4. Calm Images: The Invisible Visual Culture of Digital Image Distribution

Simon Rothöhler

The visual culture of the present is not usually suspected of being a calm one. The omnipresence of digital images distributed via various social media platforms, messaging apps and streaming services is therefore rarely associated with qualities such as discretion or restraint. On the contrary, this culture is commonly assumed to be dominated by an excess of visibility, which, in the form of viral GIFs, memes, and selfies, is linked to the instantaneous nature of digital image distribution. The ease with which they are produced and their real-time communicative mobilization have turned images, like many other digital products, into a ubiquitous phenomenon with zero marginal costs (Rifkin 2014). Generating another image, distributing it immediately, and potentially sharing it with the whole world, or at least the whole digitally connected world, via social media transport channels, costs a smartphone owner virtually nothing, other than a casual touch of the screen and a minimal amount of energy—which can nonetheless be quantified, and like every search engine query, creates a carbon footprint (Ensmenger 2018; Elegant 2019).

The image-related growth rates are correspondingly spectacular, particularly in a historical comparison. Nicholas Mirzoeff has claimed that at present, in the US alone, more photographic images are being produced every two minutes than in the entire 19th century (Mirzoeff 2016). Other estimates (whose basis, apart from the verifiable sales figures for camera-equipped smartphones, is similarly unclear) assume that humans had put a total of 3.8 trillion photographic images into circulation up until the early 2010s, but added one trillion in 2015 alone (Kane and Pear 2016). Here, we should not overlook the fact that long before the digital present, the evolution of technical visual media was already accompanied by a discourse addressing the increasing volume of images in circulation with a semantics of cultural criticism and metaphors such as that of the "flood of images," as can be demonstrated with a paradigmatic quote from Lewis Mumford in 1952:

> Between ourselves and the actual experience and the actual environment there now swells an ever-rising flood of images which come to us in every sort of medium—the camera and printing press, by motion picture and by television. A

> picture was once a rare sort of symbol, rare enough to call for attentive concentration. Now it is the actual experience that is rare, and the picture has become ubiquitous. (Mumford 1952, 96)

Whatever one's view of the assumed shifts in the economy of attention, there is no doubt that the proliferation of technologically generated images—"ubiquitous photography" (Hand 2012)—is virtually unlimited. And yet it is less possible than ever to use a mere analysis of the surfaces of visual culture—of phenomena materialized in pictorial form—to comprehend the resulting excess of visibility, which has led to massive reservoirs of digital images. We can argue here, with Claus Pias, that the reason for this lies in the nature of digital images, insofar as these can be clearly distinguished, in media technology terms, as images "that have information" (Pias 2003). We could rephrase the argument to say that digital images have a double existence, in a sense: firstly as data sets with which computers can carry out calculations, independent of their software-based transformation into visual output, i.e. beyond visual materialization, and secondly as a visual form that can be recognized and viewed as an image by human perceptions, which find non-processed image data rather uninformative—although this can only happen if the image data are mobilized pictorially by appropriate software. The pictorial nature of digital images is not simply given, but is only a potential that must be realized by computation. Without calculation, the image data sets remain completely abstract, nonpictorial, imageless. The conversion of the nongraphic state (which is the norm) into the pictorial state (which is only ever produced temporarily, by the performance of software), is updated contingently and seems fairly expendable from the point of view of the computer (Kittler 1999). Trevor Paglen has summed this up as follows:

> Digital images are fundamentally machine-readable: they can only be seen by humans in special circumstances and for short periods of time. A photograph shot on a phone creates a machine-readable file that does not reflect light in such a way as to be perceptible to a human eye. A secondary application, like a software-based photo viewer paired with a liquid crystal display and backlight may create something that a human can look at, but the image only appears to human eyes temporarily before reverting back to its immaterial machine form when the phone is put away or the display is turned off. However, the image doesn't need to be turned into human-readable form in order for a machine to do something with it. (Paglen 2016)

Without pictorial visualization, however, information (and resources) are also lost, as shown by the still-limited machine-readability of digital images. It becomes clear, for example, from forms of CAPTCHA ("Completely Automated Public Turing test to tell Computers and Humans Apart") that software-constituted image data objects that appear to the human perception, in a subjectively unproblematic, intuitive way, as images *of something*, can even be repurposed as security queries, due to their relative unreadability for machines. In image theory, the conclusion that has been drawn from this is that

> there is a yawning gap between the endless columns of numbers and the shapes which a human gaze recognizes. ... On one side there is raw data which codes images as fields of colored pixels; on the other there is a perception which has no choice but to see something: faces, people, spaces, objects. Only gradually are processes of digital vision machines beginning to find image objects within the fog of the number fields.[1] (Ernst, Heidenreich, and Holl 2003, 11, own translation)

And while (to varying degrees) impressive progress has been made by image recognition or image classification processes based on learning algorithms, processes which "train independently" using sample data (such as the data set of the ImageNet project, established as the competitive standard of machine learning), there are still some rather sobering cross-checks, such as the classification of reverse tone images, which are completely unproblematic for human actors (Hosseini and Poovendran 2017).

The continued existence of the "gap" can also be detected in the currently installed "content moderation apparatus" (Gillespie 2018) of social media platforms, where the relative unreadability of digital images is translated into a socio-technical moderation chain, in which human actors on various levels continue to do most of the work: as members of internal policy teams who decree regulation guidelines; as programmers who configure training data accumulated by crowd workers and evaluate the results of algorithmic scanning; as trainers who instruct content moderators on how to analyze images in conformity with guidelines; and as microworking reviewers, who must carry out up to 2,000 assessments of individual images per hour (Rothöhler 2018b). The conclusion that can be drawn from this is that digital images "have information" insofar as they consist of computable data, but at the same time they can also be understood as information that "has" images. The image data of digital images can be computed regardless of their pictorial manifestation, but in many cases, they then lose that aesthetically and epistemically crucial information which, at least from the point of view of human actors, is only present in the visual form, in the pictorial "exterior" of the image data. W.J.T. Mitchell puts this succinctly: "images have always given form to information" (2010, 46).

With regard to the emerging tension between *image file* and *image display*, code and form (Drucker 2014), it becomes abundantly clear that digital images are not fixed objects, but flexibly configurable processes: "no longer a solid representation of a solid world—a 'hardimage' as it were—but an unstable algorithmic configuration of a database: a 'softimage'" (Hoelzl and Marie 2017, 72). This softimage is one that is constitutionally distributed (Rothöhler 2018)—not only in that most digital images are mainly generated because they can be instantly sent and shared, but also in that these images can no longer be ascribed to a binding media location, a constant materiality, a key actor, a central agency, but can only ever be understood as both fluid and reversible phenomenalizations of an operatively distributed media logistics. Thus, for example, the specific emergence of any given digital image on a platform—of a "social photo," as Nathan Jurgenson (2019) has designated the dominant image type of present-day visual culture—depends on the operative cooperation of numerous actors, including (in generalized terms) data distribution infrastructure such as data centers and transoceanic cable networks, widely used container formats such as Matroska (MKV) and standardized image compression norms such as JPEG, as well as screen hardware and human end users—who still hold a significant proportion of the initiative when it comes to image-related practices.

Calm Images and the Internet of Things

The defining feature of digital images in general and of circulating digital images in particular, then, is their operatively distributed nature as data images. This also means that digital images consist of discrete, discretely addressable units of information, which can be computed using information technology: "The specificity of digital images lies in their operationality and processability"[2] (Schneider 2009, 191, own translation). If we concentrate, in this respect, on those processes that seem largely uncoupled from the contributions of human actors, we enter a discursive field of image theory that Paul Virilio established, with programmatic intent (and with the semantics of cultural criticism), in the late 1980s. Virilio spoke of "infographic vision machines," which would get by without human eye contact, without "visible context," and would, in future, take over the task of "analys[ing] the ambient environment and automatically interpret[ing] the meaning of events"; he also spoke of "synthetic images created by the machine for the machine" (Virilio 1994, 59–60).

The fact that digital images can have agency, the power to act, even if they are not the object of human perceptions, is captured in Trevor Paglen's paradoxical concept of a rapidly spreading "invisible visual culture":

> Visual culture has changed form. It has become detached from human eyes and has largely become invisible. Human visual culture has become a special case of vision, an exception to the rule. The overwhelming majority of images are now made by machines for other machines, with humans rarely in the loop. The advent of machine-to-machine seeing has been barely noticed at large, and poorly understood by those of us who've begun to notice the tectonic shift invisibly taking

> place before our very eyes. The landscape of invisible images and machine vision is becoming ever more active. . . . Images have begun to intervene in everyday life, their functions changing from representation and mediation, to activations, operations, and enforcement. Invisible images are actively watching us, poking and prodding, guiding our movements, inflicting pain and inducing pleasure. But all of this is hard to see. (Paglen 2016)

And in fact, for some years there has been relatively concrete proof of how—especially in the context of the Internet of things (IoT)—image data circuits are created, which have consequences for human actors but are largely invisible. In the IoT, the use of image technologies is on the increase, most notably when it comes to the addressing of nonnetworked things and environmental observation in general. The invisible visual culture evoked by Paglen is, in this sense, focused on addressing and observing. The idea is that in the IoT, even the most mundane everyday objects will be assigned addresses, and will become intelligently communicating "everyware" (Greenfield 2006) thanks to miniaturized microchips und network access. Here, the promised smartness of the IoT is largely based on things whose intelligence does not lie within themselves, but is an effect of their connectedness through media technology, as the UbiComp pioneer Mark Weiser observed over twenty years ago: "The real power of the concept comes not from any one of these devices; it emerges from the interaction of all of them" (Weiser 1999). The distributed nature of the IoT's things, the inscrutability of the computing power connecting them, which is based on infrastructure such as fiber optic cable systems and data centers, leads to connection services which, on the one hand, allow intelligent cooperation between variously competent media technology actors, but on the other hand are hardly able to be understood or controlled any more. According to Weiser, this leads to problems with the delegation of responsibility: "If the computational system is invisible as well as extensive, it becomes hard to know what is controlling what, what is connected to what, where information is flowing, how it is being used, what is broken . . . and what are the consequences of any given action" (Weiser, Gold, and Brown 1999).

But why, one might ask, this special significance of image technologies in the IoT? Firstly, simply because systems equipped with image sensors can be used to address objects that are not especially smart, such as nonnetwork-enabled fridge contents. In the case of Samsung's Family Hub RB 7500 there are in fact three high-resolution image sensors, which continuously scan the fridge interior for objects like milk cartons that need to be reordered. In the Internet of things, intelligently networked cameras are special objects (and especially intrusive ones), in that they can detect, locate, and visually store other things (and processes), and connect them with an address, without their active participation. Detection by means of imaging and image data processing can be carried out discreetly, from a distance, and requires no consent from the objects (or individuals) scrutinized by the gaze of technology. In numerous models of smartness in media technology—one example of a growth market is the application of "ambient assisted living" in elder care—the following is regarded as a fundamental rule of procedure: "vision as a high-level way of capturing information from an environment" (Augusto, Nakashima, and Aghajan 2009). Nonoptical sensor types such as motion sensors do also record things that are not networked, but work with a comparatively low level of data, which limits the options for subsequent processing. In contrast, the neutral-sounding phrase "high-level way of capturing information" indicates that, generally speaking, more patterns can be extracted and modelled algorithmically from image data sets than from data collections originating from other sensory registers. A more recent example of the "depth of information" of imaging processes is the application of Transdermal Optical Imaging, which can use any two-minute smartphone video to detect changes in a person's blood flow pattern, and thus measure systolic blood pressure with 95% accuracy (Luo et al. 2019).

In the invisible image data circuits of the IoT, information acquired from image sensors is automatically analyzed and integrated into more extensive sensor networks, which—for example in smart home concepts such as Alphabet's Nest portfolio—can also capture environmental data according to e.g. mechanical, thermoelectric or magnetic operating principles. The ultimate goal here is always the greatest possible fusion of data. The more sensors are distributed and the more diverse they are, the more information-saturated the environment becomes. Data gathered in nonoptical ways is expected to cooperate with data from image sensors. This also applies to Nest's smart home concept: if Nest detects carbon monoxide, the light is switched on, the Nest Cam is activated, and an alarm is triggered. In this model, the smart home is densely populated with a multitude of "sensing," intercommunicating detectors, which convert different stimuli first into electrical signals, and then into a universally compatible digital data form, so that operative decisions can be calculated and implemented on the basis of data collected in different places but processed centrally. So, what is especially valued about image data within sensor networks is information that it is possible to fuse with information that cannot be materialized in images.

The prerequisite for this "smartification" of environments is a "growing sensorization of environments" (Gabrys 2016). As suggested above, in many cases it is image technologies that are expected to bring the world—objects, environments, events, the behavior of human actors—into the data space. Image data circuits are created here because more and more smart things demand visual input: home security systems, fridges that do the shopping, and self-parking cars are the best-known examples of marketable applications for image-technology-competent "distributed sensing"—visual input for an invisible visual culture, in which image data are gathered and processed, but are no longer given a visible pictorial form as images. From this perspective, the IoT is being dominated increasingly by image machines that are indifferent to pictorial manifestations. Crucially, what we are dealing with here is uses of image technology whose informative value lies beyond human viewing processes. This is because the computer does not need the visual output to calculate with the image data, following its own rules—which it sometimes does very effectively, at least in clearly delimited contexts of application such as smart front doors and fridges. The Internet of things thus maintains a highly ambivalent relationship with the image: On the one hand operations involving image sensors and image processing are becoming increasingly extensive and important. On the other hand, algorithmic agencies are emerging, which are intended to filter the image data circuits according to instrumental stipulations and to minimize human viewing of images.

For most sensor network images on the IoT, there are in fact no longer any viewers, but only scripts that automatically scan image data sets. Kristin Veel has studied the new status this confers on the image, a status that must be thought of in terms of a circulation process that, fundamentally, operates in the background. Veel uses the example of a security system called Smartgate, working with facial recognition software, which is installed in Australian airports. Following Mark Weiser, Veel speaks of "calm imaging," "focusing in particular on images that are not addressing our conscious, focused attention, but that operate smoothly in the background" (Veel 2012, 120). These images "emerge and are received calmly as a matter of computational exchanges, but in the process they acquire new potentialities, qualities, and risks" (123–124). What is new is not the video monitoring principle of remote surveillance, but the closure of a network-based system in which the distribution and visual materialization of the image information is routinely delegated to subsequent algorithmic image evaluations, or is suspended in favor of these processes. Monitoring with image sensors is automated, and human viewing becomes, as a general tendency, superfluous—or is only called upon if the algorithmic image analysis encounters a pattern defined as an event.

(In)Visibility and the Regulation of Image Flows

Calm images are images that neither strive for success on social media, nor contribute to the excess of visibility mentioned at the beginning of this paper; instead, they operate discreetly in the background, remain at the ready for automatic selections, and circulate as image data in the invisible visual culture (Rothöhler 2018, 225–274). Yet the fact that these images are not intended to be viewed by human actors, or only secondarily, does not necessarily mean reduced potential for influence, control and decision-making; in fact, these calm images are sometimes more powerful than many excessively shared viral images.

Nonetheless, this new invisible landscape of calm images is not, in every respect, far removed from the excessive visibility usually associated with the platform-based visual culture of the present. There too, calming processes, occurring prior to or in parallel with the circulation of images within visible visual culture, are becoming increasingly influential. For some years, as suggested above, there have been attempts to delegate social media content moderation to technological actants ("automatic detection"), to further reduce the dependence on editorial and user practices of image screening, checking and evaluation. This is mainly a reaction to the constantly growing volume of images in circulation, as Tarleton Gillespie has commented:

> Content is policed at scale, and most complaints are fielded at scale. More important, the ways moderators understand the problems have been formed and shaped by working at this scale. . . . What to do with a questionable photo . . . when you're facing not one violation but hundreds exactly like it, and thousands much like it, but slightly different in a thousand ways. This is not just a difference of size, it is fundamentally a different problem. For large-scale platforms, moderation is industrial, not artisanal. (Gillespie 2018, 77)

The desired improvement in the efficiency of content moderation, which aims to minimize human contributions, primarily for reasons of cost, proceeds via algorithmic systems of image content processing. Thus, recent methods of automated image recognition have attempted, on the basis of artificial neural networks (ANNs), to recognize patterns in bitmaps. These patterns are meant to become operable as image content information: "They can tell what's in an image by finding patterns between pixels on ascending levels of abstraction, using thousands to millions of tiny computations on each level. New images are put through the process to match their patterns to learned patterns" (Gershgorn 2017). Adrian Mackenzie has described in greater detail these automated processes of image recognition—whose operations, while referred to as "image recognition," essentially consist of processes of algorithmic image classification—with reference to the "cat radar detector" Kittydar:

> Faced with the immense accumulation of cat images on the internet, kittydar can do little. It only detects the presence of cats that face forward. It sometimes classifies people as cats. . . . [T]he software finds cats by cutting the image into smaller windows. For each window, it measures a set of gradients . . . running from light and dark and then compares these measurements to the gradients of known cat images (the so called "training data"). The work of classification according to these simple categories of "cat" and "no cat" is given either to a neural network . . ., themselves working on images of cats among other things taken from YouTube videos, or to a support vector machine. (Mackenzie 2017, 4)

Such limitations point to the structural conservatism of machine learning—whose successes depend on "how well it makes the same distinctions that were made before" (Gillespie 2018, 107). Nonetheless, there is already a well-established trend, not only in the decision-making circuits of the Internet of things but also in the context of social media platforms, to delegate power, operatively binding images to computers. What is already in practice and works relatively efficiently is the filtering (operationalized without any human involvement in the viewing of images) of visual content whose image data has already been identified as belonging to illegal forms of image. The best-known example (and the most widespread one, because of obvious liability risks) is PhotoDNA—an image filtering application developed by Microsoft in 2009, and currently used by nearly all major platform operators. PhotoDNA compares every single user-uploaded image with a database administered by an NGO, the National Center for Missing and Exploited Children (NCMEC), in which the image data of known child pornography content are stored. The currently around 80 million images from the NCMEC database are "forensically" (Farid 2016) identifiable (even when the phenomenalized image form has been modified by user-initiated image manipulation) via a hash function, which compresses the image data set as a hash value, i.e. codes it into a numeric string that is comparatively straightforward to compute.

The problem evoked by Gillespie of the industrial scale of social media content moderation is becoming even more urgent with the proliferation of various live streaming apps, for which a timely intervention regime would need to be operationalized almost as real-time content moderation. For this and other reasons, platform operators increasingly find themselves forced to delegate the observation and evaluation of images—necessary on a large scale—to technological actants of the invisible visual culture. eGlyph, the successor project to Photo DNA (and also supervised by Hany Farid), which focuses on the extremist content of terrorist organizations, is only partially automated. Like most attempts to solve such problems, it is embedded in a "broader sociotechnical apparatus" (Gillespie 2018, 101), which works through scripts that have been co-conceived and observed by human actors and are intended for the identification and evaluation of image content. Nonetheless, it goes a step further than the calculation of image-specific hash functions, which can be regarded as the "fingerprints" of image forensics.

On the one hand, this program hashes and algorithmically scans not just images but more extensive video data (in order to optimize time efficiency by leading content moderators directly to automatically highlighted points in a video data set). On the other hand, it not only seeks to address content that has already been labelled by means of crowd work, but also to assess new image material, for which only certain image content segments are found in corresponding blacklist databases. The aim is to use algorithms to reliably detect the faces of notorious terrorists or just the relevant logos of radical groups, even in newly generated and previously unmoderated content, to identify them as known patterns, and if necessary to flag them up. Only then do they switch from the exclusively automated processes of the invisible visual culture into a "human-readable form" (Paglen 2016), and are pictorially materialized, as they still need to be subjected to human image checking. Humans are responsible for decision-making, and must make the final judgement about whether this is legitimate news content, a parody, or terror propaganda.

Image Data as Training Data

While the image-related calming in these cases is focused on channeling and filtering "human-to-human visual culture," there are quite different processes taking place at the same time in the expanding parallel world of invisible visual culture, processes that are not directly aimed at limiting or regulating the volume of image data circulating in pictorial form. Adrian MacKenzie and Anna Munster have shown, for example, that images on social media platforms circulate independent of their pictorial manifestation, and lead operative parallel lives, as it were: on the one hand as images that are distributed on user interfaces, are perceptible for human actors, and in some cases are open to interaction and can take on communicative functions, and on the other hand as "training data," used to operationalize algorithmic learning processes. As such they are inevitably opaque, nongraphic, nonvisual, but in a sense they do contribute—as a feedback loop—to the concrete phenomenality of visual culture (Mackenzie and Munster 2019).

The phenomenon referred to by Mackenzie and Munster as "the invisuality of platform seeing" is based on the observation that images on digital platforms, which are, for example, uploaded to Facebook, Instagram or YouTube by users and are then more or less freely available to the commercial operators as vast repositories of image data, are diverted into an invisible visual culture. Here, the aggregated image data is processed on various levels as sensory input—as technological perceptions serving to train and stabilize processes of algorithmic "platform seeing." This is not just about the use of these continually re-aggregated image repositories to optimize already popular image classification applications such as facial and object recognition, but also about the fundamental significance of image data sets for the operation and evolution of artificial neural networks:

> These forms of AI are driven by deep learning architectures, in which images seem to function merely as a communicative baseload that might power new "visions" for automated and autonomous decision-making and task performance. Ever since Google loaded 10 million YouTube thumbnails of cats into its neural network architecture in 2012, the conception that what an AI needs to learn to

> become more proficient is "more (image) data" has taken hold. . . . a new mode of nonrepresentational observation has become ascendant that we propose is invisual. Here, observation operates in and through the image but is not of the order of the visual. (Mackenzie and Munster 2019, 7)

When the visual information circulating on the platform interfaces as pictorial output is diverted into the invisible visual culture and becomes operable there as aggregated image data, it enters into complex computational circuits, feeding, for example, recognition algorithms such as Facebook's DeepFace or program architectures such as Google's TensorFlow, which is concerned, among other things, with the analysis of street signs in image data from Google Street View (Wojna et al. 2017). In the context of machine learning, these image data are processed not as representations but in a strictly relational way, and they feed back, not least, into algorithmic processes of image optimization:

> The techniques of training such models have intricate and heavily mathematical underpinnings, but nearly everything that happens in the construction of such models can be understood as reversing the flows of image production that have defined visuality in recent decades. Instead of generating images, these models observe images, they construct diagrammatic abstractions of features common in images, and gather these localized abstractions into predictive statements that can be operationalized as actions in the world: "place a black stone at g9 (row 7, column 9)". We might think of the calculative observation of image collections as a generalized visuality since . . . the highly variable data power consumption of a data centre, edge detection for enhancement of digital photographs, or a series of moves in a game of Go or Montezuma's Revenge all operate according to the same logic: a large image collection allows a model to be trained given the computational capacity of GPUs. (Wojna et al. 2017, 17–18)

In this context, feedback loops that retroact on registers of visual culture are most likely to emerge in situations where the learning processes retroact on image generation, that is, where they operate predictively. Hito Steyerl has spoken in this respect of "relational photography":

> By comparing what you and your network already photographed, the algorithm guesses what you might have wanted to photograph now. It creates the present picture based on earlier pictures, on your/its memory. . . . The picture might show something unexpected, because it might have cross-referenced many different databases: traffic control, medical databases, frenemy photo galleries on Facebook, credit card data, maps, and whatever else it wants. [Computational photography] is not only relational but also truly social, with countless systems and people potentially interfering with pictures before they even emerge as visible. (Steyerl 2017, 29–30)

What future images are possible, or more precisely, statistically probable, what pictorial "exteriors" actually emerge from the trillions of images making up the visual capture in the everyday digital life of the present, is based, from this perspective, to a considerable extent on feedback loops, which connect visible with invisible visual culture, pictorial interfaces with calm background activities. Insofar as algorithmic processes of scanning, recognition and filtering govern digital image distribution, and image-predicting processes feed into image-generating processes, calmly circulating image data and virally circulating image forms are not so much opposites as two cooperative sides of the same digital media culture.

Bibliography

Augusto, Juan Carlos, Hideyuki Nakashima, and Hamid Aghajan. 2009. "Ambient Intelligence and Smart Environments: A State of the Art." In *Handbook of Ambient Intelligence and Smart Environments*, edited by Juan Carlos Augusto, Hideyuki Nakashima, and Hamid Aghajan, 3–31. New York: Springer.

Drucker, Johanna. 2014. *Graphesis: Visual Forms of Knowledge Production*. Cambridge, MA: Harvard University Press.

Elegant, Naomi Xu. 2019. "The Internet Cloud Has a Dirty Secret." *Fortune*, September 18, 2019.

Ensmenger, Nathan. 2018. "The Environmental History of Computing." *Technology and Culture* 59 (4): 7–33.

Ernst, Wolfgang, Stefan Heidenreich, and Ute Holl. 2003. "Editorial: Wege zu einem visuell adressierbaren Bildarchiv." In *Suchbilder: Visuelle Kultur zwischen Algorithmen und Archiven*, edited by Wolfgang Ernst, Stefan Heidenreich, and Ute Holl, 7–15. Berlin: Kadmos.

Farid, Hany. 2016. *Photo Forensics*. Cambridge, MA: MIT Press.

Gabrys, Jennifer. 2016. *Program Earth: Environmental Sensing and the Making of a Computational Planet*. Minneapolis: Minnesota University Press.

Gershgorn, David. 2017. "It's not About the Algorithm: The Data that Transforms AI Research and Possibly the World." *qz.com*, July 26, 2017. https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/ (accessed May 3, 2020).

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

Greenfield, Adam. 2006. *Everyware: The Dawning Age of Ubiquitous Computation*. London: Pearson Education.

Hand, Martin. 2012. *Ubiquitous Photography*. Cambridge: Polity Press.

Hoelzl, Ingrid, and Rémi Marie. 2017. "From Softimage to Postimage." *Leonardo* 50 (1): 72–73.

Hosseini, Hossein, and Radha Poovendran. 2017. "Deep Neural Networks Do Not Recognize Negative Images." *arXiv.org*, March 20, 2017. https://labs.ece.uw.edu/nsl/papers/hossein_2017.pdf (accessed May 2, 2020).

Jurgenson, Nathan. 2019. *The Social Photo: On Photography and Social Media*. London: Verso Books.

Kane, Gerald C., and Alexander Pear. 2016. "The Rise of Visual Content Online." *sloanreview.mit.edu*, January 4, 2016. https://sloanreview.mit.edu/article/the-rise-of-visual-content-online/ (accessed May 3, 2020).

Kittler, Friedrich. 2010. *Optical Media: Berlin Lectures 1999*, translated by Anthony Enns, Cambridge: Polity Press.

Luo, Hong et al. 2019. "Smartphone-Based Blood Pressure Measurement Using Transdermal Optical Imaging Technology." *Circulation: Cardiovascular Imaging* 12 (8). https://www.ahajournals.org/doi/10.1161/CIRCIMAGING.119.008857 (accessed May 3, 2020).

Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice*. Cambridge, MA: MIT Press.

Mackenzie, Adrian, and Anna Munster. 2019. "Platform Seeing: Image Ensembles and their Invisualities." *Theory, Culture & Society* 36: 3–22.

Mirzoeff, Nicholas. 2016. *How to See the World: An Introduction to Images, from Self-Portraits to Selfies, Maps to Movies, and More*. New York: Basic Books.

Mitchell, W.J.T. 2010. "Image." In *Critical Terms for Media Studies*, edited by W.J.T. Mitchell, and Mark B.N. Hansen, 35–48. Chicago: University of Chicago Press.

Mumford, Lewis. 1952. *Art and Technics*. New York: Columbia University Press.

Paglen, Trevor. 2016. "Invisible Images (Your Pictures Are Looking at You)." *The New Inquiry*, December 8, 2016. https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/ (accessed May 3, 2020).

Pias, Claus. 2003. "Das digitale Bild gibt es nicht: Über das (Nicht-)Wissen der Bilder und die informatorische Illusion." *zeitenblicke* 2 (1). http://www.zeitenblicke.de/2003/01/pias/pias.pdf (accessed May 3, 2020).

Rifkin, Jeremy. 2015. *The Zero Marginal Cost Society: The Internet of Things, the Collaborative Commons, and the Eclipse of Capitalism*. London: Palgrave Macmillan.

Rothöhler, Simon. 2018. *Das verteilte Bild: Stream—Archiv—Ambiente*. Paderborn: Fink.

Rothöhler, Simon. 2018b. "Informationen, die Bilder haben: Zur Moderierbarkeit von visuellem Content." *Zeitschrift für Medienwissenschaft* 19 (2): 85–94.

Schneider, Birgit. 2009. "Wissenschaftsbilder zwischen digitaler Transformation und Manipulation: Eine Anmerkung zur Debatte des 'digitalen Bildes.'" In *Logik des Bildlichen: Zur Kritik der ikonischen Vernunft*, edited by Martina Heßler, and Dieter Mersch, 188–200. Bielefeld: transcript.

Steyerl, Hito. 2017. *Duty Free Art: Art in the Age of Planetary Civil War*. London: Verso.

Veel, Kristin 2012. "Calm Imaging: The Conquest of Overload and the Conditions of Attention." In *Throughout: Art and Culture Emerging with Ubiquitous Computing*, edited by Ulrik Ekman, 119–132. Cambridge, MA: MIT Press.

Virilio, Paul. 1994. *The Vision Machine*, translated by Julie Rose. London: British Film Institute.

Weiser, Mark. 1999. "The Computer of the 21st Century." *ACM SIGMOBILE Mobile Computing and Communications Review* 3. https://www.ics.uci.edu/~corps/phaseii/Weiser-Computer21stCentury-SciAm.pdf (accessed May 3, 2020).

Weiser, Mark, Rich Gold, and John Seely Brown. 1999. "The Origins of Ubiquitous Computing Research at PARC in the Late 1980s." *IBM Systems Journal* 39 (4). http://www.cs.cmu.edu/~jasonh/courses/ubicomp-sp2007/papers/03-weiser-origins.pdf (accessed May 3, 2020).

Wojna, Zbigniew et al. 2017. "Attention-Based Extraction of Structured Information from Street View Imagery." *arXiv.org*, August 20, 2017. https://arxiv.org/abs/1704.03549 (accessed May 3, 2020).

Notes

1 "[dass] zwischen den endlosen Ziffernkolonnen und den Gestalten, die ein menschlicher Blick erkennt, . . . eine Lücke [gähnt]. . . . Auf der einen Seite stehen Rohdaten, die Bilder als Felder farbiger Pixel kodieren; auf der anderen Seite eine Wahrnehmung, die nicht anders kann, als etwas zu sehen: Gesichter, Personen, Räume, Gegenstände. Nur langsam beginnen Prozesse digitaler Sehmaschinen, Bildobjekte im Nebel der Zahlenfelder zu finden."

2 "Das Spezifische an digitalen Bildern liegt in ihrer Operationalität und Prozessierbarkeit."