
RAGE AGAINST THE MACHINE LEARNING: a critical approach to the algorithmic mediation of information

Arthur Coelho Bezerra (1), Marco Antônio de Almeida (2)

(1) IBICT, Brasil, absurdosturos@gmail.com. (2) Universidade de São Paulo (PPGCI-ECA/USP), Brasil, marcoaa@ffclrp.usp.br

Abstract

Before being an exaltation to Luddites (the English workers from the 19th century who actually destroyed textile machinery as a form of protest) or to some sort of technophobic movement, the provocative pun contained in the title of this article carries a methodological proposal, in the field of critical theory of information, to build a diagnosis about the algorithmic filtering of information, which reveals itself to be a structural characteristic of the new regime of information that brings challenges to human emancipation. Our analysis starts from the concept of mediation to problematize the belief, widespread in much of contemporary society, that the use of machine learning and deep learning techniques for algorithmic filtering of big data will provide answers and solutions to all our questions and problems. We will argue that the algorithmic mediation of information on the internet, which is responsible for deciding which information we will have access to and which will remain invisible, is operated according to the economic interests of the companies that control the platforms we visit on the internet, acting as obstacle to the prospects of informational diversity and autonomy that are fundamental in free and democratic societies.

Keywords: critical theory of information; regime of information; algorithmic filtering; machine learning; mediation.

Who controls the past controls the future. Who controls the present controls the past
George Orwell, 1984

Who controls the past now controls the future. Who controls the present now?
Rage Against the Machine, Testify

1 Introduction

A report published on the digital page of the British newspaper The Guardian, on May 2016, brought the complaint that Facebook, a Social Network Site (SNS) described as “the biggest news distributor on the planet”, had a small and discreet editorial team responsible for interfering in the list of most commented subjects (the trending topics) of the platform. Facebook stated, at the time, that “the topics you see are based on a number of factors including engagement, timeliness, pages you’ve liked and your location”. In the same week, the company had been accused of inferring an editorial bias against conservative news agencies, which prompted requests for an investigation by the US Congress – one of many that Facebook would go through.

The headline of The Guardian's complaint says: “Facebook news selection is in hands of editors not algorithms, documents show” (The Guardian 2016a). Three and a half months later, a new report is published in the newspaper, with the following headline: “In firing human editors, Facebook has lost the fight against fake news” (The Guardian 2016b). The new article explains that, two days after Facebook announced the decision to expand the automation of its trending topics, firing 26 professionals responsible for editing the list of most commented subjects, the new list forged by the company's algorithms brought out a story entitled "Breaking: Fox News Exposes Traitor Megyn Kelly, Kicks Her Out for Backing Hillary". It was fake news – one of thousands of fake news that circulated in the context of the 2016 electoral dispute between Hillary Clinton and Donald Trump.

There is a dialectical element that arises from the intersection of these two news, that is, the complaint of the use of humans to filter information (strangely coming from a media vehicle that built its history doing the same thing), followed by denouncing the use of algorithms to filter information. What we see is the replacement of what Anthony Giddens calls “expert systems”, defined by the British sociologist as “systems of technical accomplishment or professional expertise that organize large areas of the material and social environments in which we live today”

(1990 p. 27). Giddens exemplifies the idea of expert systems with the scenario we enter when driving a car, which involves building the cars themselves and also engineering the roads, intersections, traffic lights and various other systems, whose technical functioning most people are unaware of; nevertheless, those systems enjoy the confidence of these people because of their faith in the expert knowledge of the professionals who create such systems. The same is true of the trust we place in doctors, lawyers, teachers and other specialists.

For Brazilian political scientist Luis Felipe Miguel (1999 p. 199), journalism could also be considered an expert system, as it would enjoy public confidence in three important aspects: 1) confidence in the veracity of the information reported; 2) confidence in the correctness in the selection and ranking of the important elements in the report; 3) confidence as to the correctness in the selection and hierarchy of the news in view of the stock of available facts. But when the company considered “the biggest news distributor on the planet” is accused of using journalists to act on the same aspects listed by Miguel, and when such accusation contributes to the replacement of communication professionals with algorithms, the impression we have is that journalism would have lost its status as an expert system, and that the list of the trending topics should be the responsibility of mathematical formulas that, in theory, would be free from the editors' ideological bias.

The outcome of the story shows, however, that algorithms shouldn't be granted the level of an expert system without being seen with any degree of mistrust regarding the fairness of the selection, hierarchization, organization and mediation of the information that is proposed to accomplish.

The current forms of data surveillance and algorithmic filtering of information on the internet, combined with the circulation of a large volume of false or misleading news on SNSs such as Facebook, WhatsApp and Twitter, underscore the importance of individuals' capacities for critical evaluation and ethical use of information, around which the concept of critical information literacy is built (Bezerra 2019). Considering that such consciousness cannot be achieved without a satisfactory knowledge about the structures of the current dominant regime of information, we believe it is necessary to undertake a criticism of the architecture of the algorithms that filter and decide which information is visible or invisible to users of the platforms that apply such filters, unveiling the interests that cause the mathematical formulas of these algorithms to be written the

way they are, and not so as to contribute to a broader perspective of informational diversity and autonomy.

Well known in the Brazilian field of Information Science, the concept of “régime of information” was first presented at a conference in Canada by Bernd Frohmann, in 1995, to characterize systems or networks that have specific channels, structures, producers and consumers, whose degree of stability, however, is relative. With declared inspiration in the actor-network theory of French sociologists Bruno Latour and Michel Callon, Frohmann says that “describing a régime of information means charting the agonistic processes that result in tentative and uneasy stabilizations of conflicts between social groups, interests, discourses, and even scientific and technological artifacts” (1995 p. 5).

Based on the understanding of Frohmann, who admits the stabilization of information regimes as something conflicting, it is possible to perceive that such “conflicts between social groups, interests, discourses and even scientific and technological artifacts”, to which the philosopher credits the instability of regimes, are, in fact, inscribed in the struggles of economic groups for the power to interfere in information policies in order to preserve their dominance positions in the markets in which they operate, with scientific and technological speeches and artifacts figuring as immaterial and material representations interests that guide such struggles. The mass media, including the mentioned SNSs, has become the space where business and political strategies are played out. According to Manuel Castells (2009), power now lies in the hands of those who control communication; therefore, as Frohmann himself argues, studies on information policies and regimes must take into account the fact that the dominance of information is achieved and maintained by specific groups, and that specific forms of domination are involved in the exercise of power over information (Frohmann 1995).

In this sense, our effort to critically approach the phenomenon of algorithmic filtering of information stands in the perspective of a critical theory of information, according to the definition proposed by Christian Fuchs:

Critical information theory is an endeavour that focuses ontologically on the analysis of information in the context of domination, asymmetrical power relations, exploitation, oppression, and control by employing epistemologically all theoretical and/or empirical means necessary for doing so in order to contribute at the praxeological level to the establishment of a participatory, cooperative society (Fuchs 2009 p. 243)

The methodological proposal for a critical theory of information, as argued in Bezerra (2019 p. 28), “must include the performance of interdisciplinary diagnoses that focus on the informational environment and the perspectives of production, circulation, mediation, organization, recovery and accessibility of information”, identifying not only the potentialities of a specific regime of information for human emancipation, but also, and above all, the obstacles to such emancipation. In this article, such method will be carried out by examining the role of the algorithm mediation in the new regime of information (section 3) and exploring how the structure of the algorithms found in most popular SNSs creates barriers to the informational autonomy of individuals in contemporary society (section 4).

Nevertheless, before investing on the criticism of the architecture of algorithms, more specifically in the discussion about the algorithmic mediation of information, it is worth recovering a brief history about the formulation of the concept of mediation itself, and its historical changes resulting from the challenges of incorporating technological innovation.

2 Brief considerations on the concept of Mediation

An initial point of this discussion, almost an a priori, starts from the observation that the term mediation is intrinsically polysemic. The concept both derives its meaning from the theoretical conceptions to which it is linked as a cultural, communicational or informational practice, as well as from the specific professional context in which it is used. As noted by Davallon (2007), a consensual definition of mediation seems impractical, since it is a plastic concept that extends its borders to account for very different realities.

In the field of Social Sciences, the concept of mediation is often associated with so-called “theories of social action”. These theories postulate that social actions are understood in broader systems of processes of intersubjective understanding, through which actions are coordinated, involving the role of agents ('human mediation') in these processes. Language and common action are the privileged factors of mediation. They are processes of interlocution and interaction between the members of a group or community, which allow the establishment and support of bonds of sociability, thus constituting, in a Habermasian perspective, the world of life.

A second important point about the concept of mediation is that one can observe a certain crystallization, in a significant part of the bibliography, of the conception that mediation actions would not be limited to the establishment of a simple relationship between two terms of the same level, but that they would be in themselves producers of "something more", or of a more satisfactory state in relation to the initial conditions. Thus, extrapolating the more general sociological conception, we will see that the incorporation of mediation as an activity in the "world of systems" will imply a polysemy of conceptions, related to the diversity of the institutional contexts in question. We can understand, therefore, mediation as "a set of social practices, which are developed in different institutional sectors and which aim to build a space determined by the relationships that manifest in it" (Caune 2014 p. 73). The concept of mediation, in its plasticity and flexibility, would cover quite different activities at the institutional level, ranging from the current conceptions of "customer service" to actions of cultural agents in institutions (museum, library, archive, cultural center), the elaboration of training policies or access to information and communication technologies and, of course, technological mediation provided by networked informational tools (such as SNSs and other websites), thus practically precluding a single consensual definition, as the meanings of the practices it covers derive from very different realities (Almeida 2014).

In this perspective, mediation processes would "add value" to cultural, informational or communicational processes, providing gains in terms of knowledge to the subjects involved. Therefore, mediation activities would be valued for the implicit potential of "generating cultural value".

On the other hand, in the countercurrent of this conception, certain authors would see in the mediation processes an imposition of values, a "training" of the receptive competences of the subjects. Thus, they would start to value the technological changes that enable "disintermediation", the possibility of the subjects to exercise their autonomy in the process of building their own knowledge, leveraged by the resources of the internet (Lévy 2000; Fourie 2001).

The role of information and communication technologies (ICTs) is central to this process. It was ICTs that established a global network in which societies with different ways of life and histories intertwine. This new environment had a profound impact on individuals' daily lifestyles, bringing about clashes between local cultures and globalized cultural flows. Stuart Hall (2005)

Bezerra, Arthur Coelho, Almeida, Marco Antônio de. Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, vol.14, no.2, Abr.-Jun. 2020 pp. 06-23

calls attention to a series of displacements in the daily lives of cultures, resulting from global changes in the socioeconomic reorganization of “postmodern” or “postindustrial” societies. He is particularly attentive to the role of traditional media and ICTs in the dissemination of information and symbolic content.

The context that allows the construction of a concept such as that of disintermediation is the development of increasingly sophisticated information products and services and, at the same time, of relative ease of use by individuals. This is the case with large search engines, particularly Google, created in 1998, which has become the hegemonic reference. Using a search algorithm, PageRank, Google took the academic citation procedures as a model to assess the relevance of internet pages. In the contemporary scenario of information flows, Google and its counterparts would contribute decisively to the process of “disintermediation”, playing the alleged role of precise and safe guides for users regarding the information they need. But do things really happen that way?

The cultural system of the contemporary world is increasingly characterized by its increasing complexity. This configuration demands a sophisticated apparatus of information, involving increasing physical and human resources. George Yúdice (2006) considers that culture in the contemporary world becomes an increasingly strategic resource – but problematizes this discussion by not reducing it to the role of a simple commodity. In the current context, cultural resources, as well as natural resources, do not support pure and simple exploitation. The cultural system involves management, conservation, access, distribution and investment. Managing cultural resources, aiming to achieve different objectives, has become a challenge for States, companies and social movements.

On the other hand, the idea of a network seems to refer to a universe of freedom and informational abundance available to all individuals and groups. It is a mystification supported by a false idea of “neutrality” of technologies: there is no network without choice, without organization, without hierarchy, since knowledge does not exist outside a social context, nor does it randomly reorganize. Critical media and information literacies and competences – both communicative, as well as cultural, educational and cognitive – are fundamental for individuals to contextualize information and use it, which leads to an old discussion: social inequality is not just an issue regarding the appropriate sharing of resources, but participation in determining life

opportunities both individually and collectively. It is in this perspective that stands a critical discussion about the algorithmic mediation in the new regime of information.

3 The algorithmic mediation of the new regime of information

Unprecedented practices of web scraping, data mining, algorithmic filtering and psychometric analysis emerge in what we consider a new regime of information mediation (Bezerra 2017). First introduced by Amazon to suggest books based on what users search for on its platform, algorithmic data filtering is now the preferred way of organizing information for many Internet pages and services, including popular names like Facebook, Instagram, Google, YouTube, Netflix, Spotify and many others.

The entries that feed the algorithms of these companies are extracted (web scraping) from the users' navigation on their online platforms or operating systems, internet browsers, smartphone applications and other data storage devices, which allow the identification of a large volume of information on personal data, such as location, access time, device used (computer or smartphone) and anything that can be tracked considering searches, clicks, time spent between clicks, relevance of the first search results (from search engines like Google, for example) and a number of other factors. Subsequently, all this monitoring feeds mathematical formulas that, using artificial intelligence principles such as machine learning and deep learning techniques, build a profile of each user's preferences, with precise information about age, economic class, gender identity, sexual orientation, consumption habits, cultural and sports preferences, political spectrum, participation in ethnic groups and social movements and so on. The intersection of these techniques with statistics allows the discovery of patterns in large data sets (data mining), which are used to target specific content to specific users (microtargeting).

Each platform has access to a certain set of information, and big tech companies use it to attract users, hold their attention and keep them on their page, by applying an algorithmic filtering process with a view to suggesting content from movies and series on Netflix, music on Spotify, videos on YouTube, books and other products on Amazon and posts on Facebook. In the most controversial case known, the company Cambridge Analytica applied psychometric analysis techniques of personal data scraped from tens of millions of Facebook profiles to micro-target

personalized political advertising, which would have contributed (for some, in a decisive way) to Donald Trump's victory in the 2016 presidential election in the United States (Bezerra 2019 p. 49).

Authors who focus on the field of “attention economy”, many of which on business schools, understand attention as a limited resource – we have a finite capacity to pay attention to something – that has become increasingly scarce in the “information society”. The hyperinformation that presents itself as a phenomenon inherent in the production of an increasing volume of informative content on digital networks, whether in the form of news, videos, music, texts, posts, post comments, tweets, retweets, memes, WhatsApp audios and a multitude of new forms of communication leads internet companies to hire administrators, economists, engineers, psychologists, social scientists and other professionals, in order to compete in the market constituted by our attention – a market that has great value in the current information economics.

As previously seen, platforms that use algorithms to target content and advertising depend on us to continue visiting their platforms and interacting with the content they provide, so that we can generate more and more information to integrate these companies' big data, becoming what Shoshanna Zuboff (2019) calls "surveillance capital". Therefore, the more time a person spends on a platform, the more clicks and consequently more data on personal preferences will be handed out to the company that controls the platform.

Zuboff coined the term "surveillance capitalism" to characterize the new form of information capitalism that seeks to predict and modify human behavior as a means of producing revenues and market control. For her, the triumph of Google in building a new form of market that engenders the aforementioned surveillance capitalism, understood by Zuboff as a radically extravagant variant of information capitalism, is due to the successful combination of data, extraction and analysis. This combination forges a new logic of capitalist accumulation, whose revenues, which depend on data that are mined through automated operations, constitute what the author calls “surveillance assets”, which are responsible for attracting significant investments, mainly from companies interested in their targeted advertisement expertise.

In 1971, economist, psychologist and political scientist Herbert Alexander Simon declares:

...in an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information

creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it (Simon 1971 p. 40-41)

Seven years later, Simon would receive the Nobel Prize in economics. The same concern with filtering information by individuals, so that they can select which information is relevant and use it “efficiently”, lies in the foundations of the concept of information literacy, as first proposed (also in the seventies) by Paul Zurkowski (1974) – who was at the time president of the Software and Information Industry Association – and later consolidated in documents from the Association of College and Research Libraries (ACRL), division of the American Library Association (ALA), and the International Federation of Library Associations and Institutions (IFLA), which are reference for researchers on the subject.

In the new information regime, however, algorithms have become increasingly responsible for carrying out this first filtering stage, playing the role of mediators of information that we have access to on the platforms we visit on the internet. And this happens with almost no public resistance; on the contrary, there is a widespread belief that, because algorithms deal with the mathematical organization of a large volume of data, they have the power to solve all the communication problems in the world.

The phenomenon has been called "dataism" by authors such as South Korean philosopher Byung-Chul Han (2017), who compares the current euphoria surrounding big data with that around statistics in the 18th century, considering it a kind of “second Enlightenment” whose imperative is, ultimately, to transform any and all human action and experience into quantifiable data and information. What dataism enthusiasts seem to forget is that algorithms, just like new forces of production and relations of production, as Marx reminds us in his *Grundrisse*, “do not develop out of nothing, nor drop from the sky, nor from the womb of the self-positing Idea; but from within and in antithesis to the existing development of production and the inherited, traditional relations of property” (1973 p. 278).

4 The structure of the algorithms

The word "algorithm" is said to have been coined by a Persian mathematician in the 9th century, although its concept was already present in calculations made by the Greeks centuries

Bezerra, Arthur Coelho, Almeida, Marco Antônio de. Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, vol.14, no.2, Abr.-Jun. 2020 pp. 06-23

before Christ, by Egyptians a thousand years before that, and by Babylonians two thousand years before the Greeks. The widespread use of the concept, even before it is represented in word, is the result of its elementary meaning: an algorithm is simply an instruction or set of instructions, that is, a method to solve a problem or reach a result effectively. It therefore attends to any decision making that can be expressed in mathematical language.

The most commonly used example to explain the operation of an algorithm is the cake recipe, which has information regarding the quantity of ingredients and the method of preparation, being necessary to follow the instructions and take into account the time of each operation so the cake doesn't burn. In general, this is what machine learning techniques and algorithms do most of the time when we access the Internet: from information about our preferences, the algorithms create a recipe to predict what kind of information will be of interest to each person.

As the algorithms designed to offer personalized content to Internet users understand an unread news, unseen video or unlistened music as a failure, and a "like" in a photo or comment in a post as a success, the tendency is to show more videos, photos and texts of people or on subjects for which we show interest in our previous interactions. However, that understanding overlooks two important aspects regarding the structure of the algorithms.

The first aspect is that the fact that algorithms are mathematical formulas often gives the impression that they are neutral; however, it is always good to remember that an "empty" algorithm is full of human decisions, from the objective to be achieved to the method (or set of operations) defined as the most effective to achieve that objective. In this sense, if a commercial company that controls one or more platforms on the internet aims to profit, as is the case with any commercial company; whether that profit comes from the clicks of the users in the advertisements made available on the platforms; whether the chance of someone clicking on an ad increases when that ad is targeted specifically at that certain someone; whether advertising crossed the path of this certain person and changed the direction of their navigation thanks to the platform's ability to personalize advertising and attract attention and interest; and if the ability to personalize this advertisement increases as users spend more time on that platform and generate more information to fatten the big data of the company that owns the platform, it means that the algorithms and techniques for filtering informational and advertising content will be designed, after all, by different human specialists at the service of the companies that profit from it, in order to use the

previous behavior of people on the network to predict what type of content is most immediately attractive and has the potential to generate greater engagement and length of stay of individuals in the network. It's pure math – or, as Cathy O'Neil (2016) would say, “weapons of math destruction”.

The math-powered applications that feed data economy were based on choices made by fallible human beings. Some of these choices were no doubt made with the best intentions. Nevertheless, many of these models encoded human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives. Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists (O'Neil 2016 p. 10).

Therein lies another problematic aspect, often disregarded by common sense and even in the academic environment, that algorithms show what, in fact, people want to see. First, it should be clear that Apple, Microsoft, Google, Amazon, Facebook, Netflix and Spotify are not non-profit organizations, but capitalist multinationals that aim to profit; therefore, it is not surprising that the algorithms of all these companies are private and confidential. Just like the Coca-Cola formula, which until 1929 contained cocaine, the success formulas of the internet giants also have their secret ingredients, and changes in the structure of algorithms after a period of monitoring tests represent more a rule than an exception.

PageRank, Google's first and best-known algorithm, named after one of the creators of the search engine, Larry Page, has been replaced by a number of other machine learning and deep learning mechanisms, as well as EdgeRank, the Facebook algorithm. In 2019, changes in Amazon's search engines raised suspicions that the company was artificially boosting its own products, meaning that books that appear as suggestions to us don't always align with the company's argument that people who bought a certain book also bought the other suggestions provided by the platform.

Secondly, as Eli Pariser (2011) points out, the information we want to consume now is not always the information we really want or that we judge best for us. Creator of Avaaz, a non-commercial platform that promotes digital activism around the world, Pariser is the author of the book that popularized the term “filter bubble”, or simply “my bubble”, as people say in Brazil. Using metaphors related to our “informational diet”, Pariser argues that, just as we cannot survive on junk food alone, we must balance our consumption of junk information with content that feeds our brain, our creativity, and allows us to have a vision of the world more permeated by the

Bezerra, Arthur Coelho, Almeida, Marco Antônio de. Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, vol.14, no.2, Abr.-Jun. 2020 pp. 06-23

diversity of opinions and points of view, so as not to become encumbered in our egoic universes (a fact to which the architecture of many SNSs contributes) and not to become intolerant to the differences that we invariably encounter in the globalized world.

When we talk about the emergence of algorithms and the “culture” associated with them, we are considering the set of metaphors built around the phenomenon of data explosion – big data – and its consequences for the debate not only about technology, but also about culture, economics and contemporary politics. The fact is that this buzzword took root in everyday discussions, constituting a social imaginary in which there are “data reservoirs”, composed of the different sets of user information stocks, ready to be explored, in parallel with the extraction of natural resources. The conception of data as “resources” has become commonplace, although the production of data is very different from how Nature produces its resources. Technology companies focus their advertising efforts on “sharing”, that is, on the voluntary transfer of this information by people, the first step towards building a community utopia of better services, building knowledge and sharing “experiences”. Not sharing has become the real capital sin of the present day.

However, as noted by Evgeny Morozov (2018), such “data extraction” has its economic and political consequences. In the economic field, we see immense wealth being accumulated by a handful of investors and business giants. Phenomena such as Uber, for example, can only be understood when the source of its resources is known: sovereign wealth funds and investment banks like Goldman Sachs. The application's ability to incorporate huge contingents of supposedly autonomous drivers is directly related, in turn, to the precariousness of the forms of regulation of work and services (“flexibilization” in what Orwell would call the entrepreneurship “newspeak”). In the public sector, the reform of large systems, such as healthcare, education, public administration or social security, involves the intermediation of digital service providers, with the announced advantage of saving resources but with the counterpart, almost never discussed, of loss of political and governance control over decisions that are now “technically” exercised by the algorithms. As Morozov notes, “we should take stock of the structural factors that put governments and other public institutions in the hands of these large technology companies” (Morozov 2018 p.168).

Another example given by Morozov about the penetration of this culture of algorithms is in the spread of fake news: fake news have always existed, but now they circulate in the digital

Bezerra, Arthur Coelho, Almeida, Marco Antônio de. Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, vol.14, no.2, Abr.-Jun. 2020 pp. 06-23

medium with much greater ease and speed because they thrive on click-based models. As this logic has already been installed in the collective unconscious, the way in which this spread of fake news is dealt with consists of reinforcing trust in large technology companies, assigning them the role of identifying and distinguishing what is false from what is true. Something that, paradoxically, they could only accomplish through algorithms – a task in which they did not demonstrate particular competence, just remembering the bizarre cases of distinction between artistic photos and pornographic photos established by Facebook. The hope in this type of politically imposed “neutral” algorithmic control, unfortunately, “is part of a larger effort to recruit predictive technologies – taking advantage of the huge volume of data already accumulated – in the name of control and surveillance” (Morozov 2018 p. 170).

The fact that this intensification of both data extraction regime and the impositions of surveillance and control, as well as the loss of related political autonomy, has not yet generated widespread discontent or revolt may be explained by the fascination around the myths arising from the Silicon Valley: the creative powers of individual entrepreneurship and the kind neutrality of technologies.

5 Conclusion

The practices of data mining and algorithmic mediation of information through machine learning techniques, currently operated as a new form of capitalist accumulation, are fundamental components that forge the new regime of information. This understanding follows Horkheimer's perception that the use of all physical and intellectual means dominated by nature is, in dominant social relations, subordinated to particular and conflicting interests.

From Horkheimer to Zuboff and O'Neil, the reconfigurations of capital, in the midst of technical and scientific development that allowed the production of a new commodity – big data – and the transformation of its mining and commercialization into a billionaire business model reveal that, although initiatives and services from internet multinationals that effectively contribute to life in society can be pointed out, today, as in Horkheimer's and even in Marx's time, the production is neither aimed at collectivity life nor to satisfy the demands of individuals, but it is

oriented to the demand for power of individuals and is also responsible for the shortage in the life of the collectivity (Horkheimer 1980 p.134).

By transferring our choices to the algorithms, we forge the universe as an immense (but measurable) flow of data, and we trust that the processing of this cosmic big data by the algorithms will provide all the answers we need; that is why we have no great objections to becoming “star stuff,” as Carl Sagan would say, integrating ourselves into this constellation of data through the peaceful and unreserved delivery of our biometric, geographic, financial, political and social information.

Far from being an objective mathematical expression that is capable of translating the world into quantifiable data and making infallible predictions, the algorithms and techniques of machine learning and deep learning are human creations that express specific trends, interests and objectives, being, therefore, susceptible to errors and inaccuracies. Since machine learning algorithms build a mathematical model based on the data collected to make predictions or decisions, their specific design will have strong effects on what kind of information people will have access to. The conclusion, as the epigraph of this article summarizes, is that who controls the past controls the future, and who controls the present controls the past.

The use of mathematical formulas to organize big data can be seen as a successful case of technological innovation and business model. However, if the algorithms ever showed us what they thought we would like to see on the network, it's becoming increasingly difficult to separate predictions from prescriptions, as the filtering process tends to suggest what our next steps will be based not only on our previous steps, but also on their private interests. The insertion of the individual personalization mentality of the internet in the design of the algorithms that filter the information to be displayed directly interferes in the type of information that users will have access to in the network, creating apparently comfortable environments that, in fact, encapsulate users in self-referenced and ego centered horizons, which creates obvious limits and obstacles to informational diversity and autonomy.

A fundamental political discussion is related to the architecture of the technology and how it is handled today by the data extraction sector. This should not be confused with a critique of the technology itself, but with the realization that the uncritical idea that big data and the associated

culture of algorithms are the legitimate messengers of truth (the bigger the data set, the bigger the truth can be extracted from them) contributes to the resumption of positivism and its monolithic conception of knowledge. It is necessary to recover the social dimension of human creation in science and technology and to reinforce the need for dialogue, for public debate, for a set of truly political practices as opposed to the imposition of technocratic views.

Notes

(1) Acknowledgments for financing from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through productivity grants to both authors

References

- Almeida, M. A. “Mediação e mediadores nos fluxos tecnoculturais contemporâneos”. *Informação e Informação*, Londrina, vol. 19, no. 2, maio/ago 2014, pp. 191-214.
- Bezerra, A. C. “Vigilância e cultura algorítmica no novo regime de mediação da informação”. *Perspectivas em Ciência da Informação*, Minas Gerais, v. 22, 2017, pp. 68-81.
- Bezerra, A. C. “Teoria crítica da informação: proposta teórico-metodológica de integração entre os conceitos de regime de informação e competência crítica em informação”. *iKRITIKA: estudos críticos em informação*. By Bezerra, A. C. et al. Rio de Janeiro, Garamond, pp 15-72, 2019.
- Castells, M. *Communication Power*. Oxford: Oxford University Press, 2009.
- Caune, J. *Cultura e Comunicação: convergências teóricas e lugares de mediação*. São Paulo, Ed. UNESP, 2014.
- Davallon, J. “A mediação: a comunicação em processo?” *Revista Prisma.Com*, Porto, no. 4, 2007, pp. 03-36.
- Fourie, I. “¿Debemos tomarnos en serio la desintermediación?” *Anales de Documentación: Revista de Biblioteconomía y documentación*, Murcia, vol. 4, 2001, pp. 267-282.
- Frohmann, B.” Taking information policy beyond information science: applying the actor network theory.” *23rd Annual Conference of the Canadian Association for Information Science*, 1995, Edmonton, Alberta.
- Fuchs, C. “Towards a Critical Theory of Information”. *TripleC*, Reino Unido, vol. 7, no. 2, 2009, pp. 243-292.

Bezerra, Arthur Coelho, Almeida, Marco Antônio de. Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, vol.14, no.2, Abr.-Jun. 2020 pp. 06-23

- Giddens, A. *The Consequences of Modernity*. Cambridge: Polity Press, 1990.
- Hall, S. *A Identidade Cultural na Pós-modernidade*. Rio de Janeiro: DP&A, 2005.
- Han, B. C. *Psychopolitics: newliberalism and new technologies of power*. Brooklin: Verso Books, 2017.
- Horkheimer, M. “Teoria Tradicional e Teoria Crítica”. *Textos escolhidos*. By Benjamin, Walter, et al.. São Paulo, Abril Cultural, 1980, pp. 117-154. Coleção Os Pensadores
- Lévy, P. “A revolução contemporânea em matéria de comunicação”. *Para Navegar no Século XXI: tecnologias do imaginário e cibercultura*. Organizador by Martins F. M. ; SILVA J. M. Porto Alegre, Sulina, 2000. pp. 195-216.
- Marx, K. *Grundrisse*. London: Penguin Books, 1973.
- Miguel, L. F. “O jornalismo como sistema perito”. *Tempo Social*. USP, São Paulo, vol. 11, no. 1, 1999, pp.197-208.
- 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 Publishers, 2016.
- Pariser, E. *The Filter Bubble: what the internet is hiding from you*. London, Penguin Books, 2011.
- Simon, H. A. “Designing organizations for an information-rich world”. *Computers, communications, and the public interest*. Editado by Greenberger. M. Baltimore, MD, The Johns Hopkins Press, 1971, pp. 38-72.
- The Guardian. *Facebook news selection is in hands of editors not algorithms, documents show*. 2016a. Available at: <https://www.theguardian.com/technology/2016/may/12/facebook-trending-news-leaked-documents-editor-guidelines> Access in march 2020
- The Guardian. *In firing human editors, Facebook has lost the fight against fake news*. 2016b. Available at: <https://www.theguardian.com/technology/2016/aug/29/facebook-trending-news-editors-fake-news-stories> Access in march 2020
- Yúdice, G. *A Conveniência da Cultura: usos da cultura na era global*. Belo Horizonte, UFMG, 2006.
- Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, Public Affairs, 2019.
- Zurkowski, P. *Information Services Environment Relationships and Priorities*. Related Paper, nº 5. Washington D.C: National Commission on Libraries and Information Science, 1974.

Copyright: © 2020 Bezerra, Arthur Coelho, Almeida, Marco Antônio de. This is an open-access article distributed under the terms of the Creative Commons CC Attribution-ShareAlike (CC BY-SA), which permits use, distribution, and reproduction in any medium, under the identical terms, and provided the original author and source are credited.

Received: 23/03/2020

Accepted: 07/06/2020