du centre de recherche Das Technische Bild, un département regroupant l'Institut für Kunst- und Bildgeschichte et le Hermann von Helmholtz-Zentrum für Kulturtechnik à l'université Humboldt de Berlin. De 2013 à 2019, elle a été titulaire d'une bourse de recherche postdoctorale au Center for Advanced Studies BildEvidenz History and Aesthetics à la Freie Universität Berlin. De 2014 à 2015 et en 2018, elle a été titulaire d'une bourse Volkswagen au Getty Research Institute de Los Angeles. Katja Müller-Helle est coéditrice, avec Claudia Blümle, de la revue Bildwelten des Wissens. Kunsthistorisches Jahrbuch für Bildkritik.