

INTELIGÊNCIA ARTIFICIAL E ALGORITMOS

ADRIANA GONÇALVES
LUÍSA TORRE
PAULO VICTOR MELO
[EDS.]



LABCOM
COMUNICAÇÃO
& ARTES



INTELIGÊNCIA ARTIFICIAL E ALGORITMOS

DESAFIOS E OPORTUNIDADES PARA OS MEDIA

ADRIANA GONÇALVES

LUÍSA TORRE

PAULO VICTOR MELO

[EDS]



LABCOM
COMUNICAÇÃO
& ARTES

Ficha Técnica

Título

Inteligência Artificial e Algoritmos:
Desafios e oportunidades para os media

Editores

Adriana Gonçalves, Luísa Torre & Paulo Victor Melo

Editora LabCom

www.labcom.ubi.pt

Coleção

Livros de Comunicação

Direção

Gisela Gonçalves

Design Gráfico

Cristina Lopes

[Imagens de capa e separadores criadas através
da ferramenta Midjourney AI]

ISBN

978-989-654-973-2 (papel)

978-989-654-975-6 (pdf)

978-989-654-974-9 (ePub)

Depósito Legal

527476/24

Tiragem

Print-on-demand

Universidade da Beira Interior
Rua Marquês D'Ávila e Bolama
6201-001 Covilhã
Portugal
www.ubi.pt

Covilhã, 2024

© 2024, Adriana Gonçalves, Luísa Torre & Paulo Victor Melo.

© 2024, Universidade da Beira Interior.

O conteúdo desta obra está protegido por Lei. Qualquer forma de reprodução, distribuição, comunicação pública ou transformação da totalidade ou de parte desta obra carece de expressa autorização do editor e dos seus autores. Os artigos, bem como a autorização de publicação das imagens, são da exclusiva responsabilidade dos autores.



Índice

Introdução	9
Adriana Gonçalves, Luísa Torre e Paulo Victor Melo	
PARTE I - TECNOLOGIAS E OS MEDIA	13
Tecnologia, sociedade e democracia: a questão da regulação	15
J. Paulo Serra	
Inteligência artificial, algoritmos e media: diálogos de pesquisa	31
Adriana Gonçalves, Luísa Torre e Paulo Victor Melo	
Como (e por que) os jornalistas devem investigar algoritmos que tomam decisões automatizadas nos serviços públicos?	55
Krishma Carreira	
PARTE II - INTELIGÊNCIA ARTIFICIAL, ALGORITMOS E VIESES	67
Fake News as Digital Disruption: Unravelling Algorithmic Logic in the Spread of Disinformation	69
André Lemos	
Inteligencia Artificial y <i>Deepfakes</i> : sesgos de género y agresión contra las mujeres	83
Rosa Franquet	
Por que falar de raça quando falamos de dados pessoais, inteligência artificial e algoritmos?	103
Johanna K Monagreda	
Fighting Algorithmic Racism: reactions, remediations and re-appropriations	135
Tarcizio Silva	
Biografias dos/as autores/as	161

FIGHTING ALGORITHMIC RACISM: REACTIONS, REMEDIATIONS AND RE-APPROPRIATIONS

Tarcizio Silva

/ Mozilla Foundation

There are many reactions and paths being charted by activists, developers, scientists and technologists of many fields and disciplines. Algorithmic racism is not a phenomenon that can be addressed in a simple way. More than handling “racist algorithms”, the matter is the algorithmization of racism through some key points: reproduction and machine reinforcement of social, political and cultural inequality; growth of opacity regarding racial relationships and resulting oppressions; reproduction of necropolitics biopower as the foundation of the contemporary technocrat neoliberalism; and deepening of the colonial and racialized extraction of data and labour in the Global South.

Therefore, modalities of resistance, reactions, and remediation against the algorithmic transformation of structural racism involves remembering of the diverse fronts of Black movements in social battles and in diasporic solidarity. As well as in the refuse to disaggregate the identities and to not adhere the upkeep of the *status quo*, as said by Jurema Werneck (2010), reminding us that our steps come from afar.

Public Audits and Awareness

In the studies about algorithm racism, we come across the exposing of manifestations of algorithmic racism through public audits, journalism, research and

activism. Gathering evidence about the fragility of algorithmic systems is a multidisciplinary task. The evidence are not only computational or a result of code audits, but also from data gathering, ethnographies and journalistic investigations – all can have distinct impacts depending on the power relations, professional authority and social-political contexts.

The project *Gender Shades* brought light about the intersectional disparities that computational vision systems performed by making unacceptable mistakes against black women. Beyond pointing out the problems, the researchers created a benchmark of photographs to test the systems, in a reproducible format. Through careful curatorship, the *Pilot Parliaments Benchmark* is an instrument that allows any developer or company to analyze the precision of their system regarding gender and skin tones (Buolamwini & Gebru 2018).

That way, besides the audit of the three first systems – from IBM, Microsoft and Face++, the project *Gender Shades* suited the community as an instrument of analysis that has a potential for large scale replication. Two years after, Joy Buolamwini and Deborah Raji (2019) conducted a new audit about the analyzed systems in the first phase of the study and included comparisons between other suppliers, from Amazon and Kairos. The researchers found out that the systems approached previously improved the error rates, but the ones which were analyzed for the first time on this round followed the same tendencies of inter-sectional mistakes, with bigger imprecision regarding black women.

But beyond the metrics of the softwares from the targeted companies in the study and their practices, the project can be understood as an “actionable public audit”. As far as its repercussion was based on principles of scientific disclosure built on the connection between academic spaces of power and visibility, it also generated debates in the public sphere, often mentioned in mobilized groups and even in regulatory propositions.

To address the racializing politics in the construction and functioning of algorithmic systems, however, should not be seen as something pertain-

ing solely to computer science and affiliated areas, limited by disciplinary borders. Understanding algorithmic systems do not only involve the path of analyzing code lines, but moving through its networks of delegation, which behaviors are normalized, which data is accepted, what types of mistakes are or are not considered between inputs and outputs of the system, its potential of transparency or opacity and which presences of absences are implemented – finally, the networks of racial-political relations in the changing materializations within technology.

The experiential knowledge about algorithmic systems transformed not only through scientific approaches but also through vernacular approaches reminds us that some dynamics of discrimination present themselves on the surface, even if they require enunciation. An example was the campaign *#BuscaPorIgualdade*, carried by “Desabafo Social”. In short videos, the organization exhibits queries for key words of popular categories, such as “people”, “family” or “skin”, in stock photo websites like *Shutterstock* or *Deposit Photos*, resulting in pages filled with almost exclusively white people. The contraposition with the searches supplemented with the qualifier “black” in terms like “family” is the hook to remind the stock photos banks that “Black family is also family” (Ferreira, 2017). The success of the campaign, along with other types of pressure and market reactions, which included vertical stock photos banks focused on black people, was one of the factors that came to advance the filter options in providers like Shutterstock (2017).

In an interesting typology of algorithm audits proposed by Sandvig and associates (2014), the project *#BuscaPorIgualdade* would be an example of what is called *Sock Puppet Audit*, because it simulates the behavior of users and reflects critically about the patterns of the results. Another accessible approach to non-technical investigations is the *Non-Invasive User Audit*, through of which traditional research methods from social sciences are applied, such as *surveys*, interviews or systematic observation. In a way, the initiatives coming from journalists and the activists’ campaigns fit into these categories, in carrying out “non-invasive selection of information about normal interactions from users on a platform” (Sandvig et al., 2014, p.11).

To emphasize the legitimacy of those approaches is especially relevant in order to give the appropriate importance to the role of the possible harmful impacts, regardless of explicit intentionality by the developers of the systems or of the technically measurable fragility of the codes. Thus, combating harmful impacts from algorithmic systems do not only involve the aspects seen strictly as technical or deriving from computing – neither it involves solely programmers and engineers. The audit concept, therefore, can be expanded so we can also think how the visibility of the incorporated dynamics in these systems generate arguments in order to “promote campaigns and activities which incorporate an accessible and objective language, but with impact and relevance for the citizens in general” (Nunes, 2020).

Ziv Epstein’s and associates research alerts about the “knowledge gap about artificial intelligence “ as far as the “number of unique AI systems grows faster than the number of studies that characterize these systems’ behavior” (2018, p.1). Analyzing more than 7 thousand documents published in the major conference *Neural Information Processing Systems* from 1987 to 2017, the authors discovered that there are approximately 10 times more propositions of new computing models than studies of existing models, in a growing gap.

The tendency is also identified by Pablo Nunes, coordinator of Rede Observatório da Segurança, calling out attention for the mismatch between the “reflection of the effects and the efficiency of certain implementations of algorithms with the number of projects and applications of these same technologies that are already in development” (Nunes, 2020) when we address the implementation of facial recognition for public security.

The interdisciplinary dialogue for the development of algorithmic systems and its positive applications of artificial intelligence can engender ways in which the computing approaches can be part of the solution. Rediet Abebe points some directions in how computing can be used as diagnostic and rebuttal. Abebe defends that the studies about risk algorithmic scores and criminal and behavior prediction, for example, show insurmountable limitations for the use of AI in these cases, therefore

no matter what algorithm is employed, any way of assigning risk estimates to two groups of differing base rates will necessarily produce a specific kind of disparity in outcomes between the groups; we cannot eliminate the problem through a better choice of algorithm. Formal results of this type can expose the limitations of an entire category of approaches — in this case, the assignment of numerical risks to individuals. (Abebe, 2019, p. 190).

It could also be possible to, paradoxically, formalize the problems within the algorithmic systems to the point of exposing the networks of delegations incorporated in the technologies. The computer, or the algorithmic system, becomes *synecdoche*, a discourse and critique tool about the society representing it as a part of a whole and “can offer us a tractable focus through which to notice anew, and bring renewed attention to, old problems” (Abebe, p.192).

Abebe’s proposition goes in the direction of what Charô Nunes highlighted previously in writing about the computing algorithms in overlap to the algorithms of the society understood as a “group of social, economical, ideological and even semiotic rules that are a result of the disputes and all sort of interaction between many segments of the population” (Nunes, 2018, para. 7). Paradoxically, the audits of algorithmic systems from racial critiques, even though are few in relation to numbers of disastrous implementations, are transformed in possibilities of “auditing” the same dynamics proposed by the hegemonic society elects as the central to gain the status of automatized and opaque reproduction.

“Fuck the algorithm”: Youth protests

The COVID-19 pandemic became an unexpected dream for the corporations and *startups* committed with collecting data for more layers of life on behalf of offering systems of algorithmic management. The *contact tracing* as a tool to control the transmission was rapidly co-opted by agents representing state violence, that began to use the same lexicon in order to monitor activists’ actions (Castro, 2020).

But the suspension of essential activities like basic education generated new challenges in a society that promotes competition for public resources – to manage the challenges, the technocentrism was the mistaken solution chosen around the world. In the United Kingdom an algorithmic system was implemented to attribute scores to the students about to try openings at universities, since the normality of the school year was compromised. The data that fed the system was not only based on previous performance of the students, as well as on a ranking established by the teachers about which grade they thought the students could reach at the end of the period, and also the historical performance of the school. As it was expected, the private schools were benefited – in those, the number of maximum grades went up about 5%, the double of the average of the historic difference (Yasin, 2020).

The system, therefore, favored the elites as well as the class determinism and geography in linking the school to the score – in addition to adding a discriminatory variable in the arbitrary attribution of grades by the teachers. The students protested with explicit mottoes, literally “*Fuck the algorithm!*” in front of the Department of Education, pressing to suspend the method. Despite the suspension, however, the perceived damages of the selection processes already in motion should have been adjusted through new appeals by the schools, generating barriers to the harmed students (Katwala, 2020).

The juvenile sincerity of the screams of protest rang globally and one of the posters showed the refusal to forfeit self-determination, saying: “the algorithm doesn’t know who I am”. One of the first organized mobilizations organized in a public space against the imprecision of an algorithm system add up to others aimed to pressure public institutions and corporations to not serve certain ends.

Mobilizations and leaks of internal information by employees possess a long history of impact, including globally. One of the historical peaks was the mobilization of American employees against the colluding of Polaroid with the apartheid regime in South Africa. The company secretly sold supplies for photographs used on the abject “passes” used to identify South

Africans in terms of race, ethnicity and places where they were able to circulate. From the end of the 1960's to the following years, employees of the company got organized as the “*Polaroid Revolutionary Workers Movement*”, and they advocated in favor of the termination of the contracts of the company with the racist regimen, for the public and global announcement of this refusal to make business in the country as long as apartheid was happening and for the company to contribute with the efforts in favor of the African liberation (Morgan, 2006).

Initially, Polaroid tried to hide ties with apartheid and persecuted employees who denounced the problems. Only after many years of fight and resulting coverage from the press, the company ceased the sales. However, this effort inspired the debate, close to other similar groups from the same time, which established a base for the proposition, in 1977, of the “Sullivan Principles”, a set of principles for social responsibility from companies (Alexis, 2010), referenced to this day.

During the 2019 annual Amazon Web Services conference, in New York, hundreds of citizens protested requesting that the corporation stop offering services to harmful organizations from the American government like ICE (US Immigration and Custom Enforcement), responsible for persecuting immigrants and for the implementation of concentration camps (Paul, 2019). The protest adds to a long history of mobilizations against the e-commerce and AI giant, related to its impact of gentrification (Bradshaw, 2020), local business disruption and *dumping*, besides the growing exploitation of employees and the precarious work conditions (Golledge, 2019).

During the intensification of the waves of protests against racist police violence on the United States around 2020, part of the employees from the major technology and artificial intelligence companies signed petitions in order to pressure their employers. At Google, the internal petition “*No Police Contracts*” argues that “saying Black Lives Matter is not enough, we need to show it in our thinking, in our words and in our actions” (Elias, 2020, para. 3). Google, IBM, Amazon and others reacted to the demonstrations

and regulatory pressures (Wiewiórowski, 2020) implementing temporary moratoria to sales of facial recognition to certain uses, but through elusive lexicons like the IBM delimitation on terms for applications for “mass surveillance, racial profiling, violations of basic human rights and freedoms” (Krishna, 2020, para. 9) or the short moratorium proposed by Amazon until new regulatory ethical standards are implemented by the Congress, while it lobbies to influence such rules (Matsakis, 2020).

However, if the public protests made by citizens and consumers generate financial damage to companies with valuable brands which are partially dependent on the final consumer, like Amazon and Google, the same does not occur with gigantic corporations totally directed to contracts with governments and the financial market. It’s the case of Palantir, that offers services of biometric management to persecute migrants and technologies to police forces and it’s the target of public mobilizations (Chan, 2019).

The technological elite from Silicon Valley “suppress the questionings about racism and discrimination, even when the products of the digital elites are infused with markers of race, class and gender” (Noble & Roberts, p.45). Neoliberalism and post racial technocentric myths intentionally make it harder, however, to achieve honest comprehension of the racialization of algorithmic technologies because such comprehension would be diametrically opposed to their business models, which are based in carceral imaginaries and economical disparity.

For example, Amazon developed patents of “intelligent bracelets” to track employees and the movement of their hands under auspices of efficiency, to improve alongside *big data* the work logistics and the movement of assets, reifying the employees for increased levels of productivity (Yeginsu, 2018). Subsequently, the infrastructure has been shamelessly applied to analyze, through heat maps of movement and “emotional analysis” of the employees, which stores from Whole Foods – retail company bought by Amazon in 2017 – would be under risk of unionization (Peter, 2020). Specific job positions for intelligence analysts are being molded specifically to fight worker

organization with the support of data, algorithms and new mechanisms of biometric surveillance (Holmes, 2020).

In this way, the same algorithmic management of strife inside the corporations is a tendency applied by capitalist groups that move forward with the implementation of new mechanisms of control and surveillance of their teams – reminding them of the inherent limits to mobilization directed to values of corporation responsibility.

Resistance through re-inventions

The record or erasure of inventions and technologies is a sociopolitical and historical process employed to privilege Eurocentric conceptions of scientific progress for centuries. Reactions are undertaken from the revolutionary independent research about the kemetetic civilization by Diop (Njeri & Ribeiro, 2020) to contemporary redemptions of the history of African technologies (Machado & Loras, 2017), and valuable studies about the violent appropriation of technologies (Cunha Júnior, 2010) during the slavery at the Black Atlantic. The technologist Ramon Vilarino alerts us that, despite this rich history being hidden, the Brazilian elites do not encourage the “creation of truly local and contextualized technology, which ends by frequently producing caricatures or bad thought adaptations” (Vilarino, 2020) of what is made in the technology centers of the global North.

In contemporary times, Afrodiasporic populations in countries like Brazil fight against the cumulative disadvantages in hostile environments molded by and for white supremacy, but, notwithstanding, developed many strategies of technological innovation. A typology offered by Rayvon Fouché offers a special lens to think what he calls “black vernacular technological creativity” (Fouché, 2006). The term aims to conceptualize the way in which African American inventors adapted, reinvented, or created technologies to their specific realities – despite of the constant erasure of their authorships or even the underestimating when it comes from peripheral regions.

The three categories proposed by Fouché are: a) redeploying, the process in which the material and symbolic power of technology is reinterpreted, but maintains its use and the traditional physical shape; b) reconceiving, the active redefinition of a technology in a way that it transgresses the function and/or main significance; e c) re-creation, like the redesign and production of a new material artifact after an existing shape and/or function was previously rejected (Fouché, 2006, p.658).

The non hegemonical history of the internet has many examples (McIlwain, 2019) of technological creativity initiatives, created by minority groups, like the Black populations. A software created that was especially interesting was the *Blackbird*, launched in 2008 as an experimental browser, stemming from *Mozilla Firefox*, an open code browser. Directed to African Americans, there were some specifications that rejected interstitial whiteness present on the browsers of the time, even if they presented themselves as neutral. Through its characteristics, two could be highlighted (Brock, 2020): areas of content recommendation produced by African American people, selected through curatorship; and deliberate spotlighting of fund raising for social purposes, like the initiative *Give Back*. Its creators and the development community already rejected the idea of a browser that would only function as a window to the cyberspace, not forgetting of interacting with the physical and social world beyond the online and not necessarily for the market.

Rejecting false neutralities of technologies, databases and representations is a continuous effort against the oppressions and limitations imposed to minority groups. Regarding image banks, specifically, initiatives like *Nappy*¹ or *Young, Gifted & Black*² aim to fight negative representations and fill positive gaps in providing photography banks focused in representing Black people. *Nappy*'s website explains how the modality of license and distribution works, *Creative Commons*, allowing the free use, but going beyond and saying that “we encourage it. The more you use it, the more we'll help to better the representation of Black and brown people in the media”.

1. <https://nappy.co>

2. <https://ygb.black>.

If we take a look in how generalist search engines or professional stock photo banks produce visibility, invisibility and stereotypes that penalize mainly Black women, initiatives like these respond to what Patricia Hill Collins evokes in calling for the centrality of self-determination of the image representation as well as the epistemic ones by black women in order to fight the controlling images (Collins, 2002).

Addressing the issues with the transformation of *selfies* made with social media image filters that whiten faces and promote Eurocentric standards of beauty, the Brazilian designer Joyce Gomes created the project *Black Beauty Filters*. After gathering information through qualitative and active listening of lived experiences, the designer produced filters with the augmented reality framework *Spark AR*. Beyond the filters she created for herself, she points out the importance of the project in developing a racial aesthetic literacy, in “instructing the content creator, Black or not, into having a more dedicated view of issues that concern blackness and the filter universe”, encouraging the decolonization of knowledge and the interdisciplinary creation by many groups” (Gomes, 2020).

Inverting the usual gaze of crime maps, the project *White Collar Crime Zones* (Lavigne, Clifton & Tseng, 2017) is a poignant critical parody. The developers produced a crime map³ and a risk prediction system concerned to “white collar crimes”. The project differs from the most famous spatial models of crime prediction because it does not focus in “street crimes” like drug trafficking, theft and vandalism, but in financial crimes of big sums and impact, which usually result in few penalties. The developers used data from financial regulatory institutions and crossed the information with legal drugs data, such as alcohol, besides the density of organizations that evade taxes.

Beyond that, the *White Collar Crime Zones* also constructs the average face of the criminal on its base, from the gathering and computing processing of the similarities between the face photographs of 7.000 executives from financial corporations, extracted from LinkedIn. The pictures of this

3. <https://whitecollar.thenewinquiry.com/>

“average” criminal compose, in navigating the map, variations of a prototypical face of a young white male. Ruha Benjamin argues that, in deliberately and creatively questioning the *status quo* of predictive technology, “analysts can better understand and expose the many forms of discrimination embedded in and enabled by technology” (Benjamin, 2019, p.197).

Training new imaginaries

Maybe one of the most controversial paths to overcome the damage of algorithm discrimination is the promotion of demographical diversity of the ones who develop the technologies, like computing scientists, engineers and developers. The majority of the professional positions of big apparent impact are employed by the oligopoly of “big tech”, which molds global technology.

Shortly after the *BlackLivesMatter* protests in May and June of 2020 which gained global scale after the murder of George Floyd, big USA technology corporations were pressured to act on structural racism on the society and on the industry. The main ones promised, with a lot of fuss, dozens of millions of dollars to black community initiatives and to promote diversity. From 61 thousand of dollars from Dell to 209 million from Microsoft, the numbers caused an impression due to the difficulty of glimpsing its small significance in the billion dollar scale of the big American technology corporations.

However, a journalist made the interesting comparison of the donations, stemming from the question: “If the companies were people, how much money did they donate?” (Peters, 2020). Comparing with the average annual income of 63 thousand dollars in the USA, it shows that the equivalent of the Microsoft donation would be 99 dollars while Dell’s only 4 cents.

If the technochauvinism and its resulting damages favor and are boosted by the platforms and AI oligopolies, initiatives that generate new outlooks to the learning of programming, technology, and digital safety deserve doubled praise. In Brazil, the connections between Black women around initiatives that opened paths (Barros, 2019) such as the Geledés Instituto da Mulher Negra website and the group Blogueiras Negras promote epistemic

constructions (Barros, 2020) about digital care which, in the words of Larissa Santiago, generate “philosophical and practical changes in relation to the use of technologies and tools of information and communication” (Santiago, 2020), incorporating other ways of doing and constructing.

Director of Olabi and founder of PretaLab, Silvana Bahia led an innovative data gathering project about the presence of Black and indigenous women in privileged fields of innovation and technology, defending that “the lack of representation is a problem not only for the ecosystem of technology and innovation, but also for the human rights and freedom of expression” (Pretalab, 2020). The data shows how the beginning of the contact with the area was predominantly by informal means, and that the relation with activism was the third type of motivation for their insertion in the practices of technological development.

Education initiatives like PretaLab aim to reinterpret the technologies from teaching of programming languages and *maker* cultures, as well as reflecting about the learning of the bases of social pyramids. Many innovations rise from the concreteness of the everyday life and Silvana Bahia reminds us that technology is a

big umbrella, and for us it was always important mixing analog technology, *low tech*, with the *high tech*, because we tend to think that this is an effective way of making people understand the importance of it and being able to look in a more critical way

Initiatives focused on horizontal education, sharing of knowledges and mutual emerging support about technology such as Minas Programam, Conexão Malunga, Kilombotech, Perifacode, AqualtuneLab, Tecnogueto, Afropython, Afrotech, Quebradev, InspirAda and others, arise from groups that look into producing not only formal knowledge, but also alternative narratives.

Decenter commercial technology of the discourse about programming skills is something that Bárbara Paes, co-founder of *Minas Programam*, points out when talking about how the imagination of the students about

the *hackatons*, learning experiences and collective projects promoted by the group are crossed by a duality. Some students, while they train for technology-focused industries, from startups to Silicon Valley, also collectively de-construct the usage of the learning, emerging transforming knowledges. Between the collaborative ideas during the learning process in programing, Bárbara Paes points out the inclusion of conversations about

what you can do with it, such as the learning can be one more tool for you, not only in your professional and work life, but also thinking on your community, thinking about the people that are around you, how this can be useful for *you*

In the co-organization of Perifacode, the software engineer Carla Vieira approaches this issue in a similar way, in sharing the polysemy of combating the slants in technology, adapting the term to a positive and transformative aspect in technology

when there are many different people, with their different points of view, like the world is, it makes more sense [...]. What is created will be inclusive, will represent the world as it is, with diversity: they are not only not represented in technology, as well as in other areas.

To the developer Roselma Mendes, this interdisciplinary connection should be promoted, because she believes that the importance of the digital technologies in contemporary society makes that the arising problems do not talk only about technology. They are also meaningful on “how we approach our work [...] I believe that the multidisciplinary and the inclusion/visibility are complementary” in companies that develop software. In some way, we can connect such perceptions and initiatives to a promotion of racial literacy in technologies, that aims to exhibit the false neutrality that only replicates oppressions and erases alternative imaginaries.

Jessie Daniels, Mutale Nkonde and Darakhshan Mir propose the advance of racial literacy in technology in a multisectoral way, pointing out the limitations of the propositions of big corporations. Three pillars are proposed

by the authors to leave innocuous patterns of action, including “an intellectual understanding of how structural racism operates in algorithms, social media platforms, and technologies not yet developed” (Nkonde & Mir, 2019, p.4).

More than promoting diversity as a token or in an isolated way, a transformative commitment is needed. The emphasis in fighting the damages due to racism goes through uncentering purely technical aspects from discourse (Gangadharan & Niklas, 2019) opens up possibilities for addressing different modalities of discrimination and algorithmic harm.

Regulation beyond the ethical principles

The abbreviation *FAT* or *FACcT* to account for the effort to avoid algorithm harms stems from the triad *Fairness, Accountability and Transparency*, and became a global synonym of the debate about ethics in algorithmic systems going on in the computation industry and in communities around machine learning and neural networks. Critiques about the *FACcT* approach involve specially the tendency to delimit the problem of algorithmic harm as a matter of coding or management. A common critique question the necessity of developing new concepts for prerogatives of respect to human rights, considering that framing discriminatory impacts as new can erase already known political and racial aspects of technologies.

An exploratory survey (Floridi & Cowls, 2019) studied the consensus between propositions by international initiatives about principles for artificial intelligence. It organized them into five principles: beneficence, non-maleficence, autonomy, justice and explicability. The first four are already discussed and applied frequently in bioethics, reminding how interdisciplinarity can fight tendency of perceiving problems in digital artifacts as something completely new. Regarding autonomy, Floridi and Cowls defend that humans should retain the power of deciding which types of decisions are made, having the possibility of intervention when needed, and, at last, collectively deciding in which cases the loss of control of the decision process is worth it in terms

of the benefits compared to the costs or possible damages (Floridi & Cowls, 2019).

It's more about who has the power of “classify, to determine the repercussions / policies associated thereof and their relation to historical and accumulated injustice” (Abdurahman, 2019, para. 8) in the words of J. Khadijah Abdurahman criticizing the *FaccT* approach. Catherine D’Ignazio e Lauren F. Klein agrees, in offering an alternative set of guiding concepts to the field, advocating for a transition of “data ethics” to “data justice”. The goal is to dislocate the source from the individual problems and technical systems to the comprehension of the power relations as well as fighting them (D’Ignazio & Klein, 2020), like we can see on the following table:

Table 1: From Data Ethics to Data Justice

Concepts that secure power	Concepts that challenge power
Ethics	Justice
Bias	Oppression
Fairness	Co-liberation
Transparency	Reflexivity
Understanding algorithms	Understanding history, culture, and context

Source: D’Ignazio & Klein, 2020, p.60

The proposition of this shift means to embrace concepts that dialogue with the legacy of collective organization, inters-sectional feminism and critical thinking, rejecting the idea that radically new “ethical” principles would be necessary to grasp problems based on *big data* and artificial intelligence.

In this way, the idea of algorithmic inexplicability should not be acceptable in systems that have relevant harmful potential to individuals or groups. Moving forward with the implementation of an algorithm system with inexplicable decisions means making the possible harm acceptable – which

implements computationally the hierarchies of society based on race, gender, class and others. For Abeba Birhane and Jelle van Dijk, who analyze how the debate about “robots rights” have been used as a diversion tactic about the impacts of artificial intelligence,

One of the pressing issues in this day and age is that ‘intelligent’ machines are increasingly used in sustaining forms of oppression. We do not ‘blame’ the machines (they can take no blame), nor do we say machines must bear ‘responsibility’ [15], precisely because this would relieve those actually responsible from their duties. (Birhane & van Dijk, 2020, p.6)

In this sense, the principle of explicability can be seen as an essential prerogative in fighting algorithmic racism if it’s seen as pertinent not only to code lines, but also to the processes of planning, implementing and to whom do the systems benefit or exclude.

In many cases, like in the implementation of biometric surveillance for public security, the predominance of social injustice is evident in the production of imageries – carceral and racialized – in use of the artifact and notions of how to explain the functioning of a system. The search for fairness should take into consideration the patterns of action and conceptualization around the problem that supposedly should be solved by the algorithmic system (Hanna *et al.*, 2020). Sérgio Amadeu da Silveira highlights the contradictions between opaqueness and the implementation by the State in relevant areas, questioning the “convenience and legitimacy of the use by the State of algorithmic systems that not even its employers could explain all its operations” (Silveira, 2019, p.13).

In a report to the United Nations, E. Tendayi Achiume proposes an analysis based on human rights about racial discrimination in digital emerging technologies. For Achiume, “the heart of the issue is a political, social and economic one, not solely a technological or mathematical problem” (Achiume, 2020, p.15) and, therefore, the States should establish legal commitment to perform ample scrutiny of the discriminatory possibilities against minorities.

Some of recommendations to the States to fight racial discrimination in designing and using emerging digital technologies are:

- States should adopt immediate and effective measures, particularly in the fields of teaching, education, culture and information, with a view to combating prejudices which lead to racial discrimination;
- Prevent and eliminate racial discrimination in the design and use of emerging digital technologies require addressing the “diversity crisis”;
- Make racial equality and non-discrimination human rights impact assessments a prerequisite for the adoption of systems based on such technologies by public authorities;
- States should ensure transparency and accountability for public sector use of emerging digital technologies, and enable independent analysis and oversight, including by only using systems that are auditable;
- Frameworks and guidelines developed to provide flexible, practical and effective regulation and governance of emerging digital technologies are grounded in legally binding international human rights principles. (Achiame, 2020, p.16-17)

The recommendations also re-affirm that the scope of obligations should involve a perspective based in intersectional analysis that can be applied to multiple and overlapping forms of discrimination. The relevance of the debate and the dispute in international organizations around the obligations may impact positively as much in the practices of corporations as in the participation of the civil society in the defense of human rights.

As a society, therefore, we should ask ourselves – and act – about which technologies and public policies we want to include in our possible futures and what we consider as goals and desirable results (Constanza-Schock, 2020). It’s already possible to recognize that some algorithmic systems can “function perfectly, with full enrollment, complete transparency, seamless integration and exacting discriminatory power” (Abebe, 2019, p.187),

however. It grows, as a consequence, the perception that some emerging algorithmic technologies can – and should - be the objects of collective rejection.

The *status* of full humanity is multi-faceted and connected to countless everyday processes that restrain or amplify the opportunities, barriers and possibilities of concrete, intellectual and psychological action of individuals, according to their socially perceived affiliation to a racial group. In this overview, the algorithmic racism is a phenomenon directly linked to the problem of double opacity – the way in which hegemonic groups search to present the idea of “neutrality” in technology as much as to dissipate the debate about racism and white supremacy in the West. Studying, debating and acting about the relations between technology and race, thus, becomes doubly challenging in societies dictated by technochauvinism (Broussard, 2018), the racial democracy myth (Nascimento, 2016) or post-racialism (Bonilla-Silva, 2015).

Sueli Carneiro reminds us, though, that the social idea of race has a double sense when evoked as an epistemological tool because of its social transformation. In one side, as a

methodological instrument, intends to comprehend the unequal relations between the differing human groups, more specifically the inequalities of treatment and perceived social conditions between black and whites in Brazil. As a discursive practice, the studies inspired in it aim to modify the social relationships that produce the discriminations and racial asymmetry. (Carneiro, 2005, p.52)

The dehumanization, the recovery or the maintenance of full humanity of the individuals go through understanding the counter narratives at play, as much as history as collective projects. We can connect the anti-racist thought about the technology not only as critique, but also for new emergencies (Benjamin, 2019) that have as prerogative to reject oppressive potential.

If fatalism is a colonial tool for domination (Lapa, 2018), to think horizons of alternative possibilities, to unveil the naturalization of inequalities and

strengthen paths that address local and global impacts of algorithmic racism is a task that fortifies alternative imagery for the common good of all humanity.

References

- Abdurahman, J. K. (2019, Feb 25). “FAT* Be Wilin”, @blacksirenradio. <https://medium.com/@blacksirenradio/fat-be-wilin-deb56bf92539>.
- Abebe, R. T. (2019). *Designing algorithms for social good*. [Doctoral Dissertation, Cornell University].
- Achiume, E. T. (2020). *Racial Discrimination and Emerging Digital Technologies: A Human Rights Analysis*, Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance to the United Nations Human Rights Council.
- Alexis, G. Y. (2010). “Global Sullivan Principles”, in: Nevin Cohen; Paul Robbins (orgs.), *Green Business: An A-to-Z Guide*, Thousand Oaks (EUA): SAGE Publications.
- Bahia, S. (2019). Interview given to the author via videoconference in September 2019.
- Barros, T. N. (2020). “Estamos em Marcha! Escrevivendo, Agindo e quebrando códigos”, in: Tarcízio Silva, *Comunidades, Algoritmos e Ativismos Digitais: olhares afrodiaspóricos*, São Paulo: LiteraRUA.
- Barros, Z. (2019). “Feminismo Negro e Internet”. https://www.academia.edu/1497162/Feminismo_negro_na_Internet acesso em 23 dez. 2019.
- Benjamin, R. (2019). *Race after Technology: Abolitionist Tools for the New Jim Code*, Cambridge (UK): Polity Press.
- Birhane, A., van Dijk, J., (2020). “Robot Rights? Let’s Talk about HumanWelfare Instead”, *AIES ’20*, February 7–8, 2020, New York, NY, USA.
- Bonilla-Silva, E. (2015). “The structure of racism in color-blind, “post-racial” America.” *American Behavioral Scientists*, vol. 59, 11..

- Bradshaw, K. (2017, Mar 31). “Vigil in East Palo Alto protests Amazon, Facebook policies”, *The Almanac*. <https://www.almanacnews.com/news/2017/03/31/vigil-in-east-palo-alto-protests-amazon-facebook-policies>.
- Brock, A. (2020). *Distributed Blackness: African American Cybercultures*, Nova Iorque: NYU Press.
- Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*, MIT Press.
- Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91).
- Carneiro, A. S. (2005). *A construção do outro como não-ser como fundamento do ser*, Tese de Doutorado do Programa de Pós-Graduação em Educação da Universidade de São Paulo, São Paulo – SP.
- Castro, S. (2020). “Surveilling Racialized Bodies”, *NACLA Report on the Americas*, vol. 52, n.3, 2020.
- Castro, S. (2020). “Surveilling Racialized Bodies”, *NACLA Report on the Americas*, vol. 52, n.3.
- Chan, R. (2019, Aug 17). “Protesters blocked Palantir’s cafeteria to pressure the \$20 billion big data company to drop its contracts with ICE”, *Business Insider*. <https://www.businessinsider.com/palantir-protest-palo-alto-activists-ice-contracts-2019-8>.
- Collins, P. H. (2002). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*, Londres: Routledge.
- Constanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need*, Cambridge (EUA), MIT Press.
- Cunha Júnior, H. (2010). *Tecnologia Africana na Formação Brasileira*, Rio de Janeiro: CEAP.
- D’Ignazio, C. D., & Klein, L.F. (2020). *Data Feminism*, Cambridge (USA): The MIT Press.
- Daniels, J., Nkonde, M., & Mir, D. (2019). “Advancing Racial Literacy in Tech”, Data & Society Fellowship Program Report.

- Elias, J. (2020, Jun 22). “Google employees petition company to cancel police contracts”, *CNBC*. <https://www.cnn.com/2020/06/22/google-employees-petition-company-to-cancel-police-contracts.html>.
- Epstein, Z et al. (2018). Closing the AI knowledge gap, *arXiv preprint arXiv:1803.07233*.
- Epstein, Ziv et al (2018). Closing the AI knowledge gap, *arXiv preprint arXiv:1803.07233*.
- Ferreira, M. (2017). “Projeto #BuscaPorIgualdade cobra representatividade negra dos bancos de imagens”, *CEERT*, 02 mai. 2017, available at <https://ceert.org.br/noticias/comunicacao-midia-internet/16909/projeto-buscaporigualdade-cobra-representatividade-negra-dos-bancos-de-imagens>.
- Ferreira, Matheus (2017, May 02). “Projeto #BuscaPorIgualdade cobra representatividade negra dos bancos de imagens”, *CEERT*. <https://ceert.org.br/noticias/comunicacao-midia-internet/16909/projeto-buscaporigualdade-cobra-representatividade-negra-dos-bancos-de-imagens>.
- Floridi, L, & Cowls, J. (2019). “A Unified Framework of Five Principles for AI in Society”, *Harvard Data Science Review*, vol. 1, n.1.
- Fouché, R. (2006). “Say It Loud, I’m Black and I’m Proud: African Americans, American Artifactual Culture, and Black Vernacular Technological Creativity”, *American Quarterly*, vol. 58, n.3.
- Gangadharan, S.P.; Niklas, J. (2019). “Decentering technology in discourse on discrimination”, *Information Communication & Society*, vol. 22, n.7.
- Golledge, R. (2017, Dec 14). “Protest at Amazon Rugeley over ‘hellish’ working conditions”, *Express & Star*. <https://www.expressandstar.com/news/local-hubs/staffordshire/rugeley/2017/12/14/protest-at-amazons-rugeley-warehouse-over-hellish-working-conditions/>.
- Gomes, J. (2020). Interview given to the author via email.
- Hanna, A., Denton, E., Smart, A., & Smith-Loud, J. (2020). “Towards a Critical Race Methodology in Algorithmic Fairness”, *Conference on Fairness, Accountability, and Transparency (FAT* ’20)*, Barcelona (Espanha).

- Holmes, A. (2020, Sep 01). “Amazon posted — and then deleted — a job listing for an ‘intelligence analyst’ to monitor workers’ efforts to unionize”, *Business Insider*. <https://www.businessinsider.com/amazon-posts-deletes-job-listing-intelligence-analyst-spy-worker-union-2020-9>.
- Kari Paul, “Protesters demand Amazon break ties with Ice and Homeland Security”, *The Guardian*, 11 jul. 2019, disponível em <<https://www.theguardian.com/us-news/2019/jul/11/amazon-ice-protest-immigrant-tech>>, acesso em: 20 set. 2019.
- Katwala, A. (2020, May 15). “An Algorithm Determined UK Students’ Grades. Chaos Ensued”, *Wired*. <https://www.wired.com/story/an-algorithm-determined-uk-students-grades-chaos-ensued/>.
- Krishna, A. (2020, Jun 8). “IBM CEO’s Letter to Congress on Racial Justice Reform”, *IBM THINKPolicy Blog*. <https://www.ibm.com/blogs/policy/facial-recognition-sunset-racial-justice-reforms>.
- Lapa, R. S. (2018). “O fatalismo como estratégia colonial”, *Revista Epistemologias do Sul*, v. 2.
- Lavigne, S., Clifton, B., & Tseng F. (2017). “Predicting Financial Crime: Augmenting the Predictive Policing Arsenal”, *arXiv:1704.07826*.
- Machado, C. E. D., & Loras, A. (2017). *Gênios da Humanidade: Ciência, Tecnologia e Inovação Africana e Afrodescendente*, São Paulo: DBA.
- Matsakis, L (2020 Jun 10). “Amazon Won’t Let Police Use Its Facial-Recognition Tech for One Year”, *Wired*. <https://www.wired.com/story/amazon-facial-recognition-police-one-year-ban-rekognition>.
- McIlwain, C. D. (2019). *Black Software: The Internet and Racial Justice, from the AfroNet to Black Lives Matter*, Oxford University Press, USA.
- Mendes, R. (2020). Interview given to the author.
- Morgan, E. (2006). “The world is watching: Polaroid and South Africa”, *Enterprise & Society*, vol. 7, n.3.
- Nascimento, A. (2016). *O genocídio do negro brasileiro: processo de um racismo mascarado*, São Paulo: Editora Perspectiva.
- Nelson, A., Tu, T. L. N., & Hines, A. H. (orgs.) (2001). *Technicolor: Race, technology, and everyday life*. Nova Iorque (EUA): NYU Press.

- Njeri, A., & Ribeiro, K. (2019). “Mulherismo Africana: práticas na diáspora brasileira”, *Currículo sem Fronteiras*, vol. 19, n.2.
- Noble, S. U., & Roberts, S. (2020). “Elites tecnológicas, meritocracia e mitos pós raciais no Vale do Silício”, *Fronteiras – estudos midiáticos*, vol. 22, n.1.
- Nunes, C. (2018, Jan 08). “O Algoritmo”, *Blogueiras Negras*. <http://blogueirasnegras.org/o-algoritmo>.
- Nunes, Pablo (2020). Interview given to the author via email, May 2020.
- Paes, Bárbara (2019). Interview given to the author, September 2019.
- Peters, J. (2020, Abr 20). “Whole Foods is reportedly using a heat map to track stores at risk of unionization”, *The Verge*. <https://www.theverge.com/2020/4/20/21228324/amazon-whole-foods-unionization-heat-map-union>.
- Peters, J. (2020, Aug 13). “Big Tech Pledged a Billion to Racial Justice, But it was Pocket Change”, *The Verge*. <https://www.theverge.com/21362540/racial-justice-tech-companies-donations-apple-amazon-facebook-google-microsoft>.
- PretaLab (2023). *Website Pretalab*. <https://www.pretalab.com>.
- Raji, I. D., & Buolamwini, J. (2019). “Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products”, in: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 2019.
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). “Auditing algorithms: Research methods for detecting discrimination on internet platforms”, *Data and discrimination: converting critical concerns into productive inquiry*, Seattle (USA).
- Santiago, L. (2020). “Tecnologia Antirracista: a resposta das mulheres negras”, in: FOPIR, *Mapeamento da Mídia Negra no Brasil*. http://fopir.org.br/wp-content/uploads/2020/08/ebook_mapeamento_da_midia_negra-1.pdf.
- Shutterstock. [@Shutterstock]. (2017, June 06). Refine your search further with our new People Filter. *Twitter*. <https://twitter.com/Shutterstock/status/872105821609570304>

- Silveira, S. A. (2019). “Quem governa os algoritmos? A Regulação dos Sistemas Algorítmicos no Setor Público”, *Anais do 43º. Encontro da Associação Nacional de Pós-graduação em Ciências Sociais – Anpocs*.
- Vieira, Carla (2019). Interview given to the author, October 2019.
- Vilarino, Ramon (2020). Interview given to the author via email, July 2020.
- Werneck, J. (2010). Nossos passos vêm de longe! Movimentos de mulheres negras e estratégias políticas contra o sexismo e o racismo. *Revista da Associação Brasileira de Pesquisadores/as Negros/as (ABPN)*, 1(1), 07-17.
- Werneck, Jurema (2009), “Our steps come from afar! Movimentos de mulheres negras e estratégias políticas contra o sexismo e o racismo”, in: Christine Verschuur, *Vents d’Est, vents d’Ouest: Mouvements de femmes et féminismes anticoloniaux*, Geneva (Suíça): Graduate Institute Publications, 2009.
- Wiewiórowski, W. (2020, Feb 21). “AI and Facial Recognition: Challenges and Opportunities”, *European Data Protection Supervisor*. https://edps.europa.eu/press-publications/press-news/blog/ai-and-facial-recognition-challenges-and-opportunities_en.
- Yasin, A. (2020, Aug 17). ““Fuck the Algorithm”; the Rallying Cry Of Our Youth?”, *DigitalDiplomacy*. <https://medium.com/digital-diplomacy/fuck-the-algorithm-the-rallying-cry-of-our-youth-dd2677e190c>.
- Yeginsu, C. (2018, Feb 01). “If Workers Slack Off, the Wristband Will Know. (And Amazon Has a Patent for It.)”, *NY Times*. <https://www.nytimes.com/2018/02/01/technology/amazon-wristband-tracking-privacy.html>.

Esta obra reúne contributos que discutem o papel da Inteligência Artificial e dos algoritmos na sociedade. O debate foi mobilizado por um conjunto de questões, entre elas: Quais os papéis desempenhados pela IA e pelos algoritmos nas sociedades democráticas? Que questões éticas devem ser consideradas na adoção dessas tecnologias por instituições públicas e privadas? Como garantir a aplicação da IA e dos algoritmos sem que discriminações e vieses sejam reforçados? De que modo o jornalismo é impactado pela produção de conteúdos automatizados? Quais as implicações do uso de algoritmos nas notícias? Como a IA contribui para a disseminação e combate à desinformação? A IA e os algoritmos continuarão a fazer parte das práticas comunicativas na sociedade contemporânea. O seu impacto estende-se ao ecossistema informativo, às relações sociais e à qualidade das democracias. Resta aos investigadores problematizar as lógicas algorítmicas e procurar aprofundar o conhecimento sobre as tecnologias emergentes.

Este livro conta com o apoio do MediaTrust.Lab - Laboratório de Media Regionais para a Confiança e Literacia Cívicas, executado no LabCom - Comunicação e Artes, unidade de investigação da Universidade da Beira Interior, participado pela Universidade de Coimbra e financiado pela Fundação para a Ciência e a Tecnologia (PTDC/COM-JOR/3866/2020).