

**FROM UNIVERSITY RANKING TO ELECTIONS: HOW ALGORITHMS ARE CHANGING OUR LIVES<sup>1</sup>**

***DE RANKING UNIVERSITÁRIO A ELEIÇÕES: COMO OS ALGORITMOS ESTÃO MODIFICANDO NOSSA VIDA***

***DE LAS CLASIFICACIONES UNIVERSITARIAS A LAS ELECCIONES: CÓMO LOS ALGORITMOS ESTÁN CAMBIANDO NUESTRAS VIDAS***

Luciano Lopes QUEIROZ<sup>2</sup>

We currently live in a digital society. In Brazil, it is estimated that there are around 234 million smart cell phones, despite the fact that the internet connection is limited in several regions of the country. Currently, most of our activities are carried out online. We are all constantly producing data when we Google something or tag a friend in a Facebook photo. From the data, we can determine people's behavioral patterns and even measure what they are feeling and desiring. This large volume of data, associated with greater processing capacity and advances in mathematical models and algorithms, has a great impact on our lives. However, most of us know little or nothing about how algorithms work. Cathy O'Neil, in *Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy* (published in Portuguese as *Algorithms of Mass Destruction*), brings up the discussion of how mathematical models, algorithms and Big Data have been used indiscriminately. According to the book, these components are present in almost every moment of our lives in digital platforms, in employee hiring processes, in online advertising, in public policy, in finance, in the prison system and in several other examples. O'Neil demonstrates, through a series of examples throughout the ten chapters of his book, how the false idea of fairness in algorithms can lead to an increase in inequality and injustice.

O'Neil, a PhD in mathematics and former Wall Street data scientist, was able to observe during the 2008 economic crisis (subprime crisis) how mathematical models and algorithms can influence people's lives. In the first chapter, she explains what mathematical models are and how we use them in our daily lives. During high school, we learned that mathematical models

---

<sup>1</sup> Review of the work: O'NEIL, C. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy**. New York: Crown, 2016. ISBN: 978-0553418811. Published in Portuguese as *Algoritmos de Destruição em Massa* by Editora Rua do Sabão, 2020.

<sup>2</sup> University of São Paulo (USP), São Paulo – SP – Brazil. Department of Microbiology. PhD in Biological Sciences. ORCID: <http://orcid.org/0000-0002-5260-0628>. E-mail: [lqueiroz@usp.br](mailto:lqueiroz@usp.br)

are simplifications of reality. O'Neil describes a model as nothing more than an abstract representation of some process, be it a baseball game or an oil company's supply chain. She concludes that whether the model is running on a computer or in the head, it takes the information we know and uses it to predict responses in various situations.

Some models are healthy and objective, such as those created from sports statistics; others are harmful and subjective, such as those used to calculate prisoners' chances of recidivism in some US prisons. This second type of model is what she classifies as “*Weapons of Math Destruction*” (WMD). O'Neil lists three questions to identify whether a mathematical model is a WMD: 1) is the model opaque or invisible?; 2) is the model unfair or causes harm?; and 3) does the model have the ability to grow exponentially? In short, it defines that the three elements of a WMD are opacity, scale and damage. We will return to these three elements in this review. First, let's delve into the concepts of algorithm and Big Data.

Algorithms have gained more and more importance in our lives and, therefore, understanding their definition is of fundamental importance. According to Lupton (2014), algorithms are sequences of computer commands that tell the machine how to proceed with a series of instructions to reach a specific result. In other words, Gillespie (2014) defines algorithms as coded procedures to transform input data into a desired output through specific calculations. An analogy often used to explain what an algorithm is is that of the cake recipe. A sequence of commands that will use inputs (flour, eggs, butter), perform transformations (mix all the ingredients, beat and place in the oven) and deliver an output (cake). However, although a simple sequence of commands is considered an algorithm, here we are dealing with complex algorithms, which perform a series of calculations, analyses, make choices and learn (machine learning algorithms). These algorithms mostly use Big Data as input.

The term Big Data, perhaps even more than an algorithm, has become very popular, especially among entrepreneurs and in the media. However, the concept of Big Data is not as simple as the term leads us to deduce. According to Kitchin (2014), the commonly used definition involves the 3Vs: volume, speed and variety. A huge amount of data (volume) is generated and stored daily (speed), with great diversity (variety) in data types. In addition to these three characteristics, Kitchin (2014) also lists others that have been attributed to Big Data, such as being exhaustive, flexible, scalable, having high resolution and relationships with other data. As already stated at the beginning of this review, we are constantly connected and generating data in simple activities such as liking a photo on Instagram or sending a message to a friend on WhatsApp. However, we are also generating data in our day to day offline. For example, when informing our Individual Taxpayer Identification Number (*Cadastro de Pessoas*

*Físicas* - CPF) to the supermarket cashier for the invoice, when registering for an English course or when responding to a Brazilian population census official. All this data is stored and can be used in the future as inputs to algorithms produced by private companies or the State. Thus, data and algorithms are intimately connected. The way in which the mathematical models and algorithms are produced, as well as the data used in the analyzes will be treated, can define whether they will be classified as WMD or not.

Now that we've defined mathematical models, algorithms, and Big Data, we can go back to the three elements that define an algorithm like WMD: **opacity**, **scale**, and **damage**. Cathy O'Neil has a great ability to create analogies to exemplify these three elements and demonstrates clearly and objectively the many ways that algorithms affect our lives. We will use some of the examples presented by her and add similar cases from the Brazilian reality.

A Brazilian magazine publisher created a ranking of the country's universities and colleges. Institutions were rated between one and five stars. These stars reflect the arithmetic mean of the scores given by the evaluators, who are professors and coordinators of undergraduate courses. They receive a questionnaire to evaluate courses and universities, and the same evaluator can evaluate up to 35 courses. Many students came into contact with this ranking and many must have chosen where they would study based on this assessment. Although the criteria for inclusion in a course and the choice of evaluators are clear, are the questions that make up the form fair and reflect reality? Furthermore, once the course coordinators know the evaluation criteria, could this influence their administrative choices for the course and for the university in order to raise the grade? Something similar happened in the United States in the 1980s, and O'Neil demonstrates how the criteria used to develop this model created an "arms race" among American universities.

In 1983, the U.S. News & World Report decided to evaluate 1800 universities in the United States and rank them for excellence. Similar to the Brazilian ranking, they sent a form to university presidents, who evaluated other institutions based on criteria defined by the magazine. O'Neil argues that U.S. News did not have a direct way of quantifying how the higher education process affected one student, let alone millions of them. Rather than choosing objective variables correlated with student success, they decided to select variables such as SAT (Scholastic Aptitude Test) scores, student-faculty ratio, and acceptance rates of students entering and leaving universities in a particular year. From the knowledge of the variables used by the ranking, many of which are easily manipulated, some universities began to manipulate the algorithm in order to increase their grades.

The algorithm created and powered by the U.S. News to rank North American universities meets all the criteria to be considered a WMD. It is opaque, that is, few people have complete knowledge of how the algorithm is built, which variables are considered, and which controls are performed to avoid bias. Algorithms with these characteristics are considered black-box because they have a mysterious system. We know what the inputs and outputs are, but we do not know how one is transformed into the other (PASQUALE, 2015). The U.S. algorithm News is also scalable, that is, it does not affect a small group of people and determines, based on imprecise criteria, which are the “best” universities in the US, impacting the decision of millions of students. Finally, it is also harmful, as it shapes the behavior of universities to obtain better evaluations, preventing, for example, the access of low-income students because they are scholarship holders or have lower SAT scores. In addition, it generates a vicious cycle in which universities that obtained low grades continue to be poorly evaluated and cannot avoid this situation, having less demand from students, receiving less public or private resources and increasing inequality between students and universities.

In another example, the issue of inequality and injustice of WMD algorithms becomes even more evident. In chapter five of her book, Cathy O'Neil explains how models that compute risk of recidivism are used as guides to court sentences and for police patrol. An important point that the author makes throughout the book in relation to models and algorithms is that they are based on the past, that is, the algorithms use as inputs data already collected about a reality - such as the incidence of crimes in a certain region of the city - and look for patterns, assuming they will repeat.

Many drivers of transportation apps, such as Uber and 99, report cases of violence, theft and feel unsafe to travel in areas considered dangerous. Therefore, Uber defined, based on public safety data from Brazilian cities, how its algorithm should respond to calls from these places. In São Paulo, for example, when a user from the Heliópolis neighborhood tried to request a vehicle, he received the following message: “Unfortunately, Uber is not available in your area at the moment” (SOUZA, 2017). In this way, these companies limit the possibilities of accessing the service by the simple location or Postal Address Code (Código de Endereçamento Postal - CEP) of the person making the request. We are aware of the great inequality present in Brazilian society, and cities reflect this, taking poor people to peripheral regions of municipalities with less infrastructure, quality of life and greater susceptibility to crime. In the case of the transport apps mentioned above, people living in these regions can be even more excluded, which increases social inequality.

Now, imagine using the same principle to define the policing of a city. Cathy O'Neil explains that the chief of police in an inner city in Philadelphia, as a result of his squad's downsizing due to budget cuts, decided to implement a crime prediction program from PredPol, a California-based startup, to help police department decision-making. This program is capable of predicting the possibilities of the occurrence of different types of crimes in various regions of the city from historical data. That is, the region's police department includes all data on crimes already recorded in the system, runs the algorithms and predicts in which regions new crimes may occur. From this output, the chief of police can decide where to place his agents. While these crime prediction programs do increase policing efficiency, they can incur a number of biases, mainly due to the choice of data that will be included. O'Neil argues that many of the crimes recorded are not serious, such as aggressive homeless people committing petty theft or selling and using small amounts of drugs - which would normally not be recorded if police officers were not on site. It classifies these lesser crimes as nuisance crimes. Generally, poorer regions of cities have a high rate of occurrence of this type of crime. By including them in crime prediction models, more police officers are assigned to these locations. They will register more cases that will be re-entered into the system, and the police will continue to be sent to these neighborhoods, creating a vicious cycle. While the algorithm is not necessarily racist, the choice of geographic location as a variable is. We could also say that the algorithms of transport apps reinforce the racism of our society by excluding poor regions with higher crime rates from Brazilian cities, populated mostly by black people.

From what was exposed in the previous paragraphs, it becomes more visible how algorithms based on geographic location cause harm to people. Furthermore, this type of algorithm is opaque, as it belongs to a private company, and users cannot modify it – they only include the input and receive the output. Finally, these algorithms are scalable, since large cities in the states of California, South Carolina, Washington, Tennessee, Florida, Pennsylvania and New York are already using predictive policing systems (FERGUSON, 2017).

In the last chapter of the book, Cathy O'Neil explains how Facebook's algorithm has been used in election campaigns, as well as audience segmentation. In 2012, a group of Facebook data scientists conducted an experiment with 689,000 people and demonstrated that the same emotions felt by one person can be transferred (contagious) to others, without them being aware of it (KRAMER; GUILLORY; HANCOCK, 2014). In other words, the Facebook algorithm can be used to modify the emotions of millions of people and, whoever discovers how to manipulate it, can take advantage of advertising campaigns, petitions or elections. The use of social networks for electoral campaigns gained strength from Barack Obama's

presidential campaign in 2012. Content was created in a way that was targeted to social segments, arousing greater interest in a group and generating greater emotional engagement, that is, micro-segmented content. However, content micro-segmentation had a lot of impact in the US presidential election in 2016, when Donald Trump was elected. Here in Brazil, Jair Bolsonaro's campaign in 2018 used social media in the same way: segmented content, provoking different emotions in different social groups and achieving great engagement. In the case of the last two examples, the campaigns were strongly associated with false content (fake news).

Cathy O'Neil does not classify the algorithms of large technology companies (big techs) as WMD, but indicates that they have great potential to be manipulated and misused. We know that algorithms are opaque because they represent a vital business secret for these companies. They are also scalable because social networks have billions of users worldwide. Facebook, adding its three social networks – Facebook, WhatsApp and Instagram – has about 3 billion users. In Brazil, more than 127 million people are Facebook users; about 120 million from WhatsApp, and 69 million from Instagram. They can also cause harm, as they affect the emotions of their users and, as already demonstrated, can lead to cases of depression and suicide, especially among young people aged 13 to 18 who spend more time on social media. (TWENGE *et al.*, 2018).

The manipulation of emotions through social networks, the importance and power of big tech are gaining more and more prominence in discussions about the risks to democracy (MOROZOV, 2018). Runciman (2018) lists the information technology revolution, dependence on forms of communication and information sharing – increasingly controlled by private companies – as one of the three factors that distinguish the current crisis of democracies from those faced in the past. Runciman (2018) also forcefully exposes the control of big tech over society by comparing Facebook to a Leviathan (in the Hobbesian sense), indicating a possible confrontation between Facebook and the States. Although Facebook does not have the prerogative to use force, it does have the ability to change our habits, feelings and persuade us to act the way it wants. For example, shaping the type of content produced, how people expose themselves on social networks and the data they are generating from modifying their algorithms. Although O'Neil does not delve much into these discussions, raising the discussion about the use of social networks and micro-segmentation was her great contribution to the debate in 2016.

Reading O'Neil's book opens our eyes to the importance of inspecting the most distinct algorithms present in our lives, thus preventing them from becoming WMD. As we have seen

in this review, mathematical models and algorithms can become WMD by being opaque, scalable and harmful to society. She does not advocate the banning of algorithms, if only because that would be impossible given the stage of technological advances that we are experiencing. But yes, for greater regulation and transparency. Models cannot sacrifice reality and justice for the simple fact of being and making our lives more efficient. They can make our lives better, helping with tasks that human beings could not perform, but we must select which variables will be chosen, which data will enter and also analyze if the results match reality. Political decisions that impact society must continue to be taken by us, human beings, in a democratic and transparent way, strengthening democracy and fighting for fairer and less unequal societies.

## REFERENCES

- FERGUSON, A. G. Policing predictive policing. **Washington University Law Review**, v. 94, n. 5, p. 1109-1189, 2017. Available: [https://openscholarship.wustl.edu/law\\_lawreview/vol94/iss5/5/](https://openscholarship.wustl.edu/law_lawreview/vol94/iss5/5/). Access: 14 Apr. 2021.
- GILLESPIE, T. The relevance of algorithms. *In*: GILLESPIE, T.; BOCZKOWSKI, P. J.; FOOT, K. A. **Media Technologies: Essays on Communication, Materiality, and Society**. Cambridge: MIT Press, 2014.
- KITCHIN, R. Big Data, new epistemologies and paradigm shifts. **Big Data & Society**, v. 1, n. 1, p. 1-12, jul. 2014. Available: <https://journals.sagepub.com/doi/pdf/10.1177/2053951714528481>. Access: 25 Oct. 2020.
- KRAMER, A. D. I.; GUILLORY, J. E.; HANCOCK, J. T. Experimental evidence of massive-scale emotional contagion through social networks. **Proceedings of the National Academy of Sciences**, v. 111, n. 24, p. 8788-8790, jun. 2014. Available: <https://www.pnas.org/doi/10.1073/pnas.1320040111>. Access: 8 June 2021.
- LUPTON, D. **Digital Sociology**. New York: Taylor & Francis, 2014.
- MOROZOV, E. **Big Tech: A ascensão dos dados e a morte da política**. São Paulo: Ubu Editora, 2018.
- O'NEIL, C. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy**. New York: Crown, 2016.
- O'NEIL, C. **Algoritmos de Destruição em Massa: Como o Big Data aumenta a destruição em massa e ameaça a democracia**. 1. ed. São Paulo: Editora Rua do Sabão, 2020.
- PASQUALE, F. **The Black Box Society: The Secret Algorithms That Control Money and Information**. Cambridge: Harvard University Press, 2015.

RUNCIMAN, D. **Como a democracia chega ao fim**. São Paulo: Todavia, 2018.

SOUZA, F. Uber veta bairros de SP e moradores da periferia criam a Ubra. **BBC Brasil**, mar. 2017. Available: <https://www.bbc.com/portuguese/brasil-39225699>. Access: 07 Dec. 2020.

TWENGE, J. M. *et al.* Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. **Clinical Psychological Science**, v. 6, n. 1, p. 3-17, jan. 2018. Available: <https://journals.sagepub.com/doi/abs/10.1177/2167702617723376?journalCode=cpxa>. Access: 07 Dec. 2020.

### How to reference this article

QUEIROZ, L. L. From university ranking to elections: How algorithms are changing our lives. **Estudos de Sociologia**, Araraquara, v. 27, n. 00, e022016, Jan./Dec. 2022. e-ISSN: 1982-4718. DOI: <https://doi.org/10.52780/res.v27i00.15093>

**Submitted:** 02/07/2021

**Required revisions:** 30/07/2021

**Approved:** 13/01/2022

**Published:** 30/06/2022