of digital work written in a very accessible and engaging style. With its innumerable examples on roughly 200 pages, it is recommended reading for scholars as well as practitioners or the interested public.

The Invisible Force – How Algorithms Shape Society

Matthias Philipper

O'Neal, C. 2016. *Weapons of Math Destruction: How Big Data Increases inequality and Threatens Democracy*. New York: Crown Publishers.

Big Data and algorithms have become more and more present in various aspects of modern life. They are either regarded as an efficient and objective tool for solving a broad variety of problems in the working world and governance processes or as a threat to democracy, equality, and familiar ways of life. Questions on how to deal with the ongoing data revolution and its side effects are increasingly depicted and critically discussed in mass media and literature. Hence it is not surprising that numerous publications about current developments in the field have been published over the last few years in either scholarly or popular science literature. In these Big Data analyses, algorithms and computational modelling of society are discussed from the perspective of the various academic disciplines, such as computer science, philosophy, or sociology. For readers who are interested in the subject the discourse in these fields may be demanding and the analyses thus less accessible. As well as scholarly routes into this field, there is a broad range of writing more generally accessible to the public on how data shapes the everyday life of an increasing number of people worldwide. The authors of these publications are mostly science journalists and bloggers who often present a rather critical approach to the topic. Some of these writings offer curious insights on how the Data Economy works and how it develops its state-of-the-art technology. This is especially the case, when the developers of the algorithms and techniques of machine learning want to present their inside knowledge to a wide group of readers. One of these books is *Weapons of Math Destruction* by Cathy O'Neal.

Although the author, Cathy O'Neil, is not specialised in the social sciences or questions of inequality, her knowledge and thoughts about mathematical modelling, algorithms, or, as she calls them, 'Weapons of Math Destruction' (p. 3) shed an interesting light on the different effects that these systems can have on society. Her expertise in the field derives from a broad variety of different professions

and assignments that O'Neal has worked in over the course of her diverse career. Early on she turned her enthusiasm about math and numbers that had accompanied her since childhood into an academic career by obtaining a PhD in algebraic number theory, which eventually led to a tenure track professorship at Barnard College in the United States (p. 2). More interested in the possibilities of the non-theoretical use of mathematics and its application in new contexts, O'Neal soon changed the course of her career to a more practical and 'fast-paced' (p. 33) one and joined the thriving Data Economy. The conjunction between the academic world and the practical application of math in Big Data sets the tone for her book and makes her perspective particularly interesting. Throughout the first chapters of her book she characterises her different occupations as a quantitative data analyst for a hedge fund and as a data scientist in the internet economy. The focus lies on the different, predominantly disillusioning, insights about the practical use of mathematic models she gained in these fields and in her subsequent engagement against the current practice of the mathematical modelling of society. Experiencing and even cultivating the destructive potential of these models herself in her occupations, she reflects on her trajectory and the ways in which mathematical models are used in different social contexts. Her book *Weapons of Math Destruction* is the result of this investigation into the disruptive impact of algorithms on US society.

The very clear and guiding structure of the book will help readers who may lack in-depth knowledge about mathematical models and their dominant position in the US society to easily access the topic. After the Introduction, the first chapter explains what a model is and where models can be found in daily life, even outside computational systems. In the next chapter, she explains her disillusionment about the application of this technology. From this point onward she examines different areas in which Big Data, algorithms, and mathematical models are used and where the pitfalls of their usage lie. She closes the book with a conclusion and suggestions on how math and algorithms can be used in a fairer manner for the public good. The book's general tone is somewhat pessimistic, and viewed together with the book's guiding structure it underlines O'Neal's urgent call to foster public discourse and to find different ways of using Big Data.

The Introduction begins with a description of O'Neal's personal connection to math and the trajectory of her professional life. Initially enthusiastic about the possibilities and application of mathematical models, her perspective quickly shifted after she began working in the 'Big Data economy' (p. 3). She persuasively presents one of her recurring arguments against the current use of mathematical models in Big Data: the opacity of their functioning and the common belief in their infallibility. To illustrate their harmful force and her arguments, she describes the implementation of 'a teacher assessment tool called IMPACT' (p. 4) in schools in Washington DC to rate

the performance of teachers by assigning them a certain score. This score measures a student's skills in math and language skills and gives the administration the ability to identify teachers who are not performing well. Originally intended to improve the quality of teaching in Washington's schools, O'Neal shows how the assessment system focuses only on a small and insufficient variety of relevant data to measure and score the teachers' performance. Specifically, environmental variables are not considered when calculating the skills of students and therefore also the teachers' scores (p. 5). As a result, the underlying algorithm takes a small section of reality and presents it as an unquestionable instructional guideline for the administration to act on. She furthermore argues that the algorithms behind the scoring system are not designed to receive feedback (p. 7). This leads to the problem that certain issues, such as the inaccurate measurement of a teacher's performance, are not ultimately clarified. Hence the algorithm has no chance to develop new and more accurate techniques to assess the work of the teachers. It remains blind and is therefore unable to provide a reliable evaluation of a teacher's skills. She demonstrates this by presenting the case of a competent teacher who is assessed with a low score and eventually has to leave her job and then find new employment at a richer school. The two schools use different evaluation and assessment systems. The wealthier school relies on humans to rate the quality and performance of an applicant by means of thorough interviews and observations instead of using a data driven scoring algorithm. This important observation reveals a major and recurring point of criticism that O'Neal sees in the common use of Big Data: algorithms benefit the privileged over the unprivileged.

Chapter 1 lays the foundations that are important for reading the following chapters. O'Neal explains the basic concepts of models, where they can be found in daily life, and both the possibilities and the limitations of their use. The example she uses to illustrate the basal functions of mathematical models especially is baseball. She argues that the models used to describe this sport are a perfect example of a fair and especially transparent algorithm (p. 17). The data that are used to calculate the success of a certain team or player in the various situations of a game is available to anyone interested, and how the different models function is clearly visible. More importantly, the data about the different players and their skills is very accurate and open for feedback after every game. In contrast to this, algorithms in the Data Economy work with approximate values, because the reality they are trying to depict and calculate is too complex for a model to grasp. Moreover, they do not receive any external feedback, which makes them prone to blind spots and gaps in knowledge, especially when used on a large scale. According to O'Neal, many decisions made by these models are inaccurate and often discriminating. They do not question their data themselves and as a result their decisions are made on a rather vague basis that often

reflects the prejudices of its creators (p. 23). Since most commercially used models and algorithms are opaque and their functions are kept secret, it is difficult to identify their mistakes. A striking example of this that O'Neal introduces later in this chapter are the recidivism models used to determine whether a prisoner can be released from prison early or not. She explains that the models and the data collected for them is often biased by racist perspectives and prejudices, making an early release for people of colour less likely than for whites (p. 25). Moreover, the decisions based on these toxic models are widely perceived to be a dictum of objective truth that cannot be argued against, because only the programmers of the algorithm themselves actually understand what the model is doing (p. 25). At the end of the chapter, O'Neal summarises the issues presented in the example cases and traces their origins back to the use of mathematical models. She applies these characteristics to her concept of a 'weapon of math destruction' and applies it to different areas of society in the following chapters. In each of these areas she presents different examples and explains the environment in which they are in practice and the havoc they cause in these contexts. She presents a large number of different examples in each of the chapters, giving the reader a broad idea of how Big Data are used in society. The outcome in almost any case remains the same. Weapons of math destruction support social division and benefit mainly the privileged.

The central idea of chapter two is to explain the process by which O'Neal began to question the impact of Big Data use while she was working as an analyst in different areas of the US working world. After her academic career, she began working at a hedge fund and experienced the financial crisis in 2009 as a quantitative analyst. She explains here how the procedures and weapons of math destruction in the financial economy work and how she realised the faulty impact they have on society. While the hedge funds were searching for new ways to maximise their profits through implementation of algorithms and mathematical models, many people worldwide lost their jobs (p. 40). The use of weapons of math destruction in this field clearly supports her observation that only certain people actually profit from the use of these models, whereas mainly underprivileged social groups are systematically discriminated by it. Reflecting her own responsibility of working on the math behind the models, O'Neal changed her career with the intention to prevent financial weapons of math destruction from causing harm again (p. 44). She therefore started to work for a company that analyses the risk of failure in the financial economy. Her statements about the ineffectiveness of this measure are especially striking, showing that most of the inherent risks of the use of the models are ignored (p. 45). She then changed her occupation again and began working as a data scientist in e-commerce. While describing her daily tasks in this new economy, it becomes clear that the same toxic algorithms are used throughout this field as well (p. 47). These experiences led O'Neal

to the point where she grew more and more disillusioned and became involved with the Occupy Movement. Ultimately, she quit her job in order 'to investigate the issue in earnest' (p. 49).

Beginning with Chapter 3, O'Neal presents a narrower analysis of the role of the explained models in various areas of social life, from the creation of college rankings to the micro targeting of citizens in political contexts. She begins by explaining how weapons of math destruction influence college rankings and their impact on educational infrastructure. She first gives a brief outline of how college rankings were invented in the first place and how weapons of math destruction create these rankings to find suitable applicants. She discusses how the models involved focus on the characteristics that expensive private universities throughout the US already do well in, which turns the rating system into a self-fulfilling prophecy (p. 60). O'Neal makes it clear that these evaluations do not represent high quality in education, because they are unable to measure the qualities of a college convincingly (p. 55). To illustrate this, she shows how different colleges manipulate these scores and how subjective and vague some of the selected variables used to rate the colleges are. According to O'Neal, ways of achieving better scores in the rankings include, for example, lowering standards for applicants, giving graduates better grades, or accepting a relatively small number of applicants with especially high scores to increase the average performance (p. 54). She further argues that because of the power the established models have, the majority of American colleges seek to improve their scores rather than the quality of their teaching. Students face the same issue of having to try to present themselves in a way that suits the model's expectations. Again, the weapons of math destruction in this area largely benefit the already established and expensive colleges and the students who have the opportunity to make themselves more appealing to the algorithm's preferences. Privileged students are therefore more likely to apply to a well rated college and to profit from the advantages of the given rating, which in the end supports social division (p. 65).

In Chapter 4, O'Neal analyses how weapons of math destruction are used in online advertising. The goal of the applied models is to maximise sales by identifying certain demands. In order to do this, the models find '… inequality and feast on it. The result is that they perpetuate our existing social stratification, with all of its injustices' (p. 70). The algorithms are 'trained' to find certain weaknesses, the 'pain point[s]' (p. 73) of the people browsing the internet, to create personalised adverts and to exploit them. The companies behind this deliberately target the vulnerable in order to increase sales. It is not surprising that the recruiters in these businesses search for 'Welfare Mom w/Kids. Pregnant Ladies. Recent Divorce. Low Self-Esteem. Low Income Jobs. Experienced a Recent Death. Physically/Mentally Abused …' (p. 72), as O'Neal quotes. The image of the use of algorithms and the general practice

of some companies in online advertising are further supported by the large number of examples presented by the author and analysed in more detail in reference to 'for-profit colleges' (p. 81). These colleges make heavy use of the techniques that she describes in order to target students who cannot afford to attend a regular college. They promise them 'education … and upward mobility – while plunging them deeper into dept' (p. 81). O'Neal's descriptions in this chapter make it undoubtedly clear that automated systems are used on a great scale to exploit certain groups of people and to intensify social divisions.

Chapter 5 explores the impact of weapons of math destruction in the justice system. O'Neal's analysis reveals yet again how the application of algorithms on a large scale is harmful to certain groups of people. The weapons of math destruction used in this area are designed, for example, to predict the probability of crimes being committed (p. 85). In the light of her previous observations, it is not surprising that these algorithms rely on biased data to make their predictions. If a certain area shows a high number of minor crimes, the algorithms in use recommend deploying more police into this area. The high presence of police members then makes it more likely to discover even more crimes, eventually leading to the reinforcement of a police presence (p. 87). This strongly demonstrates how algorithms verify their own decision in toxic feedback loops. She further argues that the algorithms also identify the personal probability of a crime being committed. As seen before, the data used in this process are often biased by racist and stereotypical perspectives and ideas, even if the algorithm itself has to be blind with regard to ethnicity or race.

Chapters 6 and 7 show the difficulties algorithms create for people trying to find a job and job performance assessments. O'Neal's observations again reveal the problematic use of the technology, as it is mostly applied in low-wage areas of the employment market. Applicants for higher paid jobs are more likely to be reviewed by human workers in human resources departments, as they know about 'what machines appreciate' (p. 114). By passing this first obstacle in the process, misunderstandings and problems that may occur in the application procedure are far more likely to be noticed and solved in the case of applicants with a higher level of education than in the case of applicants for low-paid jobs. The algorithms that are used also show a large rejection rate for female applicants, because the weapons of math destruction calculate the probability of a person leaving the job for maternity leave or for longer periods of time, which are too long for them to be recommended as worth being invited for an interview (p. 117). O'Neal vividly describes the formation of this kind of algorithm and how discriminatory views find their way into the underlying weapon of math destruction. The assessment of scores in employee ratings also presents itself as highly biased. Again, O'Neal shows that the data used by the algorithms is insufficient to capture certain areas and skills in the work environment. The current

algorithms can, for instance, hardly measure soft skills (p. 133). This may ultimately benefit male workers over female workers. In Chapter 7, O'Neal again inspects scoring algorithms in education and examines how these affect local and national educational policies (p. 134ff). Further, she shows how personnel planning software used to maximise the profit of certain companies impacts the everyday life of families and exploits the time of low-paid workers in different areas of the working world (p. 123ff).

The destructive force of weapons of math destruction is also observable in the financial areas of social life. O'Neal analyses their impact on 'landing credit' (Chapter 8) and 'getting insurance' (Chapter 9). Decisions on credit applications have always been influenced by markers like race, class, and gender, and in case of doubt the bank clerk decided whether the applicant was credible in each individual case (p. 141). In Chapter 8, O'Neal demonstrates how different scoring algorithms have now superseded the position of bank clerks and how the use of weapons of math destruction in this context exacerbates this harmful practice. Here, as well as elsewhere, the scoring systems are opaque and offer no opportunity for feedback, making it even more difficult for minority groups to qualify for credit. Moreover, the algorithms do not elaborate on the individual risk of illiquidity. They rate applicants' solvency based on their belonging to different social groups (p. 145). Again, the use of these systems is not open to feedback and well-situated people have a better chance of avoiding contact with a weapon of math destruction when they apply for a credit. Another reason to be distrustful of the use of algorithms in this context is the economy that developed around the different scoring systems. As O'Neal shows, the personal scores are sold to other companies and interested parties for profiling (p. 148). These profiles are increasingly used to test the quality of applicants in job interviews or to decide whether a worker deserves promotion to a major position or not (p. 148). Especially alarming is the fact that Facebook patented their own technology to determine the credit rating of applicants based on the social networks they belong to on Facebook itself (p. 155). In Chapter 9, O'Neal shows how scoring algorithms assess people's behaviour and how insurance companies use the data thus generated to determine insurance premiums. She shows that health scores are increasingly generated and used in the context of work to determine whether workers have to pay additional fees for their health insurance or not (p. 175). O'Neal fears that in the near future it may even be possible that these health data could be used in job application processes or for other purposes (p. 175).

Since Cambridge Analytica massively influenced the outcome of the presidential election in the United States in 2016, it has become clear that algorithms have the potential to alter the political course of an entire country. Micro targeting allows political parties, lobbyists, and polling institutes to target individual groups and people

based on personal data to influence their opinions. In Chapter 10, O'Neal describes the algorithms behind this and observes the different consequences. She demonstrates how various companies in the communications business, especially Facebook, are able to use the data they collect to influence democratic processes (p. 181). A particularly insightful point that O'Neal makes is how the individual newsfeeds of people on social media are curated by algorithms, which thereby affect their mood and general attitude (p. 183). Although she sees no sign of misuse of this technique by the big internet companies themselves, there undeniably lies a destructive potential within the technology (p. 185).

In the Conclusion, the O'Neal discusses her general findings from each area and once again describes the universal characteristics of a weapon of math destruction and how these 'weapons' could be altered to produce fairer algorithms. She argues that the general public and the government have to take a closer look at the ways in which these systems work and at the inequality they produce. In the second part of the Conclusion, O'Neal presents her suggestions about what can be done to improve the use of Big Data. In her opinion, one way of achieving a better practice, besides appealing to the developers of algorithms (p. 205), would be to establish a 'regulatory system for WMDs' (p.207) and a closer auditing of algorithms by experts and the government (p. 211). She calls upon 'academic support' to train 'people with the skills to build them' to monitor the results of the use of algorithms in different areas of society (p. 211). In the end she presents examples of algorithms that were actually used to improve the living conditions of marginalised groups in the US.

A dominant theme of the book is the question of how algorithms affect the equality of distinct groups in society. Even if O'Neal's analysis remains rather superficial, it becomes clear that those who profit from the use of such models are mainly well situated, white, and male. Established to prevent biased decisions in different contexts, algorithms reveal themselves as doing the exact opposite. In almost every field presented in the book, from going to college to finding a job and making a career, it is women, people of colour, and people with a low level of education who are discriminated against by the advice given by the models. This is aggravated by the fact that the choices algorithms and mathematical models make are widely perceived as objective truth, meaning that there is no discussion about the rightness of these choices. This makes it impossible to move towards a more just use of mathematical models and Big Data. Interestingly, O'Neal does not argue against the use of Big Data in society. She rather criticises the implementation and purpose of the algorithms in practice. In her view they are 'primitive tools, which hammer complexity into simplicity' (p. 166e), but ones that could be changed to serve and benefit the public.

It has to be taken into consideration that O'Neal's book is more a work of popular

science than an in-depth sociological analysis. For her writing she relies mostly on newspaper articles and government documents, and less on scholarly publications. Nevertheless, her book remains interesting to read and opens a great range of perspectives for social scientists to further pursue. Her concept of a weapon of math destruction is interesting and comprehensibly established in the beginning, but it is not elaborated on as thoroughly as one would expect for a scientific analysis. Nevertheless, its application in various areas strikingly reveals the problematic utilisation of mathematic models in almost every context of modern life. Her strong use of many examples in each chapter and her rather non-theoretical style of writing make her book easily accessible for all kinds of readers, although the examples she presents are more likely to interest an American audience. The use of baseball and profit colleges to illustrate the models may be confusing to readers outside the United States. Also, the prevalence of mathematical models in the United States differs from the situation in other parts of the world. The book thus grants valuable insights into the possible direction of the increasing use of big data even in non-US countries. Given the book's structure, the discussion remains on a practical and comprehensive level and does not offer a scientifically detailed approach to the various presented topics and example cases. As stated above, these topics and cases represent opportunities for social scientists to raise more questions and further analyse the issues presented in the book. The book does not explore in detail the reasons why mathematical models are put into practice in the first place, and while recurring motives such as efficiency and profit maximisation are mentioned, they are not linked to the general discourse about capitalism and the neoliberal economy. For a detailed look at specific areas, O'Neal's book may therefore not be the right choice, but it offers a good overview of the destructive power of algorithms and the use of mathematical models, and it makes it remarkably clear why the topic should be of interest to the general public and the social sciences.