

Sven Quadflieg / Klaus  
Neuburg / Simon Nestler

## (Dis)obeying Algorithms?

Introductory Thoughts  
on the Power of  
Algorithms and the  
Possible Necessity of  
Resisting it

We live in a time when terms such as digital transformation are constantly being used and in which—as the series in which this book is published also indicates—people often talk about our digital society. Data and algorithms, it seems, have a great influence on social developments and they are often given so much importance that they are described as shaping an era. This is reflected in the title of this book, which also addresses the concept of a digital society and refers to algorithms and data. At the same time—it should be noted at the outset—the technical aspect of this phenomenon is not of significant importance in the following pages. Of course, for reading the subsequent contributions it would be useful to understand how algorithms work and why data has such a great influence on their functionality—but it will not be explained in depth at any point in the book. This is because, instead of technical details, this book only considers the impact of this technical structure on those who consciously or unconsciously interact with it: The humans and the society they form.

In the context of digital transformation, people and computers are entering into a fascinating symbiosis, something which Joseph Carl Robnett Licklider predicted as early as the 1960s, when he said that growing computing capacities would lead to new ways of using computers. The very notion of computer use is exciting here, as we probably associate it with false images; after all, computers are ubiquitously integrated into our physical environment and not even always recognizable as computers. This new kind of computer use is all-encompassing: It affects our private lives as well as major societal developments—computers and the algorithms that work within them have an impact on our lives whose dimension is almost impossible to grasp.

This influence naturally takes up a considerable amount of space in this book—however, the title already indicates that we are most interested in strategies for how individuals, groups, or even civil society in general can break this influence through forms of disobedience. In the title of the book and its implied dichotomy, we want to illustrate the complexity of the topic: Accepting algorithms is automatically a form of obedience—it can only be actively resisted—and every form of obedience should always provoke a debate about disobedience and gathering the power to question it. After all, many of the achievements of today's society are based on active disobedience or resis-

tance.<sup>1</sup> The problem, as Howard Zinn provocatively put it, is not disobedience, but obedience.<sup>2</sup>

Disobedience, or civil disobedience, often has a moral level: For the most part, acts of civil disobedience are called for to remedy an existing injustice or to remove oneself as an active person from a political or social process and thus reduce one's own guilt. For example, Henry David Thoreau coined the term civil disobedience by arguing in his 1849 essay that one should bow to the law of the state only if it is consistent with one's own moral values. Thoreau argued that as a citizen he did not want to be complicit in the war against Mexico and the slavery that followed<sup>3</sup>, mitigating his guilt by refusing to pay taxes. Thoreau thus not only shaped the concept of civil disobedience, but also inspired many subsequent civil rights movements and civil rights activists (including Martin Luther King<sup>4</sup>)—despite his libertarian stance. Thoreau's theoretical statements are based on his practical actions and experiences—at one point he did in fact refuse to pay a tax payment and was imprisoned for a short period of time. Thus, Thoreau can already be counted as one of numerous examples of applied disobedience (whether his

- 1 David Graeber sets this out impressively: Laws are legitimized—depending on the form of government—by a constitution, which in turn was legitimized by the people. Using the examples of the USA and France, Graeber reminds us that an act of then illegal violence brought the people into the situation of being able to legitimize a constitution in the first place. Laws were formed out of resistance. Cf. Graeber, David: *The Democracy Project. A History. A Crisis. A Moment*, London 2014, p. 237–239.
- 2 He writes: "As soon as you say the topic is civil disobedience, you are saying our problem is civil disobedience. That is not our problem ... Our problem is civil obedience. Our problem is the numbers of people all over the world who have obeyed the dictates of the leaders of their government and have gone to war, and millions have been killed because of this obedience. [...] Our problem is that people are obedient all over the world, in the face of poverty and starvation and stupidity, and war and cruelty. Our problem is that people are obedient while the jails are full of petty thieves, and all the while the grand thieves are running the country. That's our problem." Zinn, Howard: *The Zinn Reader: Writings on Disobedience and Democracy*, New York City 2011, p. 405.
- 3 In 1849, he wrote for example: "In other words, when a sixth of the population of a nation which has undertaken to be the refuge of liberty are slaves, and a whole country is unjustly overrun and conquered by a foreign army, and subjected to military law, I think that it is not too soon for honest men to rebel and revolutionize." Thoreau, Henry David: *Resistance to Civil Government*. In: Myerson, Joel (Ed.): *Transcendentalism. A Reader*. Oxford 2000, p. 546–565. Here: p. 550.
- 4 Cf. Powell, Brent: Henry David Thoreau, Martin Luther King Jr., and the American Tradition of Protest. In: *OAH Magazine of History* 9, no. 2, 1995, p. 26–29.

act was actually a justified example of civil disobedience is still a matter of debate<sup>5</sup>). Prominent events followed on different scales: Numerous major political transformations of the 20th and 21st centuries were based on strategies of civil disobedience—for example, the Indian independence movement—and various resistance and protest movements of different kinds—from Solidarność to Otpor!—adopted strategies of civil disobedience to protest against governments (at this point we should also refer to Gene Sharp, whose guidance on nonviolent resistance inspired protest movements all over the world<sup>6</sup>). The moral level of civil disobedience is an important point here that is always a focus of discussion—for example, in John Rawls<sup>7</sup> or Hannah Arendt<sup>8</sup>—and certainly evokes intriguing questions in the context of this book: What legit-

5 Cf. Gros, Frédéric: *Disobey. The Philosophy of Resistance*. London/New York 2020, p. 126–127.

6 Cf. Sharp, Gene: *From Dictatorship to Democracy*, Boston 2010.

Sharp's influence has already been the subject of extensive media coverage. See, for example: Thumfart, Johannes: *Der Demokrator*. In: *Die Zeit*, Nr. 10/2011, <https://www.zeit.de/2011/10/Gene-Sharp> (April 11, 2021).

7 Vgl. Rawls, John: *A Theory of Justice*, Cambridge 1999.

8 Vgl. Arendt, Hannah: *Ziviler Ungehorsam*. In: *Id.*: *In der Gegenwart. Übungen zum politischen Denken II*, Edited by Ursula Ludz, München 2000, p. 283–321.

imizes a form of resistance?<sup>9</sup> What can be described as moral justification or even as duty?<sup>10</sup> And what are the limits of this justification?<sup>11</sup>

The questions about legitimizing possible acts of resistance (the dialectic of [dis]obedience is discussed in Florian Arnold's contribution right at the beginning of this book) take on a new quality in times of the digital transformation mentioned at the beginning. This is due to the lack of transparency (which will be discussed later) and requires its own nuanced consideration. In this book, very different perspectives are brought together in order to open up a broad discursive field. This approach is illustrated in the subtitle: *Perspectives on the Power of Algorithms and Data*. The very first word is itself up for debate, as it is not altogether precise: What appears in the following pages represents only an academic, Eurocentric subset of possible perspectives on a topic that is relevant almost everywhere, but is of course evaluated differently

- 9 It should be noted that groups outside the democratic spectrum also repeatedly invoke morality to legitimize their acts of civil disobedience. In Germany, the Reichsbürger movement could serve as an example here, as could the right-wing extremist Identitarian movement, which has achieved relatively great media attention in German-speaking countries with acts of civil disobedience. In doing so, it repeatedly presents itself as a morally-driven movement on the one hand, and refers to strategies for nonviolent resistance on the other. Martin Sellner, probably the most important leader of the Identitarian movement in German-speaking countries, quotes Gene Sharp or Srdja Popovic, the co-founder of the Serbian protest movement Otpor! in his writings. This illustrates the complexity of moral justification in this context.
- 10 These questions naturally also apply to protest in the digital society mentioned above. Prominent and controversial examples include the Internet activist Julian Assange, who uses his platform WikiLeaks as a (digital) tool to draw attention to grievances, and the whistleblower Edward Snowden, whose revelations have shaken confidence in digital communication forever. In the public debate, the spectrum of opinions about the aforementioned activists ranges from almost cult-like veneration on the one hand to contempt as "lawbreakers" on the other. Cf. Kleger, Heinz; Makswitat, Eric: Digitaler Ungehorsam. Wie das Netz den zivilen Ungehorsam verändert. In: FJ SB 4/2014, p. 8–17, <http://forschungsjournal.de/node/2654> (July 1, 2021).
- 11 Defining these limits seems easier when a system of government differs greatly from current values and norms: In totalitarian systems such as National Socialism, resistance becomes, according to contemporary values, a civic duty. In modern societies, on the other hand, democratic mechanisms are generally provided for citizens to engage in critical discourse and civic or political participation, so that resistance to state power is also directed against a democratically legitimized community. The question of justification thus becomes an even more complex one. Cf. Braune, Andreas: Zur Einführung: Definitionen, Rechtfertigungen und Funktionen politischen Ungehorsams. In: Id. (Ed.): *Ziviler Ungehorsam. Texte von Thoreau bis Occupy*, Stuttgart 2017, p.9–38.

by individuals in other cultural, geographical, or social contexts. Algorithmic power is a global phenomenon and, depending on context and perspective, a global problem, and this book can, of course, only examine a small corner of it. Algorithms are often discriminatory—particularly through the data with which they are trained—and our view of the phenomenon comes from a privileged position in that we are not usually victims of that discrimination. While some contributions at least address discrimination in the context of sexism and perhaps classism (though here not from the perspective of the victim), there is a lack of consideration of discrimination, at least in the context of racism<sup>12</sup>, so the perspectives mentioned here are fairly limited. This is particularly relevant because the book aims to demonstrate that algorithmic power does not, of course, act uniformly across society. Rather, it acts as an amplifier of the existing power structure—it repeatedly displays aspects of group-focused enmity when it makes sexist or racist decisions, for example.

But what exactly does power structure mean in the context of this book? As already indicated, algorithms shape human behavior on an individual and societal scale: They influence not only our perception of the world, but also the well-being and social interaction of us as their users. They act in the political context, create social realities, and intervene in concrete social situations—for example, in public spaces. Algorithms are thus not to be regarded as a purely technical structure, but as a social phenomenon. They embed themselves—currently still very subtly and with little transparency—in our political and social system: In digitized contexts, they determine what we read, what we consume, whom we perceive, whom we ignore, what we like, what we hate, whom we love: in short, how we live (and perhaps even how we die—after all, weapons also exist that autonomously kill people based on algorithms<sup>13</sup>). These effects of algorithms on our lives have been the subject

- 1.2 There are innumerable forms of hostility against groups, so of course other discriminations could be mentioned as well. But racism in particular is of course an important point of discussion in the context of algorithmic discrimination, which is why we highlight it here and elsewhere in the introduction. But of course, racism is also of particular importance in the context of intersectionality. We cannot duly represent an explicitly intersectional perspective with our book either.
- 1.3 Only recently a UN report triggered a discussion in the media about whether autonomous weapons had in fact already been deployed. Cf. Merlot, Julia: Autonome Waffe könnte Menschen erstmals eigenständig angegriffen haben. In: Der Spiegel, 2.6.2021, <https://www.spiegel.de/wissenschaft/technik/autonome-waffe-koennte-menschen-erstmals-eigenstaendig-angegriffen-haben-a-ad06b93d-191c-4a5f-b4ae-853e7a7a177d> (June 26, 2021).

of much discussion: Lev Manovich, for example, outlines the impact of algorithms on our cultural development and their influence on our aesthetic perception by suggesting what we should consume.<sup>14</sup> And without wanting to judge, we obey already because we let them recommend things to us by giving away our data and accepting suggestions based on it.<sup>15</sup> It is difficult to attribute agency to algorithms, since they are first and foremost merely mathematical representations of facts and processes.<sup>16</sup> Nevertheless, they can be seen as a power structure, even if, strictly speaking, it is not algorithms that exercise power—rather, power is exercised with them.<sup>17</sup> We believe, however, a legitimate form of simplification is valid here: Since social actions, following Bruno Latour, are ultimately always embedded in a network of things, and artifacts are always embedded in a network of actions<sup>18</sup>, by power structure

<sup>14</sup> He writes: "[...] recommendation engines suggesting what we should watch, listen to, read, write, or wear; devices and services that automatically adjust the aesthetic of captured media to fit certain criteria; software that rates the aesthetic quality of our photos, etc." Manovich, Lev: *AI Aesthetics*, Moscow 2018, p. 19–20.

<sup>15</sup> And at the same time, of course, they are changing the nature of cultural creation, because in the context of music, for example, it seems increasingly important that a song be algorithm-compatible.

<sup>16</sup> Cf. Reicherts, Jo: *Von Menschen und Dingen. Wer handelt hier eigentlich?* In: Poferl, Angelika; Schröer, Norbert (Eds.) *Wer oder was handelt? Zum Subjektverständnis der hermeneutischen Wissenssoziologie*, Wiesbaden 2014, p. 95–120. Here: p. 111–113.

<sup>17</sup> In his observations on the phenomena of power, sociologist Heinrich Popitz describes, among other things, a data-setting power through technical action. Even if Popitz does not explicitly refer to algorithms, these observations can be transferred very well: "In changing the object world, we set 'data' to which other people are exposed. We exercise a kind of materialized power, a data-setting power, in which the effect of the powerholder over the power-affected is mediated by objects. This effect can be unintentional, accidental, not predictable, or planned and deliberate." (translation by the authors) Popitz, Heinrich: *Phänomene der Macht*, 2. Auflage, Tübingen 1992, p. 167.

<sup>18</sup> On the role of objects as agents of agency, Bruno Latour has made an interesting contribution with his Actor-Network-Theory: "If action is restricted a priori to what 'intentional,' 'meaningful' people do, it is hard to see how a hammer, a basket, a door closer, a cat, a rug, a mug, a list, or a pendant could act". Latour, Bruno: *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford 2012, p. 71. This approach becomes even more relevant the more objects are equipped with "intelligent" capabilities. As a result, every everyday object could in future theoretically be part of a larger network, so that our everyday actions will be co-determined by an unknown number of "actors" who make data available or collect it and reuse it in other contexts. These networks are in principle all-pervasive ("evasive"), untransparent ("opaque") and exceed in their complexity what can be realized in the situation. Nevertheless, they structure our actions. Cf. Pranz, Sebastian: *Der Berliner*

in this context we mean the influence that algorithms exert on individual and social life<sup>19</sup>. This immediately leads to a major problem: algorithms often evade this discussion. It is precisely in the context of the algorithmic *black box* that the question of power urgently needs to be discussed—or, to be more precise, the question of responsibility. Andreas Matthias describes this as a “responsibility gap”: We find ourselves in a situation in which autonomous machines based on algorithms make decisions whose decision-making path (also for future decisions) disappears into the deep black of the black box. It thus raises not only the question of who exercises power at this point, but also the moral question of who has the responsibility for this power.<sup>20</sup>

A list of examples of this algorithmic power could almost be endless—our private moments in particular are shaped by it. In a digital and datafied world, algorithms have an influence on our well-being—for example, through social media—and on the way we perceive the world; and it quickly becomes clear that algorithms also have a political dimension. The influence of algorithms on a large scale has been detailed in the context of the 2016 US election campaign<sup>21</sup>, but the phenomenon can be described much more broadly. For example, Ivana Bartoletti suggests that the rise of populism can also be seen as a partial victory of algorithms over our information society—she links the electoral successes of Viktor Orbán, Jair Bolsonaro, or Rodrigo Duterte to the influence of algorithms.<sup>22</sup> It can thus be seen as an irony of history that the advancing intelligence of our technology, of all things, also seems to contribute to the political regression prevailing in some countries. Bartoletti concludes that AI has the potential to fundamentally transform our society<sup>23</sup>

Schlüssel. Bruno Latour und die Akteur-Netzwerk-Theorie. In: Döbler, Thomas; Rudeloff, Christian; Spiller, Ralf (Eds.): Schlüsselwerke der Kommunikationswissenschaft. Wiesbaden 2021, in press.

- 19 Strictly according to Latour, it would probably have to be argued that neither the algorithms, nor the programmers, nor even the users have the power, but that it is distributed dynamically in the network.
- 20 Cf. Matthias, Andreas: The responsibility gap: Ascribing responsibility for the actions of learning automata. In: Ethics and Information Technology 6, 2004, p. 175–183.
- 21 Cf. for example: Coeckelberg, Mark: AI Ethics, Cambridge/London 2020, p.100.
- 22 Cf. Bartoletti, Ivana: An Artificial Revolution. On Power, Politics and AI, London 2020, p. 14.
- 23 At this point, we must define our criticism clearly: The basic criticism that could be formulated here quickly reaches a limit. An algorithm—regardless of whether it is defined as AI or not—is first simply the subdivision of an action into a sequence of steps with the goal of solving a problem. An algorithm is therefore not per se

and links this to oppression: “And if we can understand how AI is related to oppression, then we can understand how to resist it.”<sup>24</sup>

While discussing algorithms, AI is a fascinating term, but it is deliberately omitted from the title of the book, because we include in our considerations both algorithms developed by humans and algorithms developed by computers. Computers can—without going into the many differentiations of this type of learning—develop algorithms, for example, by the computer itself learning<sup>25</sup> with a kind of neural network and a large<sup>26</sup> amount of data. In a very simplified way, the human programs the computer’s brain and the computer programs—based on data—a suitable algorithm. In the first case, people determine the behavior of the computer by programming the algorithm; in the second case, they determine this behavior primarily by selecting the data.<sup>27</sup> However, as exciting as the exact functioning of artificial intelligence may now be, there are few concrete implications that can be derived from it for the discourse in the context of this book. This book, as mentioned, is concerned neither with how algorithms work nor how they learn—but with what power structures and dependencies arise through algorithms and data in our everyday lives and in our society. The tension created by algorithmic power arises primarily from its interface with society: What is of interest to us is not the structure of algorithms, but the interaction with them or the problems created by their existence. It is not important how an algorithm is implemented, but how it acts—and above all, what the result of its existence is. However, since users usually have no knowledge at all about which parts of an algorithm come from a programmer and which parts of an

something that should be viewed critically—the criticism always refers to the actions themselves.

24 Cf. Bartoletti 2020, p. 123.

25 There are various ways in which this learning process takes place exactly, but that is not important in the context of this book.

26 Of course, large is relative and also depends on the complexity of the task. Simple image recognition—i.e. recognizing a cat in an image—an algorithm will probably be able to learn after a low four-digit number of good images.

27 Since different brains, i.e. different neural networks, are better or worse suited depending on the problem, the structure of the network also plays an important role. Since corresponding frameworks, for example Tensorflow (<https://www.tensorflow.org>), are now available, the human activity required here cannot be referred to as “programming” in the true sense—it is actually more like “configuring”. This configuration influences the structure of the computer brain, but does not directly influence the action.

algorithm have been developed with the help of a neural network, this would only distort our focus, as that discussion is on the one hand too technical and on the other hand too biased towards the perspective of computer science. Since many of the currently available books on the topic already give a lot of space to the computer science perspective, we deliberately decided not to put this perspective at the center of our discourse. The following contributions thus show perspectives from very different disciplines<sup>28</sup>—and our desire for multi-perspectivity at this point outweighs the technical view of how algorithms work.<sup>29</sup>

The perception of such algorithms is always also characterized by the lack of transparency, because it is not always communicated when and why algorithms are used. It also usually impossible to work out how and why an algorithm makes a decision, which reinforces the aforementioned problem of responsibility and the general moral dimension. At the same time, it is interesting to note that there is a relatively high level of trust in algorithms<sup>30</sup>—which seems absurd in view of their opaque decision paths and the countless reports of algorithm failures. We perceive algorithms as cold, mechanical, logical, deterministic systems that provide the only correct answer after fully considering all possible options. In the end, the decisive thing is not how an algorithm arrives at its results: What is important is that algorithms are quite prone to error and sometimes simply want to make statements about the future based on bad data<sup>31</sup> from the past (which makes them sexist or racist, for example) and that humans—as Hannah Fry, among others, ex-

28 However, the limits of a complete mapping of this range are then revealed by our reflection on the “perspectives”.

29 Thus, there are formulations in the book that deviate from our understanding of an algorithmic mode of operation, and perhaps can also be criticized in the strict sense of computer science. More important to us, however, seems to be the implications for the matter at hand. The tension between “disobey” and “obey” is not only about how algorithms act, but, if we follow the principles of human-computer interaction and interaction design, about how people perceive these algorithms; thus, different perceptions are explicitly part of the overall perspective.

30 For example, a recent poll had 51% of respondents answering that they “support reducing the number of national parliamentarians and giving those seats to an algorithm,” no doubt in part because there is not much understanding of their low performance among the general population. See: <https://docs.ie.edu/cgc/IE-CGC-European-Tech-Insights-2021-%28Part-II%29.pdf> (June 15, 2021).

31 Bad is of course a complex word in this context: What is bad data? It can simply mean data that is not suitable for training an algorithm because it is incomplete or imprecise. Of course, the word bad also implies that the data does not reflect the

plains in detail—often do not question these results<sup>32</sup>. Indeed, the question of the function and precision of algorithmic decision-making leads to a very fundamental question: Why do we obey algorithms and computers at all?<sup>33</sup> Even filter bubbles generated by algorithms are now perceived by certain political movements to be more objective than content curated by journalists.<sup>34</sup>

In the early days of computer technology, when for example Alan Turing defined his abstract Turing machine, the computer behaved analogously to a human, but ideally infallible, mathematician. If the computer was wrong, it could be traced back to a programming error or to a technical defect. But under conditions characterized by uncertain and incomplete information as well as by scarce time resources, two factors are often decisive: On the one hand, a fast approximation to the correct solution is often more important than the time-consuming, exact calculation of a solution, and on the other hand, thinking in terms of correct and incorrect solutions is in itself inappropriate. In addition, there are no exact questions in communication with people or in the search for information, so that the computer cannot provide exact answers in the mathematical sense either. To conclude the brief discourse on algorithms, the lack of transparency still needs to be addressed, because this aspect is often perceived as a central deficit of algorithms generated by artificial intelligences. But here, too, our look at social implications, our shift from white box to black box, shows that this aspect applies to all algorithms. We do not know the code or the exact workings of the algorithms that dominate our everyday lives.<sup>35</sup>

complexity of reality for various reasons, or that these data represent an unfair status quo.

- 32 Cf. Fry, Hannah: Hello World. Was Algorithmen können und wie sie unser Leben verändern, Bonn 2019, p. 184–201.
- 33 In fact, the interesting question is: When exactly does acceptance tip over? In which context do users accept obvious errors? From the world of literature, we know the powerful description of Winston Smith's horror in 1984 as he had to accept an obviously wrong result as the new truth: That 2 plus 2 equals 5. But is that so outlandish? Why do people drive their cars into rivers just because their navigation system says so?
- 34 Cf. for example: Zweig, Katharina; Deussen, Oliver; Krafft, Tobias D.: Algorithmen und Meinungsbildung. In: Informatik Spektrum 40, 2017, p. 318–326.
- 35 For example, the cheat devices uncovered during the so-called Dieselgate scandal in Germany clearly show: From the outside view of the driver, it is not apparent how the exhaust gas purification algorithm works in detail. But even the authorities responsible for testing have not noticed the exact way the algorithms work over a long time. The

This brings us to the last term in the subtitle: Data. A joke that circulated on Twitter a long time ago went like this: “What idiot called it ‘machine learning’ instead of ‘bias automation’”?<sup>36</sup> Learning processes such as machine learning are based on data, from which specific behaviors are then extracted—the quality of the learning thus depends crucially on the quality of the data used. Biases in computer systems are a complex issue—and they are not only relevant in the context of AI. Batya Friedman and Helen Nissenbaum, for example, distinguished between different types of bias: First, there are preexisting biases, i.e. existing biases that are adopted consciously or unconsciously.<sup>37</sup> Friedman and Nissenbaum also refer to technical biases—biases that arise, for example, from technical constraints or technical considerations<sup>38</sup>—and emergent biases, which can arise as a “result of changing societal knowledge, population, or cultural values” over time.<sup>39</sup> Biases in AI are an important topic to keep focus on: Since the data with which the algorithm is trained is rarely neutral<sup>40</sup>, its actions afterwards are not either—often they only recreate an unjust past<sup>41</sup> or a (discriminatory) status quo, thereby perpetuating inequalities<sup>42</sup>. Algorithmic biases are thus ultimately

question of the extent to which those responsible in the companies were familiar with how the algorithms worked has not yet been conclusively clarified.

- 36 Tracing authorship in such jokes is difficult. We refer to the tweet of the Twitter user @fasterthanlime from 2017. See: <https://twitter.com/fasterthanlime/status/868840530813353985?s=20>.
- 37 Here, we should also briefly refer to the programmers, whose role in the process should certainly be discussed. The reference to problematic data is not sufficient to clarify the responsibility for non-biased systems.
- 38 They use the example of a limited screen size that lets the algorithm make a selection.
- 39 Cf. Friedman, Batya; Nissenbaum, Helen: Bias in computer systems. *ACM Trans. Inf. Syst.* 14, 3 (July 1996). p. 330–347. DOI: <https://doi.org/10.1145/230538.230561>.
- 40 The exciting question is: Is there such a thing as neutral data? It is very likely that large amounts of data can at least be assumed not to represent a balanced picture of the world. Above all, however, algorithmic decisions should be based exclusively on aspects that are thematically relevant.
- 41 The reference to discriminatory “redlining” is certainly relevant here. Especially in the context of the city and algorithms, new forms of this practice can of course emerge here. Cf. for example: Safransky, Sara: *Geographies of Algorithmic Violence: Redlining the Smart City*. *Int. J. Urban Reg. Res.*, 44, 2020. p. 200–218. DOI: <https://doi.org/10.1111/1468-2427.12833>.
- 42 When algorithms are trained with data, they naturally also adopt the status quo of this data. For example, if images, texts, or the like show primarily white men in certain

a data problem: Large amounts of data often reflect an unjust world by reproducing a white male heteronormativity (described in this book in Katja Dill's contribution).<sup>43</sup> This is another reason why algorithms, for example, can be quickly attributed to racist behavior, as Ruha Benjamin<sup>44</sup> and Safiya Umoja Noble<sup>45</sup>, among others, have demonstrated impressively. In the context of machine learning/deep learning, the biggest problem is likely to be as outlined above: it is almost impossible to tell whether an algorithm is currently exhibiting discriminatory behavior, not only for users but increasingly also for creators.<sup>46</sup>

At the same time, however, data naturally also constitute their own (new) form of power structure: The control of large volumes of data, or the ability to evaluate complex and good data structures or to recognize patterns in them, change established power relationships and supplement (or partially replace) the means of production as the basis of power and the economy—here we can refer to Karl Marx.<sup>47</sup> In the sense of an intangible economy, companies are taking on an enormous role in the global economy on an intangible

positions of power, an algorithm can understand this as an unchanging status quo and try to reproduce it. Prominent examples of this kind are the much-discussed sexist recruiting AI of Amazon but also the sexism of Google Translate. Cf. Dastin, Jeffrey: Amazon scraps secret AI recruiting tool that showed bias against women. In: Reuters, 10. Oktober 2018, <https://www.reuters.com/article/amazon-com-jobs-automation-idINKCN1MK0AH> (Jun 3, 2021) and Prates, Marcelo O.R.; Avelar, Pedro H.; Lamb, Luís C.: Assessing gender bias in machine translation: a case study with Google Translate. In: *Neural Comput & Applic* 32, 2020. p. 6363–6381.

43 Of course, unjust data is not only a problem in algorithms: There are countless examples of how data reinforces existing privileges or disadvantages marginalized people. Cf. for example Cirado-Perez, Caroline: *Invisible Women: Exposing Data Bias in a World Designed for Men*, London 2019.

44 Vgl. Benjamin, Ruha: *Race After Technology: Abolitionist Tools for the New Jim Code*, Hoboken 2019.

45 Noble, Safiya Umoja: *Algorithms of Oppression. How Search Engines Reinforce Racism*. New York City 2018.

46 Depending on the context, checking the training data or controlling the algorithm can already be important steps. This can be explained with a simple example: If parameters in the usage of a system are changed (for example, the name, the place of residence or the gender) and the algorithm then changes its behavior, whether it should not do so, then problematic behavioral pattern could quickly be identified.

47 With regard to the significance of the social means of production and the power over them, one could of course refer directly to *Das Kapital*. Because this cross-reference would go beyond the dimensions of this introduction, we refer here to the text by Michael Seemann, who writes about digital capitalism and also refers to Karl Marx.

basis—and this is primarily due to the large quantities of analyzable data. This data generates insights into individual, societal and political behavior<sup>48</sup>, it reveals invisible connections, makes accurate forecasts—and it can be used to persuade people, influencing behavior in a targeted and large-scale way. At the same time, we seem to be unwilling to acknowledge this. As Timothy Snyder points out, cognitive dissonance is an important issue here—by insisting on being the authors of our own actions, we provide alibis for digital creatures.<sup>49</sup>

The current situation is thus that algorithms have a major impact on individual and social life—and their ubiquitous dissemination means that they touch on various disciplines in scientific discourse. These different fields of consideration also open up the field of possible strategies of disobedience and allow for a wide-ranging discussion. If we speak of disobedience in this interdisciplinary field of tension, then the vagueness of the term in our context allows for quite different interpretations of the term. It includes the fact that disobedience also refers to the regulation of algorithms, i.e. by sparking a discourse on possible the legal consequences of an algorithmic power structure—also in order to make clear that this power is something that has to be negotiated socially and politically (which is discussed in this book, for example, in the contribution by Johanna Mellentin and Francesca Schmidt and also in Matthias Pfeffer's contribution). The legal approach to algorithmic power is especially a topic of immense importance, since it is necessary, among other things, to shed legal light on the black box of algorithms—not least in order to exclude the potential discriminatory treatment of human beings (in the book, this discussion can be found, for example, in the contributions from Fabian Lütz and Victoria Guijarro Santos). Disobedience also involves a general discussion about the use of algorithms, and of course also the individual strategy for evading algorithmic power, even if algorithms are also used at different scales as instruments of surveillance. The enormous possibilities

Cf. Seemann, Michael: Eine beunruhigende Frage an den digitalen Kapitalismus. In: APuZ. Aus Politik und Zeitgeschichte 69, 2019, p. 10–15.

- 48 At this point, it should of course be pointed out that data and the use of the same through machine learning turn the consideration of the complex relationship between humans and computers, since users of services simultaneously become data suppliers and thus—as Jaron Lanier puts it—data models. Cf: Lanier, Jaron: *Who owns the future?* New York City 2013.
- 49 Cf. Snyder, Timothy: *Und wie elektrische Schafe träumen wir Humanität, Sexualität, Digitalität.* Wien 2020, p. 41.

that arise, for example, for the surveillance of a society—with algorithms tracking people and their movement patterns in public space or in concrete urban spaces—must at least generate a debate about potential disobedience strategies. This means both that there should be discussion on a societal level about the use of such strategies (as in Lotte Houwing’s contribution) and that people (as, for example, Fabian Weiss, Moritz Ahlert and Bernd Friedrich Schon suggest in their contributions) develop individual strategies for disobedience. Especially in political activism, disobedience to algorithms can be a necessary strategy that makes protests possible in the first place, both in an urban space (here we refer to Carolin Höfner’s contribution) and in the digital space that serves mobilization and exchange. The importance of the concrete urban space for the formation and execution of a political protest has been explained many times, especially in the post-Tahrir years<sup>50</sup>, which are characterized by the occupation of public squares. Countless protest groups, from Occupy to Movimiento 15M, have used this method, creating an exciting field of discussion. An algorithmic monitoring of the public space is thus also the potential control of spaces for demonstrations, which in a democratic society should at least provoke a critical discourse.

The contributions in this book oscillate between the poles of obedience and disobedience. Again and again, the aim is to show the manifold, complex dependencies between data and algorithms as part of the reality of life (in the context of work, for example, this is done in Christina Hecht’s contribution), in order to then document strategies of disobedience that can be based on this (Harald Trapp and Robert Thum do this in their contributions, in which they lay out the dependencies between digital platforms and their data and urban space and the work commissioned there). The consideration of disobedience—we emphasize once again—is not meant to serve a general critique of algorithmic processes (whose comprehensive significance for the contemporary world we are of course aware of). Nor is it meant to be understood as a legitimatization of structures (the datafied world of work, for example, can

50 A detailed note here would go beyond the scope of this introduction, because here both the detailed reference to the general importance of public space and the reference to the concrete urban space, with reference to Georges-Eugène Haussmann etc., would be important. But Tahrir Square in particular can be highlighted here, as the political events there in 2011 have influenced many subsequent protest movements. Cf.: Mohamed, Abdelbaseer A.; van Nes, Akkelies; Salheen, Mohamed A.: Space and protest: A tale of two Egyptian squares. In: SSS10: Proceedings of the 10th International Space Syntax Symposium, London, UK, 13–17 July 2015, p. 110:1–110:18.

also be viewed critically, even if there are possibilities for “resistance”). Furthermore, our discussions are also not to be understood as a direct injunction to commit acts of disobedience (we would prefer if this were not necessary), even if we attach great importance to the strategies for this. Rather, it is intended to focus on very fundamental techno-ethical questions and to initiate a comprehensive discourse, at the end of which concrete strategies of disobedience may indeed be posited. In this sense, the book’s contributions span both very concrete ideas and strategies and much more abstract considerations. The book begins on an epistemological level, where the implied problem of disobedience and obedience in the context of algorithms is laid out in various contexts, before that is discussed more empirically and analytically in the middle section of the book. From these considerations, the normative thoughts or the concrete descriptions of possible disobedience at the end of the book can then be well framed: Despite our academic context, we allow this book to end almost on an activist note. As will become apparent in the following pages, strategies for disobedience can be found on very different levels: In the individual act of resistance as well as in a broad political debate. It can concern very private areas, but equally important social issues. And it ranges from humorous interventions, which perhaps are rather meant to initiate a discussion, to strategies for ensuring survival or freedom in repressive regimes. What unites these different levels is that it is always about questioning and breaking the dominance of technologies. We hope to contribute to a discourse whose importance for a society in the process of digital transformation can hardly be overestimated.

## Literature

Arendt, Hannah: *Ziviler Ungehorsam*. In: Id.: *In der Gegenwart. Übungen zum politischen Denken II*, Edited by Ursula Ludz, München 2000, p. 283–321.

Bartoletti, Ivana: *An Artificial Revolution. On Power, Politics and AI*, London 2020.

Benjamin, Ruha: *Race After Technology: Abolitionist Tools for the New Jim Code*, Hoboken 2019.

Braune, Andreas: Zur Einführung: Definitionen, Rechtfertigungen und Funktionen politischen Ungehorsams. In: Id. (Ed.): *Ziviler Ungehorsam. Texte von Thoreau bis Occupy*, Stuttgart 2017, p. 9–38.

Cirado-Perez, Caroline: *Invisible Women: Exposing Data Bias in a World Designed for Men*, London 2019.

Coeckelberg, Mark: *AI Ethics*, Cambridge/London 2020.

Dastin, Jeffrey: Amazon scraps secret AI recruiting tool that showed bias against women. In: Reuters, 10. Oktober 2018, <https://www.reuters.com/article/amazon-com-jobs-automation-idINKCN1MK0AH> (June 3, 2021).

Friedman, Batya; Nissenbaum, Helen: Bias in computer systems. In: *ACM Trans. Inf. Syst.* 14, 3, July 1996, p. 330–347. DOI: <https://doi.org/10.1145/230538.230561>

Fry, Hannah: *Hello World. Was Algorithmen können und wie sie unser Leben verändern*, Bonn 2019.

Graeber, David: *The Democracy Project. A History. A Crisis. A Moment*, London 2014.

Gros, Frédéric: *Disobey. The Philosophy of Resistance*. London/New York 2020.

Kleger, Heinz; Makswitat, Eric: Digitaler Ungehorsam. Wie das Netz den zivilen Ungehorsam verändert. In: *FJ SB* 4/2014, p. 8–17, <http://forschungsjournal.de/node/2654> (July 1, 2021).

Lanier, Jaron: *Who owns the future?* New York City 2013.

Latour, Bruno: *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford 2012.

Manovich, Lev: *AI Aesthetics*, Moscow 2018.

Matthias, Andreas: The responsibility gap: Ascribing responsibility for the actions of learning automata. In: *Ethics and Information Technology* 6, 2004, p. 175–183.

Merlot, Julia: Autonome Waffe könnte Menschen erstmals eigenständig angegriffen haben. In: *Der Spiegel*, 2.6.2021, <https://www.spiegel.de/wissenschaft/technik/autonome-waffe-koennte-menschen-erstmals-eigenstaendig-angegriffen-haben-a-ad06b93d-191c-4a5f-b4ae-853e7a7a177d> (June 2, 2021).

Mohamed, Abdelbaseer A.; van Nes, Akkelies; Salheen, Mohamed A.: Space and protest: A tale of two Egyptian squares. In: *SSS10: Proceedings of the 10th International Space Syntax Symposium*, London, UK, 13–17 July 2015, p. 110:1–110:18.

Noble, Safiya Umoja: Algorithms of Oppression. How Search Engines Reinforce Racism. New York City 2018.

Popitz, Heinrich: Phänomene der Macht, 2. Auflage, Tübingen 1992.

Powell, Brent: Henry David Thoreau, Martin Luther King Jr., and the American Tradition of Protest. In: OAH Magazine of History 9, no. 2, 1995, p. 26–29.

Pranz, Sebastian: Der Berliner Schlüssel. Bruno Latour und die Akteur-Netzwerk-Theorie. In: Döbler, Thomas; Rudeloff, Christian; Spiller, Ralf (Eds.), Schlüsselwerke der Kommunikationswissenschaft, Wiesbaden 2021, in press.

Prates, Marcelo O.R.; Avelar, Pedro H.; Lamb, Luís C.: Assessing gender bias in machine translation: a case study with Google Translate. In: Neural Comput & Applic 32, 2020, p. 6363–6381.

Rawls, John: A Theory of Justice, Cambridge 1999.

Reichert, Jo: Von Menschen und Dingen. Wer handelt hier eigentlich? In: Poferl, Angelika; Schröer, Norbert (Eds.) Wer oder was handelt? Zum Subjektverständnis der hermeneutischen Wissenssoziologie, Wiesbaden 2014, p. 95–120.

Safransky, Sara: Geographies of Algorithmic Violence: Redlining the Smart City. Int. J. Urban Reg. Res., 44, 2020. p. 200–218. DOI: <https://doi.org/10.1111/1468-2427.12833>

Seemann, Michael: Eine beunruhigende Frage an den digitalen Kapitalismus. In: APuZ. Aus Politik und Zeitgeschichte 69, 2019, p. 10–15.

Sharp, Gene: From Dictatorship to Democracy, Boston 2010.

Snyder, Timothy: Und wie elektrische Schafe träumen wir Humanität, Sexualität, Digitalität, Wien 2020.

Thoreau, Henry David: Resistance to Civil Government. In: Myerson, Joel (Ed.): Transcendentalism. A Reader. Oxford 2000, p. 546–565.

Thumfart, Johannes: Der Demokrator. In: Die Zeit, Nr. 10/2011, <https://www.zeit.de/2011/10/Gene-Sharp> (April 11, 2021).

Zinn, Howard: The Zinn Reader: Writings on Disobedience and Democracy, New York City 2011.

Zweig, Katharina; Deussen, Oliver; Krafft, Tobias D.: Algorithmen und Meinungsbildung. In: Informatik Spektrum 40, 2017, p. 318–326.